Retrieval practice prevents stress-induced inference impairment by restoring rapid memory reactivation

Jinpeng Guo[1,2], Ruixin Chen[1,2], Qi Zhao[1,2], Xiaojun Sun[1,2*], Wei Liu[1,2*]

1. Key Laboratory of Adolescent Cyberpsychology and Behavior (CCNU), Ministry of Education, Wuhan, China
2. Key Laboratory of Human Development and Mental Health of Hubei Province, School of Psychology, Central China Normal University, Wuhan, China

Correspondence:

Prof. Xiaojun Sun

School of Psychology, Central China Normal University

Address: Nanhu Complex Building, No 152 Luoyu Road, Wuhan. Postal Code: 430079

E-mail: sunxiaojun@ccnu.edu.cn

Dr. Wei Liu

School of Psychology, Central China Normal University

Address: Nanhu Complex Building, No 152 Luoyu Road, Wuhan. Postal Code: 430079

E-mail: weiliu1991@ccnu.edu.cn

Abstract

A hallmark of human memory is its ability to form novel inferences by linking discrete but related events. Our study examined whether acute stress impairs memory inference in humans and assessed the potential of retrieval practice to buffer this effect. Participants were trained on pairs of images, AB and BC, to establish interconnected triads (ABC) with a shared bridge element B. Twenty-four hours later, we induced acute stress in half of the participants and then tested for their capacity to infer the indirect AC associations. Behavioral results indicated that acute stress impaired memory inference, yet targeted retrieval practice of the AB and BC pairs after encoding could prevent the stress-induced impairment in A-C inference. Using multivariate decoding analysis of human electroencephalogram (EEG) recordings, we provided neural evidence that bridge element B is rapidly reactivated during the inferential process, a neural signature that is predictive for subsequent successful inference. Importantly, we showed that stress disrupts this rapid neural reactivation of the brdige element, but retrieval practice can buffer the stress effect and even enhanced the strength of reactivation signals beyond the non-stress condition. Collectively, our findings pinpoint rapid memory reactivation of bridge information as an essential neural mechanism underpinning memory-based inference, which, although susceptible to stress, can be enhanced through retrieval practice. These insights suggest that building robust memory traces could enable subsequent memory inferences to be resilient to stress, highlighting that retrieval practice could sustain normal flexible cognitions under stress.

Keywords: retrieval practice; acute stress; neural reactivation; memory inference; retrieval

Introduction

The ability to make novel inferences from distinct memories (i.e., memory inference) is essential for complex and adaptive behaviors. It is believed that successful inference, at least partly, depends on the retrieval of established memories. However, in many real-world settings (e.g., exam), the ability to retrieve may be negatively influenced by stress[1]. Consequently, this may impair memory inference, negatively impacting flexible behaviors and decision making. In the present study, we examined whether, and how, acute stress impairs the cognitive and neural mechanisms underlying memory inference and further investigated the efficacy of retrieval practice in buffering this behavioral impairment and restoring underlying neural processes.

To study the neural processes underlying memory inference in humans, we used electroencephalogram (EEG) to record neural activity during memory inference task. We adapted the associative inference paradigm[2,3], which involves recombining information across overlapping experiences (i.e., AB, BC) to enable novel inference (i.e., A to C). Previous studies mainly focused on the integration phase (i.e., encoding AB and BC associations), with neuroimaging evidence highlighting prefrontal-hippocampal interaction during retrieval-mediated learning[4,5]. One study distinctly highlighted the different roles of the medial temporal lobe (MTL) and prefrontal cortex (PFC) during integration and inference phases: PFC activation was associated with trial-by-trial inferential success, whereas MTL activity correlated with overall transfer performance across participants[6]. Recent advancements in computational methods of fMRI[7], such as representational similarity analysis and neural decoding analyses, have begun to elucidate how bridging information B

was represented in the brain and reactivated to support subsequent efficient inference[8,9]. Nevertheless, the direct role of bridging information in supporting A-C inference remains elusive, as previous studies lacked neural data analysis during the active online A-C inference task. In our study, we focused on the cognitive and neural mechanisms during A-C inference, postulating that the rapid reactivation of bridging information B is crucial for memory inference. Given the limitations of self-reported memory assessments and fMRI's temporal resolution, we leveraged EEG's superior temporal resolution and employed neural decoding to capture this reactivation of bridge information without interrupting the inferential process. Specifically, we used distinct categories of pictures (i.e., faces and buildings) as bridging information, eliciting unique neural patterns in the human brain[10,11], whereas pictures from another category (objects) represented A and C events. We applied time-resolved neural decoding[12,13] to EEG signals to pinpoint the reactivation of bridging information (i.e., face or building) during the inferential process. Importantly, during inference, only object pictures were visually presented, ensuring we were decoding reactivated internal mnemonic representations during the mental inference.

Previous studies have shown that acute stress impairs episodic retrieval in both animals and humans[14–19] by disrupting prefrontal-hippocampal circuit [17,20–25]. Additionally, under stress, there is a tendency for individuals to transition from flexible, MTL-dependent behaviors to simpler, striatal-dependent habitual behaviors[26]. For example, stressed participants were less likely to found novel shortcuts and more reply on familiar path during the navigation task. This behavioral change was associated with disrupt hippocampal-prefrontal function during prospective navigation planning[27]. Given these neurocognitive

shifts under acute stress, we hypothesized that stress would hinder the flexible use of stored memories (i.e., AB and BC) for making inferences (i.e., A to C).

Given the potential negative impact of stress on memory inference, we investigated methods to protect memory inference against stress. Several approaches have been used to buffer the negative effects of stress on memory[28], including pharmacological[29,30] and brain stimulation approaches[31]. Our focus is on the retrieval practice strategy—the act of taking practice tests—as a highly effective learning approach[32,33]. This strategy enhances the resilience of participants' memory retrieval to stress by reinforcing memory representations before stress induction and retrieval. One study found that retrieval practice not only improve overall memory performance but also prevents the typical stress-related memory impairments observed in in the free recall task[34]. Given that episodic memory retrieval is crucial for making inferences, we suggest that retrieval practice serves as an effective defense against stress in memory inference tasks. The neural mechanisms underlying the protective effects of retrieval practice remain to be fully elucidated; however, its contribution to rapid memory consolidation might offer insights[35]. Recent studies have shown that retrieval practice accelerates memory consolidation, leading to the rapid formation of adaptive cortical memory representations independent of the hippocampus[36–38]. Stress-induced glucocorticoid release primarily impairs MTL functions[39] , rendering these rapidly consolidated cortical representations more resistant to stress. Crucially, to emphasize the protective effects on inference rather than retrieval in the current study, we implemented retrieval practice to strengthen AB and BC associations separately, avoiding direct AC inference practice before inference task under stress. We predict that retrieval practice creates stress-resilient memory

traces, providing robust memory components (i.e., AB and BC) for inference and thus preventing stress-induced impairments in memory inference.

The present study brings the above theoretical framing together (**Figure 1**) to investigate (1) the neural mechanism of memory inference in humans. We proposed that the fast reactivation of neural representation of bridge information B within interconnected triads (ABC) facilitates succesful memory inferences and provided neural evidence from the cross-task multivaraite decoding analyses of human EEG data; (2) the impact of acute stress on this rapid memory reactivation process, potentially impairing memory inferences. We predicted that participants under stress would exhibit impaired inference performance and disrupted neural reactivation of bridge information B during the mental simulation of A-C inference. Leveraging the optimal temporal resolution of EEG data allowed us to pinpoint the specific timeframe in which stress inhibited rapid memory reactivation; (3) the efficacy of retrieval practice in buffering stress-induced memory inference impairments and the neural mechanisms of this protective effect. We investigated whether independent retrieval practice of AB and BC pairs could enhance AC inference and protect it against stress. We further predicted that the protective effect of retrieval practice under stress could be attributed to the restoration of rapid memory reactivation of bridge information.

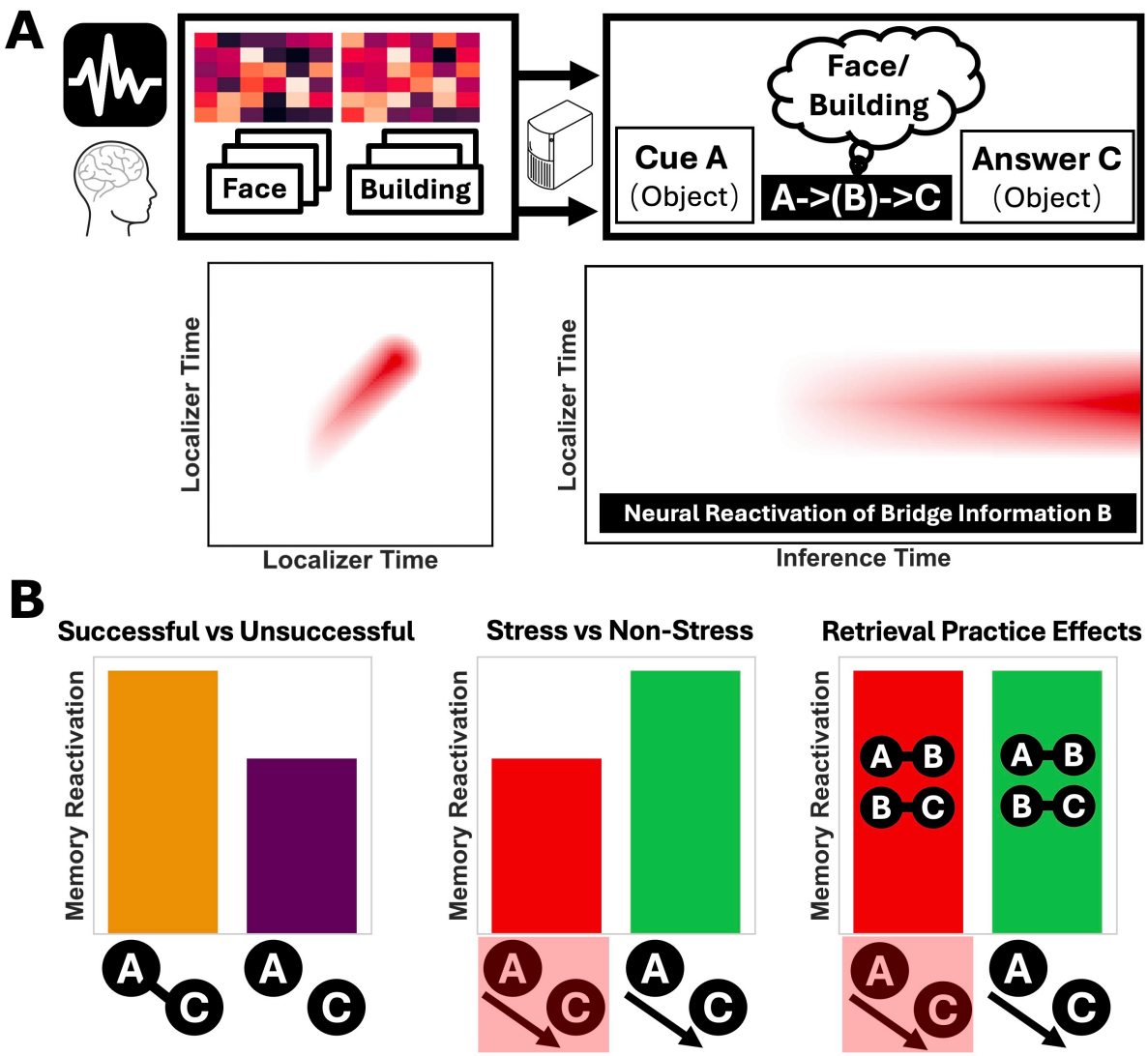

Figure 1 Design for EEG experiment and neural decoding for memory reactivation. **(A)** In the EEG experiment, participants performed an inference task in which they were asked to infer the correct answer-C based on the presented cue-A, using pretrained AB and BC associations. Element B, the "bridge information," could be either a face or a building image. Both the cue-A and the answer-C were pictures of daily-life objects. EEG data were recorded during both the inference and an independent localizer phase. During the localizer phase, participants actively viewed a different set of face and building images. A classifier was trained on category-specific neural response patterns from the localizer phase and then applied to the inference phase. This approach enabled time-series decoding analyses to detect, within individual trials, whether category information of the bridge element-B was reactivated at any moment. **(B)** Using neural decoding to capture the memory reactivation of bridge information-B, we examined whether this neural signature was stronger during successful inference trials (left panel) and weaker under acute stress (middle panel). Additionally, we assessed whether the neural signature could be protected through retrieval practice (right panel).

