# Reactivating Positive Personality Traits During Sleep Impacts Self-Evaluative Memories

Ziqing Yao[1], Tao Xia[1], Zhiguo Zhang[2], Xuanyi Lin[1],

Pengmin Qin [3], Yina Ma[4,5] & Xiaoqing Hu[1,6,7*]

1 Department of Psychology, The University of Hong Kong, Hong Kong SAR, China
2 Institute of Computing and Intelligence, Harbin Institute of Technology, Shenzhen, China,
Peng Cheng Laboratory, Shenzhen, Guangdong, 518055, China,
Marshall Laboratory of Biomedical Engineering, Shenzhen, Guangdong, 518060, China
3 Key Laboratory of Brain, Cognition and Education Sciences, Ministry of Education,
School of Psychology, Center for Studies of Psychological Application, and
Guangdong Key Laboratory of Mental Health and Cognitive Science, South China
Normal University, Guangzhou, Guangdong 510631, China
4 State Key Laboratory of Cognitive Neuroscience and Learning, IDG/McGovern
Institute for Brain Research, Beijing Key Laboratory of Brain Imaging and Connectomics,
Beijing Normal University, Beijing, China
5 Chinese Institute for Brain Research, Beijing, China
6 The State Key Laboratory of Brain and Cognitive Sciences,
The University of Hong Kong, Hong Kong SAR, China
7 HKU, Shenzhen Institute of Research and Innovation, Shenzhen, China

Corresponding Author:
Xiaoqing Hu,
Department of Psychology, The University of Hong Kong,
Pokfulam, Hong Kong SAR, China.
E-mail: xiaoqinghu@hku.hk
Phone: (852) 3917-2291

## Abstract

People tend to view themselves through rose-tinted glasses, as evidenced by preferential recall of positive personality traits. We asked whether reactivating positive personality traits during sleep could enhance peoples' positive self-evaluative memories. After a baseline self-referential encoding task in which participants endorsed positive and negative traits as self-descriptive, participants were trained to give timely responses to positive traits in a cue-approach training (CAT) task. Once participants had entered slow-wave sleep during a subsequent nap, half of the trained positive traits were unobtrusively re-played to them to promote consolidation (targeted memory reactivation, TMR). Participants completed free-recall tasks about self-descriptive traits to measure their self-evaluative memories. Our findings revealed that TMR prioritized the recall of positive traits that were strongly memorized before sleep, while impairing the recall of intermediate traits. The results suggest pre-TMR self-evaluative memory strength modulated the TMR benefits. Sleep EEG analyses revealed that compared with weak/intermediate/control traits, re-playing strongly memorized traits during sleep elicited greater sigma power changes, which likely reflect preferential memory reactivation. Our results demonstrate the potential implication of wakeful cue-approach training and sleep-based memory reactivation in strengthening positive self-evaluative memories.

***Keywords: targeted memory reactivation, sleep, cue-approach training, self-evaluation, positivity bias***

## Introduction

People view themselves through rose-tinted lenses; when presented with positive and negative personality traits, people not only tend to endorse positive traits more frequently, they also preferentially remember positive rather than negative traits[1–4]. This positive self-evaluative memory bias[5] may safeguard mental wellness, thereby reducing feelings of worthlessness and depressive symptoms in individuals especially when faced with self-threatening information[6]. Despite its importance[7], few studies have examined how positive self-evaluative memories can be enhanced. To expand this seldom-explored area of research, we hypothesized that one strategy to enhance positive self-evaluative memories is to increase the salience and memory accessibility of positive traits when people describe themselves. To this end, we tested the efficacy of two recently developed paradigms to enhance positive self-evaluative memories: (1) a cued-approach training (CAT[8]) task that increases stimulus salience via approach motor training, and (2) sleep-based targeted memory reactivation (TMR[9]) that promotes memory consolidation during post-training sleep.

CAT is a behavioral training task used to induce choice and preference changes for given stimuli in the absence of external reinforcement[8,10,11]. Specifically, when people are cued to manually respond to a stimulus (e.g., ‘Go’), they are more likely to prefer and more frequently choose Go stimuli over NoGo stimuli despite having comparable initial preferences for Go and NoGo stimuli. Previous research has used CAT to change individuals’ evaluation and choices of low-level stimuli such as snacks, meaningless fractals, and positive images from the International Affective Picture System (IAPS)[10]. However, it remains unknown whether CAT can change individuals’ evaluation of high-level personality traits. Inspired by CAT research, and to address this gap, we examined whether prompting participants to respond to positive self-evaluative traits (e.g., ‘brilliant’, ‘conscientious’, ‘brave’) could enhance positive self-evaluative memories and self-evaluation.