Results

To investigate how acute stress affects behavioral performance and neural processing of memory inferences, and to explore the potential strategy (i.e., retrieval practice) for mitigating stress-induced impairments, we conducted two independent EEG experiments with healthy young adults of both sexes (total n = 128, $n_{experiment1}$=68, $n_{experiment2}$=60). On Day 1, participants learned associations AB and BC, where each association consisted of pairs of images, forming ABC triads. Critical to our aims, the ABC triads provided participants with an opportunity to leverage associative memories to flexibly infer picture C from Cue A on Day 2. In Experiment 1, we examined how acute stress affects underlying memory reactivation of bridging information and subsequent inference performance by assigning participants at the start of Day 2 to either a stress group or a non-stress control group (**Figure 2**). In Experiment 2, participants underwent the same experimental procedure, but immediately after encoding, they engaged in retrieval practice (i.e., taking memory tests on the AB and BC associations with feedback on correct associations) to enhance AB and BC associative memories. We examined whether retrieval practice could prevent stress-induced inference impairment by modulating underlying memory reactivations. The following analyses were performed and presented in the order: (1) manipulation checks of stress induction at both the psychological and physiological levels; (2) the effect of acute stress on memory inference and how retrieval practice mitigates stress-induced impairment; (3) multivariate EEG decoding of neural reactivation during inference processing; and (4) group comparisons of time-resolved decoding outcomes, focusing on the effects of stress and retrieval practice.

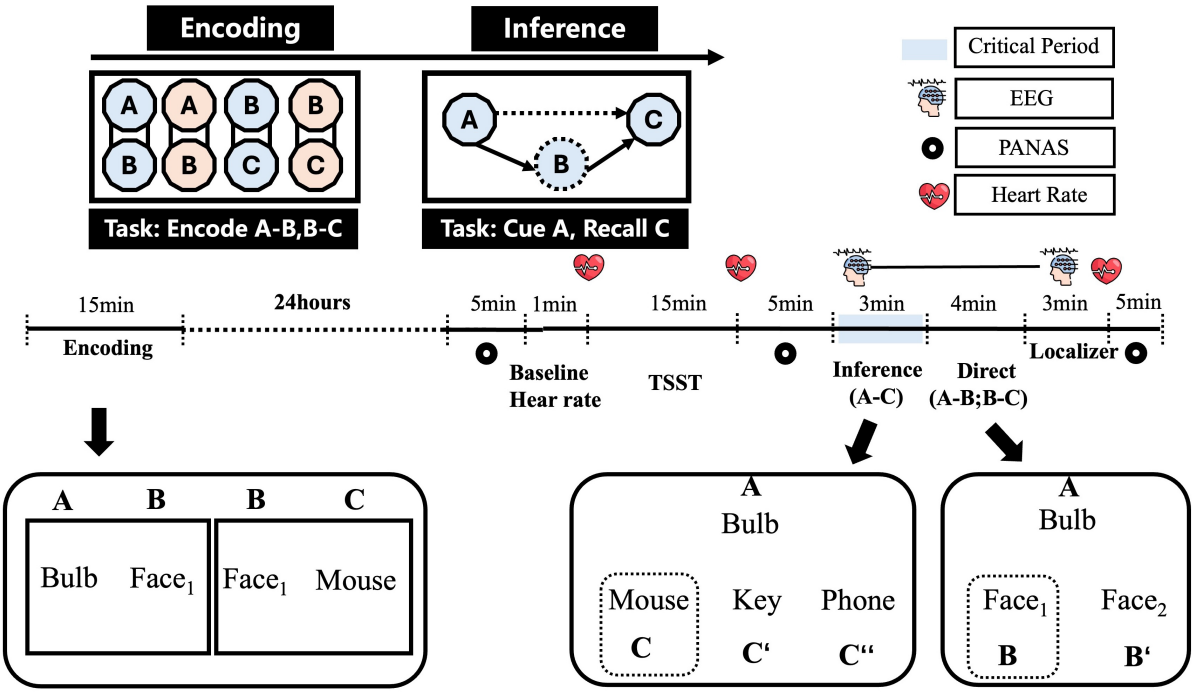

Figure 2: Task Overview. During the encoding phase, participants learned two types of associations: AB and BC, where each association consisted of pairs of images, forming interconnected triads (ABC). During the inference phase, picture A was presented as a cue to recall picture C. Importantly, participants were not informed about the inference requirement during the AB and BC encoding phases. The task was divided into two sessions, separated by a 24-hour interval. On Day 1, the encoding phase lasted approximately 15 minutes, during which two pictures were presented side-by-side on the screen. On Day 2, participants underwent either a stress (i.e., Trier Social Stress Test [TSST]) or a control manipulation, and then finish the memory tasks. Acute stress responses were assessed using heart rate measurements and subjective emotion ratings (i.e., Positive and Negative Affect Schedule [PANAS]). EEG data were recorded during both the inference and localizer phases, with the inference phase being the critical period of our analysis. During the inference phase, Cue-A was presented alongside the correct answer-C and two lure images from other learned ABC triads. The original AB and BC associations were also tested to enable subsequent trial-by-trial analyses of memory inference. Experiment1 and Experiment2 used the same task procedure, with experiment2 included the additional retrieval practice session in Day1, immedialy after the encoding (See **Figure S1** for the procedure of Retrieval Practice phase).

2.1 Manipulation checks of stress induction

In both experiment1 and 2, we evaluated the effectiveness of acute stress induction using both physiological (heart rate in beats per minute, BPM) measures and psychological (self-reported Positive and Negative Affect Schedule, PANAS).

For physiological responses, we monitored heart rate throughout the stress induction procedure both experiments and during the memory inference phase in Experiment 1 and Experiment 2. Heart rate responses were quantified as the difference between the average during the Trier Social Stress Test (TSST) phase and the baseline phase (i.e., before stress induction). Independent t-tests showed significant higher heart rate elevation in the stress condition compared to the control condition in both experiments (exp1: $t(66) = 9.377$, $p < .001$, Cohen's d = 2.274, 95% CI = [8.04,12.40], **Figure 3A**; exp2: $t(58) = 5.643$, $p < .001$, Cohen's d = 1.457, 95% CI = [4.19,8.81], **FigureS2A**).

For psychological stress responses, we analyzed changes in PANAS scores. Differences between the second measurement (after stress induction) and the first measurement (before stress induction) served as indicators of stress responses for both negative and positive affect. Independent samples t-tests were conducted with group (stress vs. control) as the between-subject factor. In both experiments, the stress induction procedure led to significantly increased negative affect (exp1: $t(66) = 2.335$, $p = 0.023$, Cohen's d = 0.566, 95% CI = [0.38,4.86], **Figure 3B**; exp2: $t(58) = 3.023$, $p = 0.004$, Cohen's d = 0.781, 95% CI = [1.05,5.15], **Figure S2B**), and decreased positive affect in the stress group compared to the control group (exp1: $t(66) = -2.155$, $p = 0.034$, Cohen's d = -0.523, 95% CI = [-6.35,-0.24], **Figure 3C**; exp2: $t(58) = -2.149$, $p = 0.036$, Cohen's d = -0.555, 95% CI = [-3.73,-0.13], **Figure S2C**).

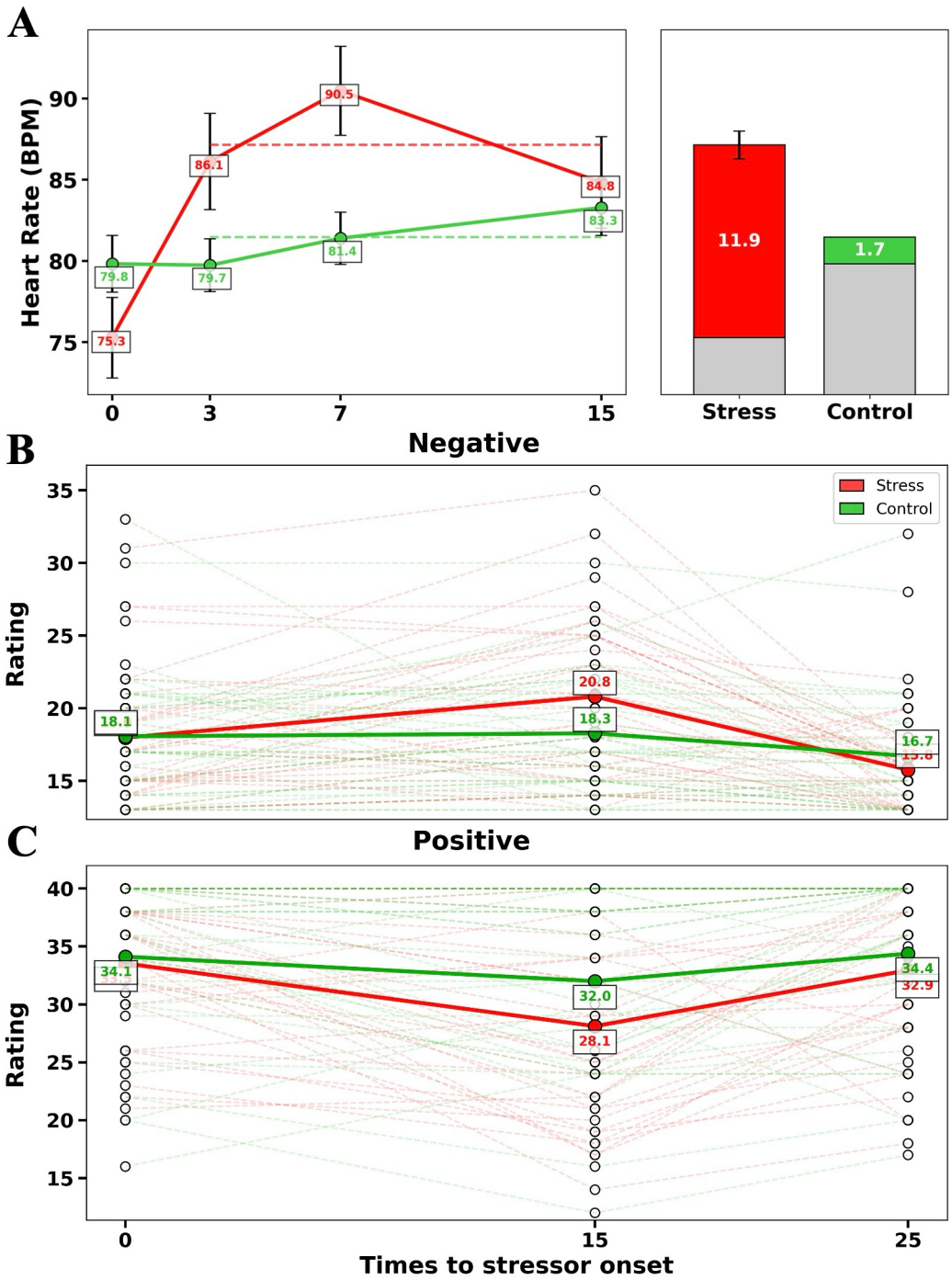

Figure 3. Manipulation Checks. (A) Acute stress induction significantly increased heart rate (HR), measured in beats per minute (BPM), during the Trier Social Stress Test (TSST). The heart rate response was defined as the mean BPM difference between the TSST and baseline phases, with the stress group showing a mean increase of 11.9 BPM (SD = 4.95) and the control group showing a mean increase of 1.7 BPM (SD = 3.98). **(B)** Stress significantly increased subjective ratings of negative valence. **(C)** Stress significantly decreased subjective ratings of positive valence. Note: Results from Experiment 1 are presented here. Similar stress responses were observed in Experiment 2, which are shown in **Figure S2**.

2.2 Memory inference performance

In Experiment 1, we examined whether acute stress impaired memory inference (i.e., the ability to recall item C based on cue A in previously learned ABC triads). Memory inference performance was calculated as the percentage of correct AC inferences, contingent on successful recall of both AB and BC pairs to control for individual differences in memory encoding. Results revealed that acute stress significantly impaired memory inference compared to the control condition (t(66) = -2.160, p = 0.034, Cohen's d = -0.523, 95% CI = [-0.10，-0.00], **Figure 4A left panel**).

We further conducted exploratory correlational analyses to investigate relationships between individual differences in stress responses and memory inference. Stronger heart rate responses to stress were associated with poorer memory inference (r(126) = -0.175, p =0.048; **Figure 4B**), while higher reductions in positive affect correlated with worse performance (r(126) = -0.162, p =0.068; **Figure 4C**). Subsequent correlation analyses, although not reaching strict statistical significance, suggest that subjective positive affect may serve as a potential resilience factor in protecting memory inference under acute stress.

Next, Experiment 2 tested whether retrieval practice could mitigate the stress-induced inference impairment reported in the Experiment1. Compared to Experiment 1, Experiment 2 included a retrieval practice phase immediately following the initial encoding of AB and BC pairs. On Day 2 participants underwent the same stress procedure and memory inference task as in Experiment 1. Unlike Experiment 1, no significant difference in memory inference performance was observed between the stress and control group (t(58) = -0.488, p = 0.627, Cohen's d = -0.125, 95% CI = [-0.06，0.04], **Figure 4A right panel**). Notably, the retrieval

practice phase included only AB and BC pairs, with no AC inferences. Participants were informed that they would be tested on these practiced pairs and were unaware that they would instead be assessed on AC associations on Day 2, thereby ruling out the possibility that the observed protective effects of retrieval practice resulted from explicit AC practice on Day 1.

To further examine the interaction between acute stress and retrieval practice, we combined data from both experiments and revealed significant main effect of retrieval practice ($F(1, 124) = 4.589$, $p = 0.034$, $\eta 2 = 0.036$), suggesting that the retrieval practice enhanced inference performance regardless of acute stress induction. However, the main effect of acute stress was not significant ($F(1, 124) = 3.609$, $p = 0.060$, $\eta^2 = 0.028$), largely because the retrieval practice reduced the stress effect in experiment2. In summary, we found that acute stress significantly impairs memory inference ($F(1, 124) =4.599$, $p = 0.034$, $\eta^2 = 0.036$, 95% CI = [0.004,0.095]), while with the protection of retrieval practice, acute stress did not affect memory inference anymore ($F(1, 124) =0.242$, $p =0.624$, $\eta^2 = 0.002$, 95% CI = [-0.037,0.061])).

An alternative explanation is that retrieval practice in Experiment 2 reduced stress levels during the memory reference phase, resulting in undamaged performance under stress. However, this is unlikely for two reasons: (1) between-group design ensured participants were unaware of retrieval practice manipulation in the other group, preventing confidence biases; and (2) retrieval practice and memory inference are distinct tasks, such that inference task-related stress would not be alleviated by prior retrieval practice on AB and BC. Additionally, we conducted data analyses to rule out this alternative explanation. We analyzed the mean PANAS scores before and after the memory inference phase. Independent

samples t-tests were performed with retrieval practice as the between-subjects factor and subjective positive and negative scores as the dependent variables. For both subjective positive and negative ratings, no significant differences between groups were observed in the memory inference phase between Experiment 1 and Experiment 2 (subjective positive ratings: $t(62) = 1.895$, $p = 0.063$, Cohen's d = 0.475, 95% CI = [-0.15,6.21]; subjective negative ratings: $t(62) = -0.985$, $p = 0.329$, Cohen's d = -0.247, 95% CI = [-3.85， 1.37]).

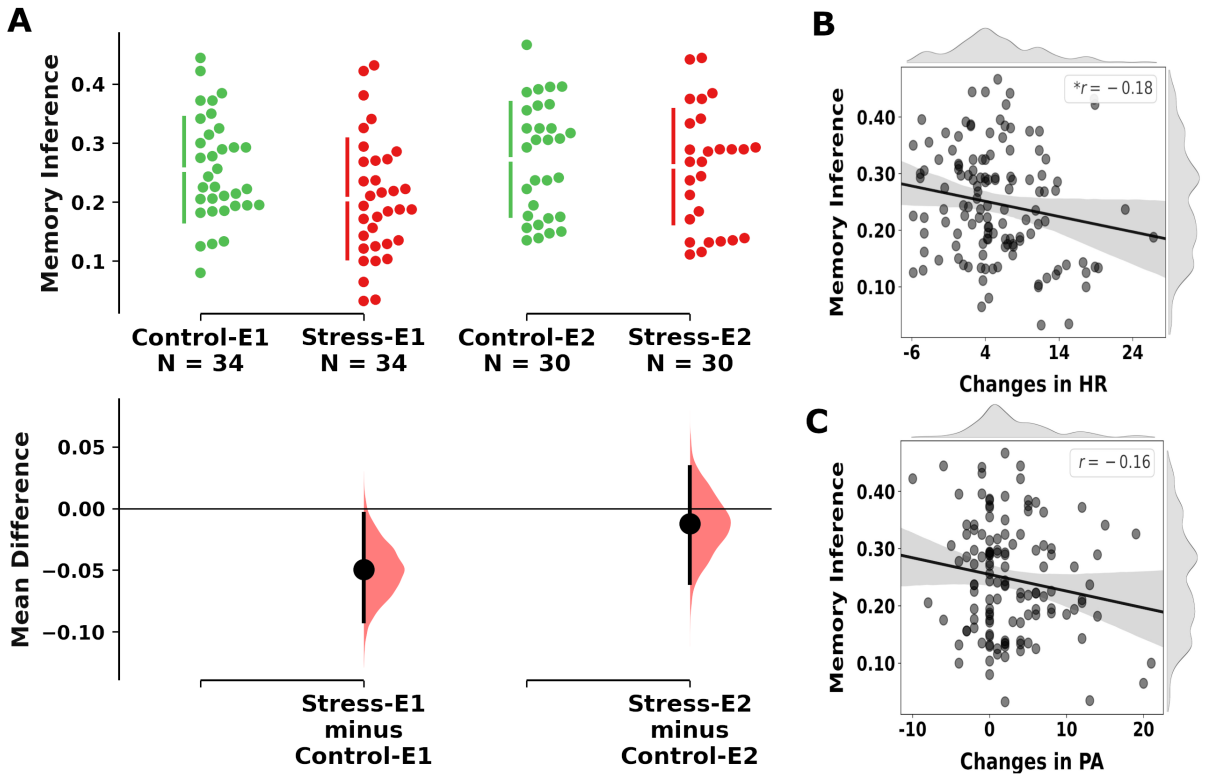

Figure4 Memory inference performance. (A) In Experiment 1, the stress group (i.e., Stress-E1) demonstrated significantly lower memory inference performance compared to the control group (i.e., Control-E1). In Experiment 2, where retrieval practice was conducted immediately after encoding to enhance memory representations, no significant difference in memory inference performance was observed between the stress and control groups (i.e., Stress-E2 vs Control-E2). Additionally, we observed a general improvement in inference performance following retrieval practice, regardless of the stress manipulation. **(B)** Larger increases in heart rate during the stress induction phase predicted greater impairment in memory inference performance. **(C)** Regression analysis revealed a negative relationship between changes in positive affect before and after stress induction and memory inference performance.

2.3 Rapid neural reactivation of bridge information supports subsequent memory inference

To capture the neural dynamics underlying memory inference, we employed multivariate pattern analysis (MVPA) on time-resolved EEG data. Our rationale was that if bridge information (i.e., picture B) is mentally reactivated during inference, then a classifier trained to discriminate between face and building images—using neural activity collected during an independent localizer phase—would successfully classify the category of the reactivated bridge information during inference, even though only object images (i.e., cue A and answer C) were presented (**Figure 5A**). We used this neural decoding strategy as a means to capture underlying mental processes that are difficult to assess via self-report and examined how this neural measure of memory reactivation supports memory inference. Following the principle of subsequent memory effect, we defined the subsequent inference effect contrast (i.e., successful vs. unsuccessful inference trials) to investigate whether neural reactivation of bridge information predicted subsequent correct behavioral responses (**Figure 5B**). We present the EEG results in the following order: (1) within-task classification, in which we trained and tested the model to classify the category of pictures on localizer phase data; (2) cross-task classification, where the classifier was trained on localizer data and tested on inference phase data; and (3) analyses of the subsequent inference effect intergrating outcomes of neural reactivation analyses and trial-level behavioral responses.

First, we validated the classifier's ability to capture the temporal dynamics of category representations. Specifically, we performed within-task classification by analyzing EEG data acquired while participants viewed a separate set of images from the same categories (e.g.,

faces and buildings) used in the main memory task, though with different specific images (see Methods: Localizer Phase for more details). We observed significant within-task classification, ranging from 90 ms to 950 ms in the decoding matrix (**Figure 5C**), suggesting that the classifier was well-trained and ready to be generalizd to the inference phase.

Next, to investigate whether and when the visual category (i.e., face or building) of bridge information (B) can be decoded from brain activity during an A–C inference trial, we applied cross-task classification to the time-resolved EEG data collected during the inference phase. The classifier was trained on neural activity elicited when participants viewed pictures from specific categories and then applied to the time window during which memory cue A was presented, prior to the selection of answer C. We hypothesized that the category information of bridge information (B) would be reactivated, despite the absence of face or building stimuli during this phase. To examine the temporal dynamics of rapid neural reactivation at the single-trial level, we conducted a series of time-resolved decoding analyses. Linear discriminant analysis (LDA; see Methods) was employed to classify visual category information based on the EEG topography at each time point following the presentation of memory cue A. Two primary cross-task classification clusters were identified in the decoding accuracy maps. The first cluster extended from 180 to 340 ms during the localizer phase and from 710 to 3930 ms during the inference phase. The second cluster spanned 1100 to 1500 ms during the localizer phase and comprised two distinct intervals during the inference phase: 180 to 1520 ms and 2760 to 3440 ms (**Figure S3**).

Lastly, to validate the behavioral relevance of the identified neural reactivation of bridge information (B), we examined whether this reactivation predicted subsequent successful

inference (i.e., Subsequent Inference Effect). Each trial was labeled as either successful or unsuccessful, and we anticipated that rapid neural reactivation would be more pronounced in successful trials compared to unsuccessful ones. For each participant, two decoding matrices were generated: one for successful inference and one for unsuccessful inference. Two participants were excluded from the Subsequent Inference Effect analyses because two types of trials could not be defined for them, resulting in a final sample of 126 participants. When focusing on neural activity preceding successful inference, we observed significant cross-task classification, ranging from 80 ms to 260 ms in the localizer phase and from 420 ms to 4000 ms in the inference phase (**Figure 5D, upper organge panel**). In contrast, when applying the classifier to neural activity preceding unsuccessful inference, we observed significant cross-task classification, ranging from 570 ms to 1500 ms in the localizer phase and from 70 ms to 610 ms in the inference phase (**Figure 5D, lower purple panel)**. Moreover, we identified the optimal time window to capture memory reactivation (See Methods for details), and then average each participant's 2D maps across time points within that time window and adjusting by subtracting the 50% chance level, thereby setting 0% as the new baseline. An initial paired t-test revealed that the classifier decoded reactivated bridge information (B) more accurately prior to successful inference than unsuccessful inference ($t(124) = 2.063$, $p = 0.04$, Cohen's d = 0.260, 95% CI = [0.126, 5.450]; **Figure 5E**). Follow-up tests indicated that the category of bridge information was significantly reactivated above chance levels prior to a successful inference response ($t(125) = 1.941$, $p = 0.055$, Cohen's d = 0.173, 95% CI = [–0.041, 4.221]). Conversely, for unsuccessful inference trials, decoding accuracy did not significantly differ from chance ($t(125) = -0.855$, $p = 0.394$, Cohen's d = –0.076, 95% CI = [–2.313, 0.917]).

These findings underscore the behavioral relevance of the detected reactivation of bridge information during the inference phase.