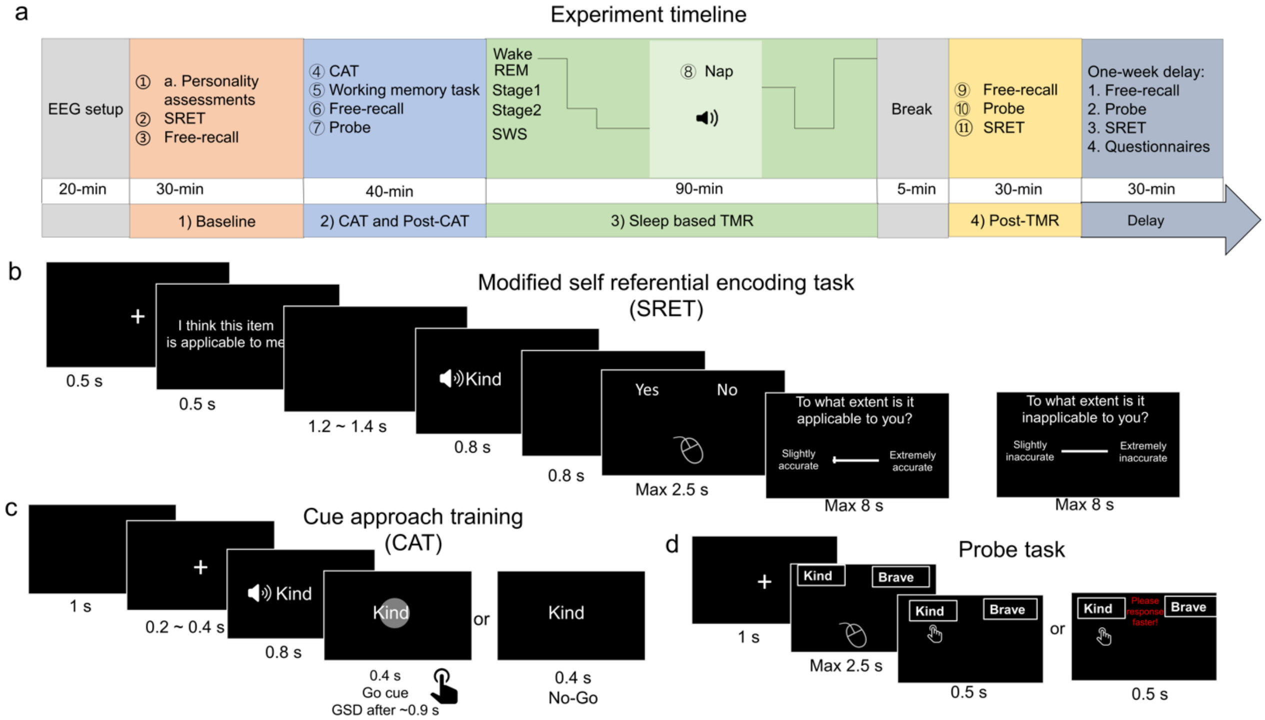
Sleep is pivotal in memory consolidation[12,13]. A plethora of research suggests that post-learning sleep contributes to the stabilization of newly acquired memories, mediated by repeated memory reactivation particularly during non-rapid eye movement (NREM) sleep [9,13–15]. Notably, memory reactivation not only occurs spontaneously, but can also be initiated via re-playing memory-related sensory cues to individuals during post-learning NREM sleep. TMR has been shown to promote memory consolidation for cued memories, as evidenced by improved cued (versus uncued) memory performance in post-sleep tests [16–22] (see Hu et al.[23] for a meta-analysis on TMR). However, it remains unclear whether TMR can consolidate highly self-referential memories such as personality traits. The present study aimed to investigate whether TMR could enhance positive self-evaluative memories via repeatedly reactivating positive-personality traits during sleep.

While mounting evidence suggests that sleep and TMR promote memory consolidation, not all memories are equally benefited [18,24,25]. A range of factors influence sleep-mediated memory consolidation and TMR effect, such as pre-TMR memory strength (strong or weak) and motivational salience (high versus low arousal, reward versus non-reward) [26–29]. Relatedly, in self-evaluation, people tend to have stronger memories of highly self-descriptive traits compared to less self-descriptive items[30,31]. Given that highly self-descriptive traits are preferentially encoded and that they are highly motivationally salient, we hypothesized that TMR benefits would be more evident among positive traits that were strongly endorsed and better memorized than traits that were poorly memorized or less likely to be endorsed before TMR.

We ask whether we could enhance participants' positive self-evaluative memories via reactivating CAT-trained, positive trait memories during sleep (see Figure. 1a). In a self-referential encoding task (SRET), participants viewed a series of positive and negative personality traits and rated the extent to which they could be applied to themselves, ranging from "extremely inaccurate" to "extremely accurate" (see Figure. 1b). In the subsequent CAT, participants manually responded to certain positive traits, presented visually on a screen and aurally through speakers, see Figure. 1c). Our key experimental manipulation happened during the post-CAT nap phase; during slow-wave sleep, the experimenter re-played half the previous CAT-trained positive traits to participants to reactivate memories. Upon waking, participants completed a free-recall task to assess the TMR's impact on their self-evaluative memories as well as a probe task to assess its effect on self-evaluative preferences. In addition to assessing TMR's immediate effect, we invited participants to come back to the lab one week after to examine TMR's potential long-term effects.