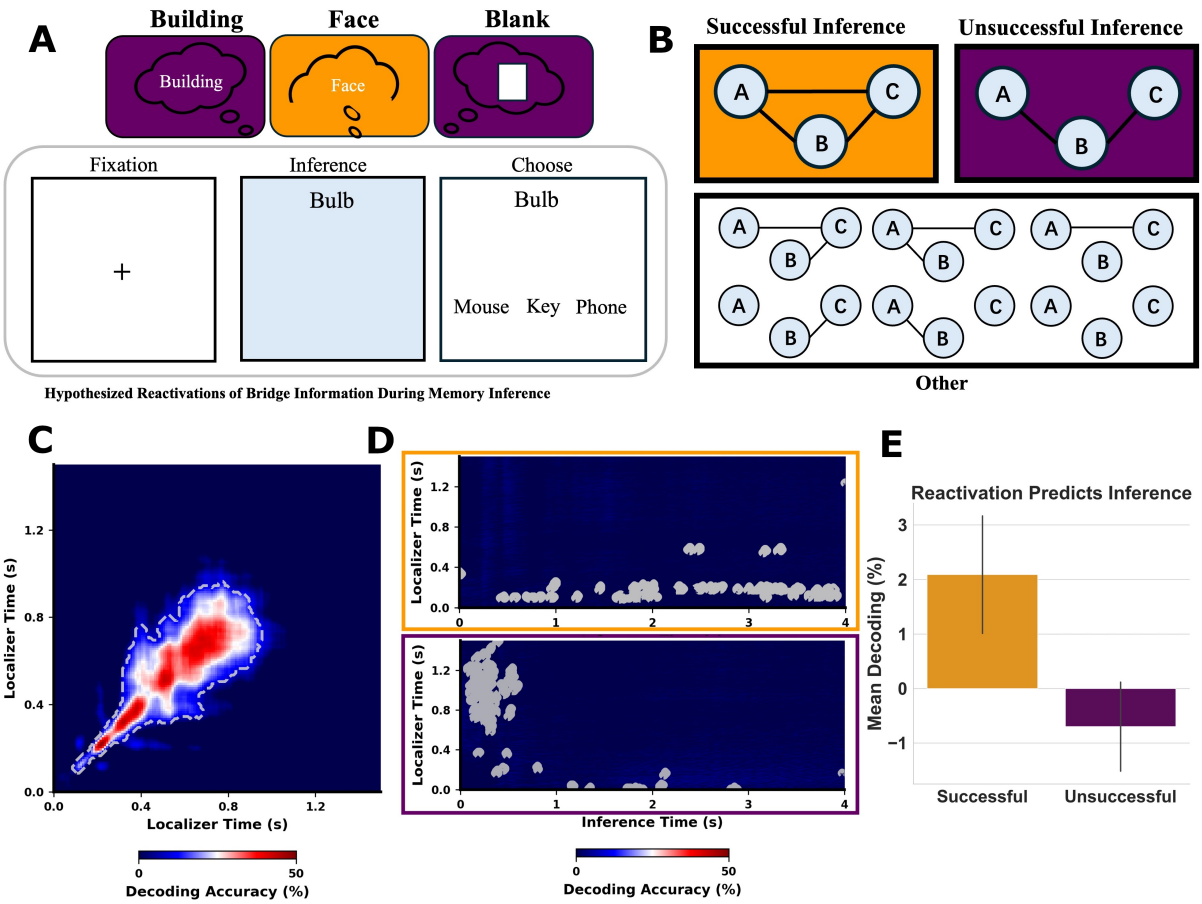

Figure5 Bridge information were reactivated to support subsequent successful inference. (A) Hypothesized reactivation of bridge information during the inference phase. To select the correct image-C based on cue-A, participants were hypothesized to mentally retrieve the bridge information-B (e.g., a face in the example trial, in orange). In some cases, the incorrect image (e.g., a building, in purple) might also be reactivated, or no image might be recalled during this phase. **(B)** Definition of successful and unsuccessful inference trials. In successful trials, both associations (A–B and B–C) were remembered, allowing for the correct inference of the A–C relationship (Orange Box). In contrast, in unsuccessful trials, despite remembering A–B and B–C, participants failed to correctly infer the A–C relationship (Purple Box). Trials in which either A–B or B–C associations were not encoded prior to inference were excluded from the analysis (White Box). **(C)** Results of time-generalization analysis from within-task classification.

EEG data from the localizer phase were used to train a machine learning model to classify faces and buildings based on localizer-phase EEG activity. **(D)** Results of time-generalization analysis from cross-task classification. EEG data from the localizer phase were used to train a machine learning model to classify faces and buildings using EEG data from the inference phase. Decoding accuracy matrices were computed separately for successful and unsuccessful inference trials (depicted in orange and purple, respectively). **(E)** Mean decoding accuracy within the optimal decoding time window, relative to chance, for successful and unsuccessful inference trials. Significantly above-chance decoding of bridge information predicted subsequent successful inference, demonstrating the Subsequent Inference Effect.

2.4 Acute stress disrupts reactivated neural representation of bridge information, but retrieval practice restores this process

Behavioral data indicate that acute stress impairs memory inferences in humans. To investigate how acute stress influences the neural mechanisms underlying the flexible use of encoded memories to infer A–C associations, we employed rapid neural reactivation of bridge information as a neural lens. We used two neural indices to capture the dynamics of memory reactivation: *Decodability* and *Mean Decoding Accuracy* (see Methods for details). *Decodability* reflects the temporal consistency of memory reactivation, with higher scores suggesting sustained activation during the inference phase. In contrast, *Mean Decoding Accuracy* quantifies the strength of memory reactivation by averaging the accuracy of decoding bridging information from brain activity, adjusted for chance levels; higher values indicate stronger retrieval of memory representations.

Initially, we provided neural evidence that stress impairs inference by disrupting the reactivation of bridge information. In Experiment 1, we compared *Decodability* and *Mean*

Decoding Accuracy between stress (Stress-E1) and control participants (Control-E1). Acute stress disrupted bridge information reactivation relative to the control condition, as indicated by lower Decodability measures (t(66) = -1.535, p = 0.065, Cohen's d = -0.372, 95% CI = [-22.707, 2.971]; **Figure 6A, left panel**) and a similar, near-significant trend for Mean Decoding Accuracy (t(66) = -1.271, p = 0.104, Cohen's d = -0.308, 95% CI = [-5.432, 1.206]; **Figure 6B, left panel**). To further support the notion that acute stress disrupts neural reactivation of bridge information, we correlated individual differences in acute stress responses (i.e., changes in HR and PA) with *Decodability* and *Mean Decoding Accuracy* measures. Stronger HR responses under stress predicted weaker subsequent neural reactivation of bridge information, as shown by both *Decodability* (r(62) = -0.338, p = 0.006; **Figure 6C**) and *Mean Decoding Accuracy* (r(62) = -0.259, p = 0.039; **Figure 6E**). Critically, HR—a rapid biological response to acute stress—had already returned to baseline during the inference phase, indicating that HR changes were measured prior to the reactivation of bridge information. This temporal ordering further strengthens the conclusion that acute stress causally impairs subsequent neural reactivation, as measured by EEG decoding. We also examined the associations between changes in PA and neural reactivation measures, which revealed a similar negative relationship that did not reach significance (Decodability: r(62) = -0.22, p = 0.08; **Figure 6D**; Mean Decoding Accuracy: r(62) = -0.20, p = 0.10; **Figure 6F**).

In Experiment 1, participants under stress were less likely to infer correct A–C associations. However, behavioral data from Experiment 2 indicate that retrieval practice of AB and BC associations can prevent this stress-induced impairment and enhance overall inference performance. If neural reactivation of bridge information contributes to successful

inference and reflects the impact of stress, then it should be modulated by retrieval practice. Accordingly, we compared the decoded neural dynamics of bridge information during inference between participants who did and did not use retrieval practice (i.e., Experiment 1 vs. Experiment 2), irrespective of stress manipulation, to reveal the general effect of retrieval practice. We found that participants in Experiment 2 exhibited significantly stronger reactivation of bridge information than those in Experiment 1, as indexed by higher *Decodability* ($t(126) = 1.995$, $p = 0.048$, Cohen's d = 0.355, 95% CI = [0.076, 20.034]) and higher *Mean Decoding Accuracy* ($t(126) = 2.099$, $p = 0.038$, Cohen's d = 0.374, 95% CI = [0.162, 5.609]).

Subsequently, we examined participants in Experiment 2 who all underwent retrieval practice but received either stress (Stress-E2) or control (Control-E2) manipulations prior to inference, to elucidate why the memory inference process is resilient to stress following retrieval practice. We observed that participants in Experiment 2 who received stress not only avoided the stress-induced impairment in reactivation but also exhibited significantly stronger reactivation of bridge information compared to those receiving control manipulations, as reflected in both higher Decodability ($t(126) = 3.132$, $p = 0.003$, Cohen's d = 0.809, 95% CI = [8.068, 36.972]; **Figure 6A, right panel**) and higher Mean Decoding Accuracy ($t(126) = 2.549$, $p = 0.014$, Cohen's d = 0.658, 95% CI = [1.136, 9.513]; **Figure 6B, right panel**).

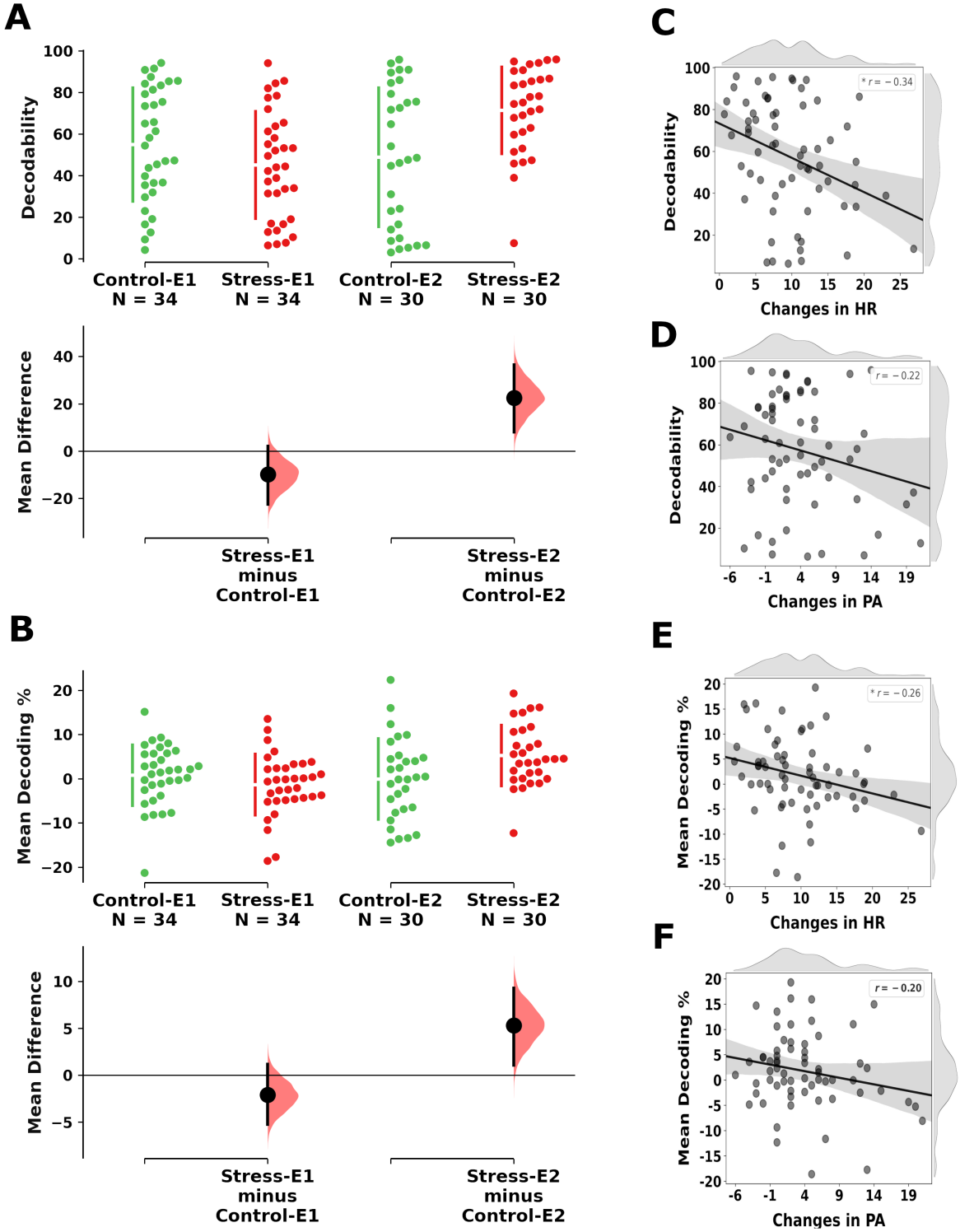

Figure 6 Neural reactivation under acute stress and after retrieval practice. (A) Decodability analysis as a function of stress and retrieval practice. In Experiment 1, stressed participants exhibited lower Decodability during inference compared to controls (Stress-E1 vs. Control-E1). In Experiment 2, following retrieval practice to enhance AB and BC memory representations, stressed participants showed higher

Decodability than controls (Stress-E2 vs. Control-E2). (B) Stress-induced increases in heart rate (HR) predicted lower Decodability. (C) Stress-induced decreases in positive affect (PA) were associated with lower Decodability. (D) Mean Decoding Accuracy analysis as a function of stress and retrieval practice. In Experiment 1, stressed participants exhibited lower Mean Decoding Accuracy during inference compared to controls. In Experiment 2, following retrieval practice, stressed participants demonstrated higher Mean Decoding Accuracy than controls. (E) Stress-induced HR increases predicted lower Mean Decoding Accuracy. (F) Stress-induced PA decreases were associated with lower Mean Decoding Accuracy.

3. Discussion

Our study provides evidence that (1) acute stress impairs memory inference and disrupt its underlying neural mechanisms; (2) retrieval practice can prevent the stress-induced inference impairment at both behavioral and neural levels. Critically, using EEG and multivariate time-resolved decoding approach, we captured the rapid neural reactivation of bridge information during the inferential process. This rapid memory reactivation, predictive of subsequent successful memory inference, was disrupted in stress participants but restored by retrieval practice. Collectively, our data offers a novel understanding of memory inference in humans, demonstrate its disruption under stress, and highlight the protective potential of retrieval practice.

3.1 Rapid neural reactivation of bridge information for successful memory inference

The associative inference paradigm—which involves recombining information across overlapping experiences (e.g., AB, BC) to enable novel inference (e.g., A→C)—has been widely applied in memory research[4–6,8,9], yielding significant insights into the online integration of separately learned information. For example, human fMRI studies have demonstrated that neural activity in the hippocampus, prefrontal cortex[4–6,8] and medial temporal lobe[6,9] play critical roles in memory integration. Furthermore, human MEG evidence suggests that theta oscillations support information exchange between the hippocampus and prefrontal cortex[39]. Instead of focusing on online memory integration, we leveraged high-temporal-resolution EEG and neural decoding techniques to investigate the neural mechanisms underlying the memory inference phase. Specifically, we recorded brain activity during the active A–C inference phase and observed that the rapid neural reactivation

of bridging information (B) differed between successful and unsuccessful memory inference. Our finding that bridging information is reactivated during inference and can be decoded from neural activity extends the traditional view of encoding-retrieval[40] and perception-imagery[41] similarity. Previous studies employing similar neural decoding approaches have demonstrated that neural activation patterns are shared by encoding-retrieval[42,43] and perception-imagery processes[44,45]. In contrast to retrieval or imagery—mental processes in which participants receive explicit instructions to consciously reactivate neural patterns—our neural signature resembles spontaneous neural reactivation that supports subsequent inferential processing. We argue that because the bridge information in our design—a necessary step for achieving the final cognitive goal—is reactivated too rapidly and transiently, it is particularly well suited for investigation using EEG and MEG methodologies. Building on this mechanistic insight, our study extends prior research by shifting the focus of associative memory processing from encoding and retrieval to inference, and it provides a neural signature of successful inference. Although memory integration has predominantly been investigated during encoding, memory is fundamentally a multi-stage process encompassing encoding, consolidation, and retrieval. This temporal expansion offers a unique opportunity to study the effects of experimental manipulations (in this study, acute stress) on memory inference. Specifically, we induced acute stress immediately before memory inference, thereby isolating its effects to the inference stage and avoiding potential confounds arising from stress impacts on other stages, such as encoding and consolidation. This investigation is highly relevant to educational policy and interventions in real-life scenarios, such as the effects of acute stress prior to an exam.

3.2 Disruptions of memory inference and rapid neural reactivation under stress

Having elucidated the neural mechanisms underlying memory inference, we next examined how this process is disrupted by acute stress—a pervasive real-world challenge. Behaviorally, we provided evidence from healthy human participants indicating that the ability to make inferences is impaired under stress. While it is well established that acute stress impairs simple memory functions, such as direct memory retrieval[14,15,26,34,46–48], our findings extend this understanding by showing that stress also disrupts complex memory functions, specifically A-C memory inference in this study. In addition to our results, previous studies have shown that stress reduces cognitive flexibility in spatial navigation; for example, stressed rats perform poorly in water mazes[49,50] and humans tend to favor familiar routes over shortcuts[44,45]. These findings support the notion that acute stress shifts cognitive memory systems toward more reflexive, habitual systems[26]. Specifically, stress induces greater reliance on the dorsal striatum (habitual memory system) over the hippocampus (cognitive memory system)[51,52]. These neural circuit-level observations help explain the broader stress-induced impairments in flexible cognitive functions: stress also disrupts flexible updating of aversion values[53] and schema learning[54–56]. Notably, glucocorticoids primarily affect hippocampal function, leading to increased dependence on the dorsal striatum during memory inference. This hippocampal vulnerability is further exacerbated by stress-induced prefrontal dysfunction[57], as drug-induced glucocorticoid secretion reduces prefrontal activity during retrieval tasks[23]. Although our study lacks direct human neuroimaging evidence from the prefrontal-hippocampal loop due to the EEG method employed, we provide complementary evidence through neural reactivation, leveraging the

optimal temporal resolution of EEG. This allowed us to pinpoint the specific time window in which acute stress exerts a profound disruptive effect. Collectively, these findings offer new evidence from the associative memory paradigm, linking acute stress responses in humans with reduced cognitive flexibility.

3.3 Retrieval practice protects inference and restores neural reactivation under stress

Despite stress-induced impairments in memory inference, our investigation identified a potent cognitive-psychology-based countermeasure—retrieval practice—and examined its underlying neural mechanisms. Behaviorally, we found that retrieval practice not only enhanced overall inference performance but also mitigated the adverse effects of stress on inference. Specifically, participants who repeatedly learned AB and BC associations through retrieval practice were immune to stress-induced impairments in A–C inferences, as reported in Study 1.

Decades of research has supported a reboust consensus that retrieval practice enhances learning and memory [32,33,58,59]. This enhancement is evident across a wide range of tasks, from simple word recall[34,60] to more complex activities such as comprehending academic terminology[61] and popular science articles[33], and even extends to real-life scenarios including face-scene pairing[37] visual-spatial location tasks[62], and classroom learning[58].

One behavioral study employing the demonstrated that retrieval practice protects memory against acute stress [34], and another spatial navigation study found that the second navigation attempt, compared to the first, is less affected by acute stress[27]. Although both studies suggest that retrieval practice exerts a protective effect against stress, they employed identical tests during both the retrieval practice and final test phases. By contrast, our

findings, for the first time, reveal that retrieval practice can enhance task performance and provide protection against stress even when the practiced material differs from that which is tested. Conventionally, the enhancement and protection induced by retrieval practice is attributed to the alignment between the formats of retrieval practice and the final test[34,63,64]. However, our results suggest that the benefits of retrieval practice may depend less on content-specific alignment than previously assumed. From both an educational and practical perspective, our findings offer a new viewpoint on the effectiveness of retrieval practice as a learning strategy. Although some have criticized retrieval practice as a low-level, time-consuming memory enhancement strategy that does not facilitate the flexible application of knowledge, we challenge this view by providing both behavioral and neural evidence that retrieval practice—as a conscious-level intervention applied to previously learned information—can facilitate future flexible inference and recombination while also rendering the process more resilient to stress.

At the neural level, our findings suggest a novel effect of retrieval practice on associative memory—specifically, the enhancement of flexible memory inference via a potential preplay mechanism[65,66] that links indirectly encoded associations (i.e., A–C). In contrast to previous studies that employed identical content for retrieval practice and final tests (e.g., practicing AB and testing AB), we instructed participants to use retrieval practice to strengthen the memory representations of previously encoded AB and BC associations and subsequently assessed their inference of A–C associations. One recent study employing a similar multielement memory paradigm demonstrated that sleep can selectively benefit memory for associations that were not directly encoded, attributing this effect to sleep-

dependent memory consolidation as a reinforcement mechanism[63]. Another study on retrieval practice showed that its memory benefits are not limited to retrieved information[67]; instead, they generalize to nonretrieved information, potentially because the two types of information share a similar temporal context. We interpret our findings as reflecting a general enhancement of the ABC structure resulting from retrieval practice of previously encoded information, and we propose two phases during which the A–C associations may be spontaneously linked and enhanced. The first phase occurs during retrieval practice itself, wherein the ABC structure is unconsciously established and strengthened. The second phase involves memory consolidation, which occurs during both wakefulness and sleep. Retrieval practice effectively tags the ABC structure for spontaneous reactivation, thereby facilitating future flexible inference of A–C associations.

3.4 Limitation and further directions

Although these findings advance our understanding of the interplay between memory inference, stress, and retrieval practice, two limitations warrant discussion. First, the use of specific stimuli (i.e., pictures of famous faces, buildings, and everyday objects) to form the ABC structure and investigate AC inference may limit the ecological validity of our design. Future studies should employ more universally applicable materials to enhance generalizability and more accurately mimic the processes of knowledge acquisition and generalization. Second, the EEG signal cannot be used to infer region-specific information representation, circular or network-level interactions, or reorganization under stress or following retrieval practice. Future studies using fMRI and similar designs may provide complementary evidence to address these issues. Mechanistically, we propose two

complementary explanations for the protective effect of retrieval practice based on previous related fMRI studies: (1) retrieval practice may optimize neural representations in memory-related brain regions, as evidenced by altered activation patterns in the prefrontal, parietal, and temporal lobes[55–59]. This representational strength may buffer against stress and enable flexible memory inference; and (2) retrieval practice may facilitate the rapid transfer of memory representations from the hippocampus to the cortex. Unlike traditional memory consolidation theory[60], recent theories and experimental evidence suggest that retrieval practice accelerates this transfer, enabling hippocampal-dependent memory traces to quickly form cortical representations[61], thereby safeguarding hippocampal-dependent memory function from stress-induced impairment.

3.5 Conclusions

Converging evidence from studies in both humans and animals indicates that acute stress impairs memory retrieval by disrupting the interaction between the prefrontal cortex and the hippocampus. Here, we leveraged the associative inference paradigm to ask whether and how acute stress might also impair memory inference, and whether retrieval practice could mitigate the adverse effects of stress on inference. Using EEG and multivariate decoding analyses, we identified the rapid reactivation of neural representations of bridging information as a predictive neural signature of successful memory inference. We observed that both inference performance and this neural reactivation were compromised under stress. Notably, retrieval practice was found to prevent stress-induced inference impairment, and to restore the rapid neural reactivation of bridge information, essential for the inferential process. These findings collectively enhance our understanding of how negative emotional

states limit the capacity to use memories flexibly for making inferences and underscore the potential of retrieval practice, an effective learning strategy, to counteract stress-induced cognitive impairments. The implications of these results are significant for real-world situations, such as taking exams, surgical procedures, and public speaking, by offering mechanism insights into high-level cognitive impairments under stress and suggesting effective behavioral strategies to buffer these effects.

Method

1. Participants

We tested a total of 128 healthy college students across both experiments. Experiment 1 involved 68 participants (50 females; M_{age}=20.43 ± 1.80 years), who were randomly assigned to either a stress condition (n=34; M_{age} = 20.35 years, SD_{age} = 1.87, 26 females) or a control condition (n=34; M_{age} = 20.5 years, SD_{age} = 1.75, 24 females). There were no significant differences in age or gender between the two conditions (age: t(66) = -0.335, p = 0.739, Cohen's d = -0.081, 95% CI = [-1.02,0.73]; gender: $\chi^2(1,68)=0.302$, p=0.582). The remaining 60 participants (50 females; M_{age} 20.12 ± 1.46 years) were tested in the Experiment 2, which involved random assignment to inference under stress (n=30; M_{age} = 19.93 years, SD_{age} = 1.41, 28 females) or inference without stress (n=30; M_{age} = 20.3 years, SD_{age} = 1.51, 23 females). Again, there were no significant differences in age or gender between these groups (age: t(58) = -0.971, p = 0.336, Cohen's d = -0.251, 95% CI = [-1.12,0.39]; gender: χ^2（1,60）=2.092, p=0.148). All participants were right-handed and had normal or corrected-to-normal vision. The study received approval from the Ethics Review Committee for Life Sciences at Central China Normal University, and all participants provided informed consent prior to participation.

2. Experimental procedure

2.1 Task overview

We conducted two two-day experiments to investigate whether retrieval practice can mitigate stress-induced memory inference impairments and explore the associated neural reactivation mechanisms. The only difference between the two experiments was that

Experiment 2 included a retrieval practice phase following the initial encoding phase. Detailed procedures are illustrated in **Figure S1**.

On Day 2, they returned for baseline stress measurements, which included heart rate recording and the Positive and Negative Affect Scale (PANAS, PA1, and NA1). Participants were then randomly assigned to either the stress or control group. After the stress induction or control procedure, and once cortisol levels peaked, they completed a three-alternative forced-choice test assessing inference (AC) while EEG recorded brain activity. Memory for directly learned associations (AB, BC) was also tested for subsequent trial-by-trial behavioral analysis. The second PANAS (PA3 and NA3, PA4 and NA4) was administered 10 minutes post-stress, with continuous heart rate monitoring. Finally, participants completed a localizer task to train an EEG-based machine learning model to classify two distinct types of visual stimuli—famous faces and famous buildings—based on neural response patterns. Data from the localizer task served as the training dataset for the model, which was used to decode neural responses during the memory inference phase, specifically targeting the reactivation of bridge information B. We primarily used behavioral data and neural responses during the AC inference phase to address our main research questions.

2.2 Stimuli

We employed two distinct sets of visual stimuli for this study. One set was used in the main memory task, comprising the encoding, inference, and retrieval phases, while the other was used for the localizer phase. The two sets were independent of each other, and all images are publicly available (see our statement on Data and code availability below) for future experiments or replication purposes.

Stimuli for the main memory task: The main memory task used 72 images divided into three categories: faces, buildings, and objects. Specifically, this set included 12 images of famous faces (e.g., Andy Lau), 12 images of iconic buildings (e.g., the Great Wall), and 48 images of everyday objects (e.g., an iPad). These images were selected to reflect items commonly encountered in the daily lives of Chinese college students. The stimuli were organized into 24 "ABC" memory triads, where A and C were always two distinct object images, and B was either a face or a building image. Stimuli for the localizer phase: The localizer phase included 108 images, evenly distributed across the same three categories (36 famous faces, 36 iconic buildings, and 36 everyday objects). The images of everyday objects were sourced from the THINGS database[68], while images of famous faces and buildings were obtained from publicly available online sources. All images were resized to a resolution of 474 × 658 pixels and edited to remove distracting elements.

2.3 Encoding phase

During the encoding phase, participants learned 48 memory pairs (AB and BC) derived from ABC triads, with each triad divided into two separate components (AB and BC) for independent encoding. Participants were explicitly instructed to learn the relationship between the two images displayed on the screen; however, no hints were provided regarding the potential to form a complete triad. Each encoding trial began with a fixation cross presented for 100 milliseconds, followed by the display of an AB or BC image pair for 2500 milliseconds. The next trial began immediately after the fixation cross was presented again. Each AB or BC pair was presented twice, with the positions of the images reversed between presentations. For example, if image A appeared on the left and image B on the right during

the first presentation, their positions were swapped during the second presentation, with image B on the left and image A on the right. Notably, images A and C were never presented together during encoding, enabling a memory inference test on Day 2. To control the temporal interval between the presentation of AB and BC pairs, a mini-block design was employed, ensuring a trial interval of 1 to 4 trials between corresponding AB and BC pairs. The encoding phase consisted of two runs, with participants allowed to take breaks between runs to maintain focus throughout the task(**Figure S4**).

2.4 Retrieval practice phase

During the retrieval practice phase (included only in Experiment 2, but not Experiment 1 as to experimenter's manipulation), participants were asked to freely recall the memory pairs (i.e., both AB and BC) learned during the encoding phase using paper and pen. They were instructed to recall as many pairs as possible within a two-minute time limit. Immediately following each encoding run, participants performed the retrieval practice run, resulting in two retrieval practice trials in total. In other words, two encoding and retrieval practice trials were administered in an alternating sequence.

2.5 Stress procedure

To ensure equivalent encoding contexts for both control and stress groups and to confine the stress effects to Day 2, all participants underwent the same encoding procedure on Day 1, without knowledge of their Day 2 group assignment (i.e., stress or control).

2.5.1 Stress induction

Acute stress was induced using a modified version of the Trier Social Stress Test (TSST; Kirshbaum et al., 1993), a well-established method for inducing acute stress in a laboratory

setting. Previous research shows that this paradigm reliably triggers increases in both hypothalamic-pituitary-adrenal (HPA) axis and sympathetic nervous system (SNS) activity. Our adaptation included a videotaped free speech and a mental arithmetic task. This protocol consisted of a 3-minute preparation period, a 4-minute public speaking task, and an 8-minute mental arithmetic task. Participants prepared an impromptu speech for a mock high school teaching certification interview as a Chinese teacher, tasked with teaching a collection of prose poems. They were instructed to emphasize organization and clarity, as they would be evaluated on these criteria, and to continue speaking for the full 4 minutes, even if content ran out. Following the speech, participants completed a multiplication task requiring rapid, accurate responses. We replaced the original subtraction problems with two-digit multiplication challenges (e.g., 56×47), issuing an alarm for incorrect or delayed responses within a 5-second window. This modification was intended to increase mental stress compared to the original tasks. Before implementation, our independent pilot study confirmed that completing these calculations within the time limit was nearly unfeasible. Experimenters maintained a neutral demeanor, providing only necessary instructions. Participants were informed that the entire process was video recorded for later evaluation to enhance stress induction, though no recordings were actually preserved to protect participants' privacy.

2.5.2 Control condition

The control condition matched the stress condition in duration and took place in the same room to ensure consistency in environmental cues. However, it lacked the social and self-relevant threat aspects. It consisted of a 7-minute solitary reading period (with no

experimenters present or instructions for preparing a speech) and an 8-minute simple arithmetic task (e.g., 2+3). Participants were not disturbed, even for incorrect or delayed responses.

2.6 Memory inference phase

Before the formal test, we used slides and practice trials featuring completely unrelated images to explain the structure of the ABC task and to clarify that picture A and C were indirectly related through B. After confirming participants' understanding of the inference task, they completed a three-alternative forced-choice memory inference phase. Each inference trial began with the presentation of a fixation point, displayed for a randomly selected duration (500 ms, 1000 ms, 1500 ms, or 2000 ms), followed by the memory cue (i.e., picture A) for 4000 ms. Participants were instructed to vividly retrieve and imagine the answer (i.e., picture C) during this period. This phase was the primary window of interest for EEG analyses, as successful inference required the recall of picture B (which we referred to as the "bridge information") and picture C during this phase. Following the presentation of the memory cue, three candidate pictures were shown, and participants were required to select the correct answer within 3000 ms using the keyboard. One of the pictures was the correct answer (i.e., picture C), while the other two were distractor object images. These distractors had been presented during the Day 1 encoding phase but were not paired with the current memory cue A(**Figure S5**). This setup ensured that participants had to perform the full A-B, B-C inference to select the correct answer, rather than choosing based on familiarity alone (e.g., recognizing the picture from earlier encoding phase). If participants selected the correct option, the trial was labeled as correct. Additionally, we took further steps, analyzing

trial-by-trial data from the retrieval phase to rule out trials where the correct answer was selected due to guessing (See *Statistical Analyses* below).

2.7 Direct memory retrieval phase

In addition to the A-C memory inference task, participants were also tested on their originally learned AB/BC pairs. We refer to this phase as direct memory retrieval to emphasize its distinction from the A-C inference, which involves both memory retrieval and inference. At the start of each direct memory retrieval trial, a fixation point was presented for a randomly selected duration (500 ms, 1000 ms, 1500 ms, or 2000 ms). Next, a memory cue (i.e., picture A or C) and two response options (one being the correct picture B) were presented simultaneously, with a 3000 ms window for participants to make a selection. Similar to the inference phase, the distractor image was a previously presented image from another AB or BC pair and belonged to the same category as the correct option. This task required direct recalling information from the encoding phase without engaging in memory inference. The direct retrieval phase consisted of 48 trials, testing all 24 AB and 24 BC pairs learned(**Figure S6**).

2.8 Localizer phase

At the end of the experimental procedure, participants completed a visual category localizer task employing a 1-back paradigm with images from three distinct categories: famous buildings, famous faces (male and female), and common objects. EEG was used to measure brain activity. The stimuli for this task were unique and did not overlap with those used in the main memory tasks (encoding, inference, and retrieval phases). Each run of the task comprised nine mini-blocks, with three mini-blocks assigned to each image category,

and each mini-block containing 12 trials. During each trial, an image was displayed for 1500 ms, followed by a 500ms inter-trial interval. Participants were instructed to respond quickly and accurately to indicate whether the current image was identical to the previous one(**Figure S7**. This task requirement served as a cover story to ensure participants remained engaged with the stimuli; however, their behavioral responses were not included in subsequent analyses. To prevent category-specific fatigue or habituation, the order of mini-blocks was pseudo-randomized, ensuring that two mini-blocks from the same category did not appear consecutively. The primary objective of the localizer task was to collect neural response patterns specific to each image category, which were later used to train a machine learning model for image classification based on brain activity (see *Decoding of EEG Signals* below).

2.9 Physiological and Psychological measures of acute stress response

In Experiment 1, heart rate, as a physiological measure of acute stress, was continuously monitored using the BIOPAC MP160 system. In Experiment 2, heart rate was tracked with portable equipment, specifically the POLAR V2 system, which includes a Polar WearLink and a heart rate monitor. Heart rate was initially measured for 1 minute to establish a baseline and was continuously monitored throughout the subsequent tasks. Average heart rates were calculated for each experimental stage: a 3-minute preparation, a 4-minute public speaking, and an 8-minute arithmetic stage.

The Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) was employed to assess participants' current subjective affective states as a psychological measure of acute stress. This 20-item scale comprises 10 items for positive affects (e.g., "interested," "excited") and 10 items for negative affects (e.g., "nervous," "scared"). Participants rated

each item on a 5-point scale reflecting their current state, from 1 (very slightly or not at all) to 5 (extremely). Average scores for both positive and negative affect were then computed.

3. EEG acquisition and preprocessing

EEG data were collected using Brain Products GmbH actiCHamp Plus amplifiers at a 1000 Hz sampling rate. Sixty-four electrodes recorded continuous EEG, referenced to the vertex (FCz), with signals amplified through a 0.1-100 Hz band-pass filter and digitized at 250 Hz for storage. Offline processing involved band-pass filtering between 0.1 Hz and 30 Hz, down-sampling to 500 Hz, and re-referencing to the average of the left and right mastoids. To reduce line noise at 50 Hz, a band-stop filter was applied between 49 and 51 Hz. Subsequently, we visually inspected the ERP data and discarded any trials with residual ocular or other artifacts, such as those caused by movement. If an artifact was detected at any scalp site, data from all sites for that trial were excluded from further analysis. Interpolation was used to replace bad channels. Independent component analysis was performed to remove eye-blink and horizontal eye movement artifacts. The continuous EEG signal was segmented into 2500 ms periods for the localizer task and 4500 ms for the inference task, each starting 500 ms before stimulus presentation. Segments were baseline-corrected against the mean voltage of the 500 ms pre-stimulus period. Behavioral data were used to label each trial with a true or false category, indicating whether the trial involved a famous building or a male or female famous face. The EEG data were preprocessed using the MNE-Python software package (Version 1.9.0), applying a band-pass filter from 0.1 to 30 Hz. Following filtering, the entire 0.1–30 Hz bandwidth was analyzed without further frequency decomposition.

4. Multivariate time-resolved decoding of EEG signals

4.1 General overview

Time-resolved decoding was performed using Linear Discriminant Analysis (LDA) implemented in custom Python 3.12 code adapted from NeuroRA[69]. Preprocessed raw amplitudes from 63 EEG channels at discrete time points served as features for classifier training and testing. Two-category LDA classification (i.e., faces vs. buildings) was conducted for each participant and time point, yielding decoding accuracy for each time point within each trial. Decoding accuracy exceeding chance levels indicated significant memory reactivation. Our study employed both within-task and cross-task classification with distinct objectives. Specifically, within-task classification was trained and tested on localizer data to validate classifier performance, whereas cross-task classification was trained on localizer data and tested on inference data to capture the reactivation of category-level information of the bridge information-B. Within-task classification generated two-dimensional decoding accuracy maps for each participant spanning 1500 ms. Cross-task classification produced decoding accuracy maps for each participant, spanning 1500 ms in the training time domain (i.e., localizer time) and 4000 ms in the testing time domain (i.e., inference time). Additionally, to analyze subsequent inference effects, trials were separated based on whether they led to successful or unsuccessful inference, resulting in two maps per participant.

4.2 Within-task classfication

A sliding time window approach was employed to extract features from the raw EEG signals. Specifically, a window encompassing 5 consecutive time points was shifted along the recording in increments of 5 time points. Rather than averaging the data within each window, we retained individual time point values by adopting a feature-based approach, thereby

preserving the fine-grained temporal structure. To enhance the signal-to-noise ratio, trials were grouped by averaging 7 trials per category. For each subject, the available trials were randomized to mitigate potential ordering biases and subsequently partitioned into non-overlapping groups based on the minimum number of trials available across categories. This procedure yielded a balanced set of averaged trials for each class, ensuring that subsequent analyses were not biased by unequal trial counts. A cross-temporal decoding framework was implemented to assess the generalizability of neural representations over time. For each subject, a stratified k-fold cross-validation procedure (with 5 folds and 2 independent repetitions) was applied to the averaged data. In each cross-validation fold, the following steps were executed: (1) Standardization: Feature vectors from the training set were standardized using z-scoring to ensure that scaling differences did not influence the classifier. (2) Classification: A Linear Discriminant Analysis (LDA) classifier was employed to decode category-specific information from the neural features. The model was trained on the standardized data from the training folds and subsequently tested on the held-out fold. (3) Cross-Temporal Generalization: Decoding accuracy was evaluated not only for the time window corresponding to the training data but also for all other time windows, yielding a cross-temporal generalization matrix for each subject, wherein each matrix element reflects the classifier's performance when trained on one time window and tested on another. (4) Smoothing of Decoding Accuracies: To enhance the robustness of the decoding results, the cross-temporal accuracy matrices were refined through spatial smoothing. For each subject, a two-dimensional sliding window (covering a neighborhood of approximately 5 consecutive time points in each dimension) was applied to the accuracy matrices. Local averaging across

neighboring elements reduced the impact of noise and highlighted consistent patterns of cross-temporal generalization.

4.3 Cross-task classfication

EEG data from both the localizer and inference tasks, along with the corresponding category labels, were used for cross-task classification. Similar to within-task classification, decoding analyses were conducted over sliding temporal windows, where individual time points within each window were preserved as distinct features. The following steps, implemented in within-task classification, were applied: trial averaging, class balancing, decoding analysis and classifier training, as well as spatial smoothing of decoding accuracies. The key difference between within-task and cross-task classification is that, in the latter, the classifier was trained and fixed for application to the inference task, even though no face or building images were presented during this phase.

5. Statistical analysis

5.1 Validation of Acute Stress Induction:

To assess the induction of acute stress, two-way mixed ANOVAs were performed on heart rate and subjective affect (separately for positive and negative affect), with group (stress vs. control) as a between-subjects factor and acquisition period as a within-subjects factor. We also defined three indices to represent individual differences in acute stress responses. For heart rate and negative affect ratings, we subtracted the baseline value from the post-induction value, with higher values indicating stronger stress responses. In contrast, for positive affect ratings, we subtracted the post-induction value from the baseline value, with a larger reduction indicating a stronger stress response.

5.2 Analysis of Memory Inference Performance

To examine whether retrieval practice mitigated stress-induced impairments in memory inference, two-way ANOVAs were conducted on memory inference performance, with group (stress vs. control) and strategy (Study 1: no retrieval practice vs. retrieval practice) as between-subjects factors. Additionally, an independent t-test was performed in Study 1 to isolate the effect of acute stress on memory inference performance (i.e., whether acute stress impaired memory inference). Another independent t-test was conducted in Study 2 to compare the stress and control groups, aiming to determine whether retrieval practice could prevent stress-induced inference impairments. Furthermore, individual differences in memory inference were correlated with acute stress response indices and memory reactivation indices to establish a comprehensive model incorporating physiological, neural, and cognitive measures.

5.3 Analyses of Decoding Accuracy Maps

We employed two approaches to analyze the generated decoding accuracy maps. The first analysis focused on determining whether mean classification accuracy exceeded the chance level of 50%. We applied permutation-based cluster-level correction to identify time windows with significant decoding accuracy, thereby suggesting the optimal training features and the timing of memory trace reactivation. Specifically, a cluster-based permutation test was applied to the 2D spatial data, which is well-suited for analyzing spatially correlated data, such as neuroimaging and EEG results. Initially, a one-sample t-test was performed for each spatial point in the 2D grid, comparing the results to chance level. A one-sided test with a p-value threshold of 0.05 identified significant points, which were assigned a value of 1,

while non-significant points were marked as 0. Significant points were then grouped into clusters of adjacent significant points based on spatial connectivity. To assess cluster significance, a permutation procedure was applied, where the data were shuffled, and t-statistics were recalculated for 5000 iterations. The maximum t-statistic from each iteration was recorded, and clusters with observed t-statistics exceeding the 95th percentile of the permutation distribution were considered significant. The final output was a 2D matrix of binary values indicating the significance of each spatial point, ensuring robust control over Type I errors and identifying spatially contiguous regions of significant clusters.

For comparisons between experimental conditions, such as stress and retrieval practice, we employed an averaging approach to avoid the multiple comparison problem associated with comparing each time point individually. In this approach, decoding accuracy within the optimal time window was averaged. The averaging process followed these steps: (1) Identifying the optimal training time window: We first analyzed decoding accuracy over time for each subject by identifying clusters of time points where decoding accuracy exceeded a predefined threshold (50%) for a minimum duration of 10 ms. The peak latency was determined by calculating the center-of-mass of the cluster with the highest summed accuracy, and peak latencies were estimated for each subject. Group-level analyses were then conducted to calculate the mean and standard deviation of peak latencies, providing the optimal time window for classifier training. At the group level (i.e., across all 128 participants), the peak latency was located at 207.17 ms, with a standard deviation of 228.17 ms. (2) Validating the optimal training time window through cross-task decoding: We employed a time-generalization variant of MVPA to account for potential differences in

neural processing across tasks. This method generated classifier performance across all possible time-point combinations for both the localizer and inference tasks. The optimal time window was independently estimated from the cross-task classification decoding matrix. Using the same procedure as in Step 1, we estimated the optimal training time window for capturing neural reactivation as 230.48 ms (SD = 139.81 ms), which did not differ significantly from the time window estimated during model training ($t = 1.39$, $p = 0.16$), suggesting no delayed or accelerated neural processing during memory reactivation. Thus, we used the peak latency at 207.17 ms, with a 50 ms range, to define the training time window for subsequent statistical analyses.

5.4 Linking Neural Reactivation to Subsequent Inference Effect, Stress, and Retrieval Practice

Each participant generated three decoding accuracy matrices: (1) all trials, (2) trials with correct inferences (successful inferences), and (3) trials with incorrect inferences (unsuccessful inferences). Since each matrix is a high-dimensional structure containing over 200 × 450 accuracy outputs, we applied dimensionality reduction by averaging the values within the identified optimal training time window prior to further analysis. To quantify individual differences in memory reactivation within this window, we defined two indices representing the temporal and strength aspects of memory reactivation: *decodability* and *mean decoding accuracy*. Decodability was calculated by determining the percentage of time points with decoding accuracy greater than chance within the specified time window. This index ranges from 0 to 1, with higher values indicating greater memory reactivation across time points. From a cognitive processing perspective, higher decodability suggests sustained

memory reactivation during the inference phase, reflecting the temporal aspect of reactivation. In contrast, mean decoding accuracy represents the average decoding accuracy and ranges from 0 to 50, as raw accuracy was corrected for chance and scaled by 100. Higher mean decoding accuracy indicates more accurate decoding of bridging information from brain activity, which may suggest stronger underlying memory representations.

To compare the neural indices of memory reactivation across experimental conditions, we defined the subsequent inference effect as the comparison of mean decoding accuracy and decodability between successful and unsuccessful inference trials, analogous to the subsequent memory effect, which compares neural activity between correctly and incorrectly recalled trials. Because successful and unsuccessful trials were compared within participants, a paired t-test was conducted on the two memory reactivation indices with a threshold of $p = 0.05$ (one-tailed), hypothesizing that memory reactivation would be stronger for successful trials. Additionally, we compared the two neural indices between Experiment 1 and Experiment 2 to examine the general effects of retrieval practice on memory reactivation, between stress and control participants in Experiment 1 to assess the effect of acute stress on memory reactivation, and between stress and control participants in Experiment 2 to investigate the protective effect of retrieval practice on memory reactivation under stress. These analyses were performed using independent t-tests with a threshold of $p = 0.05$ (one-tailed), with hypotheses specified in the Results section.

5. Code availability

The custom code used in this study is available in https://osf.io/v54g8/?view_only=dd87e13d50b04c30b9ed839c58b80d95

6. Data availability

The data and that support the findings of this study are in

https://osf.io/v54g8/?view_only=dd87e13d50b04c30b9ed839c58b80d95

Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments and Funding

This study was supported by the National Natural Science Foundation of China (grant No. 32300879), Humanities and Social Sciences Fund, Ministry of Education (grant No. 22YJCZH109), Natural Science Foundation of Hubei (grant No. 2022CFB793), Open Research Fund of the State Key Laboratory of Cognitive Neuroscience and Learning (CNLYB2103), Open Research Fund of the Key Laboratory of Adolescent Cyber Psychology and Behavior (grant No. CCNUCYPSYLAB2022B10), the Major Program of the National Social Science Foundation of China (grant No. 22&ZD187), self-determined research funds of CCNU from the colleges' basic research and operation of MOE(CCNU22QN020), and National Natural Science Foundation of China (62077010)

Contributions

W.L, X.J.S and J.P.G conceived the Study; J.P.G, R.X.C, Q.Z collected the data; W.L. and J.P.G analyzed the data; W.L. and J.P.G prepared the first draft. W.L, X.J.S and J.P.G reviewed and edited the manuscript. W.L and X.J.S provided supervision and obtained funding.

Reference

1. Vogel, S. & Schwabe, L. Learning and memory under stress: implications for the classroom. *NPJ Sci Learn* **1**, (2016).
2. Schlichting, M. L. & Preston, A. R. Memory integration: neural mechanisms and implications for behavior. *Curr Opin Behav Sci* **1**, 1–8 (2015).
3. Preston, A. R. & Eichenbaum, H. Interplay of Hippocampus and Prefrontal Cortex in Memory. *Current Biology* **23**, R764–R773 (2013).
4. Backus, A. R., Schoffelen, J.-M., Szebényi, S., Hanslmayr, S. & Doeller, C. F. Hippocampal-Prefrontal Theta Oscillations Support Memory Integration. *Current Biology* **26**, 450–457 (2016).
5. Zeithamova, D., Dominick, A. L. & Preston, A. R. Hippocampal and Ventral Medial Prefrontal Activation during Retrieval-Mediated Learning Supports Novel Inference. *Neuron* **75**, 168–179 (2012).
6. Zeithamova, D. & Preston, A. R. Flexible Memories: Differential Roles for Medial Temporal Lobe and Prefrontal Cortex in Cross-Episode Binding. *The Journal of Neuroscience* **30**, 14676–14684 (2010).
7. Cohen, J. D. *et al.* Computational approaches to fMRI analysis. *Nat Neurosci* **20**, 304–313 (2017).
8. Schlichting, M. L., Guarino, K. F., Roome, H. E. & Preston, A. R. Developmental differences in memory reactivation relate to encoding and inference in the human brain. *Nat Hum Behav* **6**, 415–428 (2021).
9. Morton, N. W., Schlichting, M. L. & Preston, A. R. Representations of common event structure in medial temporal lobe and frontoparietal cortex support efficient inference. *Proceedings of the National Academy of Sciences* **117**, 29338–29345 (2020).
10. Liu, J., Harris, A. & Kanwisher, N. Stages of processing in face perception: an MEG study. *Nat Neurosci* **5**, 910–916 (2002).
11. Kanwisher, N., McDermott, J. & Chun, M. M. The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *The Journal of Neuroscience* **17**, 4302–4311 (1997).
12. Linde-Domingo, J., Treder, M. S., Kerrén, C. & Wimber, M. Evidence that neural information flow is reversed between object perception and object reconstruction from memory. *Nat Commun* **10**, 179 (2019).
13. Cichy, R. M., Pantazis, D. & Oliva, A. Resolving human object recognition in space and time. *Nat Neurosci* **17**, 455–462 (2014).
14. Wolf, O. T. Stress and memory retrieval: mechanisms and consequences. *Curr Opin Behav Sci* **14**, 40–46 (2017).
15. Kuhlmann, S., Piel, M. & Wolf, O. T. Impaired Memory Retrieval after Psychosocial Stress in Healthy Young Men. *The Journal of Neuroscience* **25**, 2977–2982 (2005).
16. Smeets, T. Acute stress impairs memory retrieval independent of time of day. *Psychoneuroendocrinology* **36**, 495–501 (2011).
17. de Quervain, D. J.-F., Roozendaal, B. & McGaugh, J. L. Stress and glucocorticoids impair retrieval of long-term spatial memory. *Nature* **394**, 787–790 (1998).

18. Gagnon, S. A., Waskom, M. L., Brown, T. I. & Wagner, A. D. Stress Impairs Episodic Retrieval by Disrupting Hippocampal and Cortical Mechanisms of Remembering. *Cerebral Cortex* **29**, 2947–2964 (2019).
19. Buchanan, T. W., Tranel, D. & Adolphs, R. Impaired memory retrieval correlates with individual differences in cortisol response but not autonomic response. *Learning & Memory* **13**, 382–387 (2006).
20. Qin, S., Hermans, E. J., van Marle, H. J. F., Luo, J. & Fernández, G. Acute Psychological Stress Reduces Working Memory-Related Activity in the Dorsolateral Prefrontal Cortex. *Biol Psychiatry* **66**, 25–32 (2009).
21. Yuen, E. Y. *et al.* Acute stress enhances glutamatergic transmission in prefrontal cortex and facilitates working memory. *Proceedings of the National Academy of Sciences* **106**, 14075–14079 (2009).
22. Gärtner, M., Rohde-Liebenau, L., Grimm, S. & Bajbouj, M. Working memory-related frontal theta activity is decreased under acute stress. *Psychoneuroendocrinology* **43**, 105–113 (2014).
23. Oei, N. Y. L. *et al.* Glucocorticoids Decrease Hippocampal and Prefrontal Activation during Declarative Memory Retrieval in Young Men. *Brain Imaging Behav* **1**, 31–41 (2007).
24. Hermans, E. J., Henckens, M. J. A. G., Joëls, M. & Fernández, G. Dynamic adaptation of large-scale brain networks in response to acute stressors. *Trends Neurosci* **37**, 304–314 (2014).
25. Hermans, E. J. *et al.* Stress-Related Noradrenergic Activity Prompts Large-Scale Neural Network Reconfiguration. *Science (1979)* **334**, 1151–1153 (2011).
26. Schwabe, L., Hermans, E. J., Joëls, M. & Roozendaal, B. Mechanisms of memory under stress. *Neuron* vol. 110 1450–1467 Preprint at https://doi.org/10.1016/j.neuron.2022.02.020 (2022).
27. Brown, T. I., Gagnon, S. A. & Wagner, A. D. Stress Disrupts Human Hippocampal-Prefrontal Function during Prospective Spatial Navigation and Hinders Flexible Behavior. *Current Biology* **30**, 1821-1833.e8 (2020).
28. Smith, A. M. & Thomas, A. K. Reducing the consequences of acute stress on memory retrieval. *J Appl Res Mem Cogn* **7**, 219–229 (2018).
29. Schwabe, L. *et al.* Stress effects on declarative memory retrieval are blocked by a β-adrenoceptor antagonist in humans. *Psychoneuroendocrinology* **34**, 446–454 (2009).
30. de Quervain, D. J.-F., Aerni, A. & Roozendaal, B. Preventive Effect of β-Adrenoceptor Blockade on Glucocorticoid-Induced Memory Retrieval Deficits. *American Journal of Psychiatry* **164**, 967–969 (2007).
31. Bogdanov, M. & Schwabe, L. Transcranial Stimulation of the Dorsolateral Prefrontal Cortex Prevents Stress-Induced Working Memory Deficits. *Journal of Neuroscience* **36**, 1429–1437 (2016).
32. Roediger, H. L. & Butler, A. C. The critical role of retrieval practice in long-term retention. *Trends Cogn Sci* **15**, 20–27 (2011).
33. Karpicke, J. D. & Blunt, J. R. Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping. *Science (1979)* **331**, 772–775 (2011).

34. Smith, A. M., Floerke, V. A. & Thomas, A. K. Retrieval practice protects memory against acute stress. *Science (1979)* **354**, 1046–1048 (2016).
35. Antony, J. W., Ferreira, C. S., Norman, K. A. & Wimber, M. Retrieval as a Fast Route to Memory Consolidation. *Trends Cogn Sci* **21**, 573–576 (2017).
36. Liu, W., Kohn, N. & Fernández, G. Probing the neural dynamics of mnemonic representations after the initial consolidation. *Neuroimage* **221**, 117213 (2020).
37. Zhuang, L. *et al.* Rapid neural reorganization during retrieval practice predicts subsequent long-term retention and false memory. *Nat Hum Behav* **6**, 134–145 (2021).
38. Ye, Z., Shi, L., Li, A., Chen, C. & Xue, G. Retrieval practice facilitates memory updating by enhancing and differentiating medial prefrontal cortex representations. *Elife* **9**, (2020).
39. De Quervain, D. J. -F. *et al.* Glucocorticoid-induced impairment of declarative memory retrieval is associated with reduced blood flow in the medial temporal lobe. *European Journal of Neuroscience* **17**, 1296–1302 (2003).
40. Xue, G. The Neural Representations Underlying Human Episodic Memory. *Trends in Cognitive Sciences* vol. 22 544–561 Preprint at https://doi.org/10.1016/j.tics.2018.03.004 (2018).
41. Dijkstra, N., Bosch, S. E. & van Gerven, M. A. J. Shared Neural Mechanisms of Visual Perception and Imagery. *Trends in Cognitive Sciences* vol. 23 423–434 Preprint at https://doi.org/10.1016/j.tics.2019.02.004 (2019).
42. Chadwick, M. J., Hassabis, D., Weiskopf, N. & Maguire, E. A. Decoding Individual Episodic Memory Traces in the Human Hippocampus. *Current Biology* **20**, 544–547 (2010).
43. Polyn, S. M., Natu, V. S., Cohen, J. D. & Norman, K. A. Category-Specific Cortical Activity Precedes Retrieval During Memory Search. *Science (1979)* **310**, 1963–1966 (2005).
44. Xie, S., Kaiser, D. & Cichy, R. M. Visual Imagery and Perception Share Neural Representations in the Alpha Frequency Band. *Current Biology* **30**, 2621-2627.e5 (2020).
45. Dijkstra, N., Mostert, P., Lange, F. P. de, Bosch, S. & van Gerven, M. A. Differential temporal dynamics during visual imagery and perception. *Elife* **7**, (2018).
46. Schwabe, L., Joëls, M., Roozendaal, B., Wolf, O. T. & Oitzl, M. S. Stress effects on memory: An update and integration. *Neurosci Biobehav Rev* **36**, 1740–1749 (2012).
47. Roediger, H. L. & Abel, M. The double-edged sword of memory retrieval. *Nature Reviews Psychology* **1**, 708–720 (2022).
48. Shields, G. S., Sazma, M. A., McCullough, A. M. & Yonelinas, A. P. The effects of acute stress on episodic memory: A meta-analysis and integrative review. *Psychol Bull* **143**, 636–675 (2017).
49. Kim, J. J., Lee, H. J., Han, J.-S. & Packard, M. G. Amygdala Is Critical for Stress-Induced Modulation of Hippocampal Long-Term Potentiation and Learning. *The Journal of Neuroscience* **21**, 5222–5228 (2001).
50. Elliott, A. E. & Packard, M. G. Intra-amygdala anxiogenic drug infusion prior to retrieval biases rats towards the use of habit memory. *Neurobiol Learn Mem* **90**, 616–623 (2008).
51. Knowlton, B. J., Mangels, J. A. & Squire, L. R. A Neostriatal Habit Learning System in Humans. *Science (1979)* **273**, 1399–1402 (1996).

52. Zerbes, G., Kausche, F. M. & Schwabe, L. Stress-induced cortisol modulates the control of memory retrieval towards the dorsal striatum. *European Journal of Neuroscience* **55**, 2699–2713 (2022).
53. Raio, C. M., Hartley, C. A., Orederu, T. A., Li, J. & Phelps, E. A. Stress attenuates the flexible updating of aversive value. *Proceedings of the National Academy of Sciences* **114**, 11241–11246 (2017).
54. Kluen, L. M., Agorastos, A., Wiedemann, K. & Schwabe, L. Noradrenergic Stimulation Impairs Memory Generalization in Women. *J Cogn Neurosci* **29**, 1279–1291 (2017).
55. Kluen, L. M., Nixon, P., Agorastos, A., Wiedemann, K. & Schwabe, L. Impact of Stress and Glucocorticoids on Schema-Based Learning. *Neuropsychopharmacology* **42**, 1254–1261 (2017).
56. Vogel, S., Kluen, L. M., Fernández, G. & Schwabe, L. Stress affects the neural ensemble for integrating new information and prior knowledge. *Neuroimage* **173**, 176–187 (2018).
57. Arnsten, A. F. T. Stress signalling pathways that impair prefrontal cortex structure and function. *Nat Rev Neurosci* **10**, 410–422 (2009).
58. Yang, C., Luo, L., Vadillo, M. A., Yu, R. & Shanks, D. R. Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychol Bull* **147**, 399–435 (2021).
59. Carpenter, S. K., Pan, S. C. & Butler, A. C. The science of effective learning with spacing and retrieval practice. *Nature Reviews Psychology* **1**, 496–511 (2022).
60. Roediger, H. L. & Karpicke, J. D. Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. *Psychol Sci* **17**, 249–255 (2006).
61. Fendos, J. Anatomy Terminology Performance is Improved by Combining Jigsaws, Retrieval Practice, and Cumulative Quizzing. *Anat Sci Educ* **14**, 641–657 (2021).
62. Carpenter, S. K. & Kelly, J. W. Tests enhance retention and transfer of spatial learning. *Psychon Bull Rev* **19**, 443–448 (2012).
63. Lutz, N. D., Martínez-Albert, E., Friedrich, H., Born, J. & Besedovsky, L. Sleep shapes the associative structure underlying pattern completion in multielement event memory. *Proceedings of the National Academy of Sciences* **121**, (2024).
64. Wing, E. A., Marsh, E. J. & Cabeza, R. Neural correlates of retrieval-based memory enhancement: An fMRI study of the testing effect. *Neuropsychologia* **51**, 2360–2370 (2013).
65. Dragoi, G. & Tonegawa, S. Distinct preplay of multiple novel spatial experiences in the rat. *Proceedings of the National Academy of Sciences* **110**, 9100–9105 (2013).
66. Dragoi, G. & Tonegawa, S. Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature* **469**, 397–401 (2011).
67. Bäuml, K.-H. T. & Trißl, L. Selective memory retrieval can revive forgotten memories. *Proceedings of the National Academy of Sciences* **119**, (2022).
68. Hebart, M. N. *et al.* THINGS: A database of 1,854 object concepts and more than 26,000 naturalistic object images. *PLoS One* **14**, e0223792 (2019).
69. Lu, Z. & Ku, Y. NeuroRA: A Python Toolbox of Representational Analysis From Multi-Modal Neural Data. *Front Neuroinform* **14**, (2020).