**Figure 1**

*An Overview of Experimental Design and Main Tasks*



*Note.* (a) Task flow demonstrating the four task phases used following arrival, consent-form signing, and EEG setup, including baseline tests (phase 1), CAT and post-CAT/pre-TMR tests (phase 2), sleep-based TMR (phase 3), and post-TMR tests (phase 4). (b) Modified SRET, in which participants rated how accurately specific traits reflected themselves, followed by a self-evaluative memory free-recall task. (c) Exemplar trial of CAT, in which participants responded to positive traits presented visually and aurally. (d) Probe test, in which participants were presented with Go and NoGo trait-word pairings and asked to select the word from the pair that was most self-descriptive (see Methods section for a full description of the procedure and experimental tasks).

## Results

### Behavioral results

We measured the following outcomes related to self-evaluation: (1) self-evaluative memories, via free-recall of personality trait in the free-recall task; (2) self-evaluative preferences, via preference choices among two traits in the probe task; and (3) self-evaluative judgements, via endorsement judgements and ratings of personality traits in the SRET.

### *Evidence of Self-Evaluative Positivity Bias at Baseline Session*

Consistent with previous research that examined self-evaluation in healthy populations [32], participants endorsed more positive than negative traits as self-descriptive ($t$ = -11.1 $p < 0.001$) and were faster at endorsing positive traits ($t$ = 2.30 $p$ = 0.028).

In the baseline recall task (1st recall), participants recalled more positive than negative traits, indicating a positive self-evaluative memory bias ($t$ = 8.42, $p < 0.001$). Participants who showed a larger positivity bias also reported a lower score from the Beck Depression Inventory-II (BDI) ($r$ = -0.68, $p < 0.001$; Figure. 2a), replicating the relationship between self-positivity bias and depressive symptoms[33].

Together, via successfully replicating the self-evaluation positivity bias, we confirmed the validity of using SRET and the trait free-recall task to assess self-evaluation and self-evaluative memories.

### *Effects of CAT and TMR on Self-Evaluative Memories*

Given that we primarily focused on positive self-evaluation and self-evaluative memories, we presented results on positive traits using (generalized) linear mixed models ((G)LMM) to include participants and/or traits as random effects during post-CAT, post-TMR, and delay tasks.

To test the CAT effect (i.e., Go versus NoGo) on the recall of positive self-evaluative memories while controlling for baseline recall performance, we ran (G)LMM using CAT (Go versus NoGo) and pre-CAT recall (recalled versus not recalled) as fixed effect, using participant factor as the random effect to predict post-CAT recall. We analyzed whether a trait was recalled or not, and if recalled, its recall order (the earlier, the better). Results showed a significant interaction between baseline recall (1st recall) and CAT on post-CAT recall order, $F$ (1, 1073.7) = 4.04, $p = 0.045$. However, the post-hoc comparison did not reach significance. No other significant effects were found for recall percentage during the post-CAT recall task, all $ps > 0.1$. These results suggest CAT alone might not be capable of changing self-evaluative memories.

Next, we focused on the TMR effect (cued vs. uncued) by examining post-TMR recall performance (3rd recall), again using recall outcome and recall order. To understand how pre-sleep memory strength might influence TMR effects, we categorized each trait into one of three conditions (weak, intermediate, and strong) based on recall from the baseline (1st) and post-CAT (2nd) tasks. A trait would be categorized as weak if it was not recalled in neither baseline and post-CAT tests, intermediate if it was recalled only once in either baseline or post-CAT tests, and strong if it was recalled in both tests. Because the categorization of traits depended on the recall performance of baseline and post-CAT tests, 17 participants were excluded from subsequent analyses due to lack of traits assigned to one or more memory-strength conditions. This left us with 18 participants. Given the small sample size, we regarded the following analyses as exploratory, and the results should be interpreted with caution. The number of traits assigned to weak, intermediate, and strong conditions are presented in Table 1.

**Table 1**

Number of traits in each pre-memory strength condition (mean ±SEM).

| **Cue condition** | **Weak** | **Intermediate** | **Strong** |
|---|---|---|---|
| Go-cued | 10.16 ± 0.35 | 3.00 ± 0.29 | 1.83 ± 0.22 |
| Go-uncued | 9.88 ± 0.37 | 3.22 ± 0.34 | 1.88 ± 0.23 |
| NoGo | 21.61 ± 0.52 | 4.83 ± 0.40 | 3.56 ± 0.45 |

For the binary recall outcome (i.e., recalled or not), we found a significant interaction effect between TMR and pre-TMR memory strength, $\chi 2(4) = 12.28$, $p = 0.015$. Post-hoc results showed that cueing intermediate traits impaired their later recall, $\beta = -1.15$, $SE = 0.45$, $p = 0.027$, but there was no significant cueing effect for strong traits, $\beta = 1.34$, $SE = 0.75$, $p = 0.17$, or weak traits, $\beta = 0.30$, $SE = 0.37$, $p = 0.70$ (see Figure. 2b). No other significant contrasts were found, *p*s > 0.1.

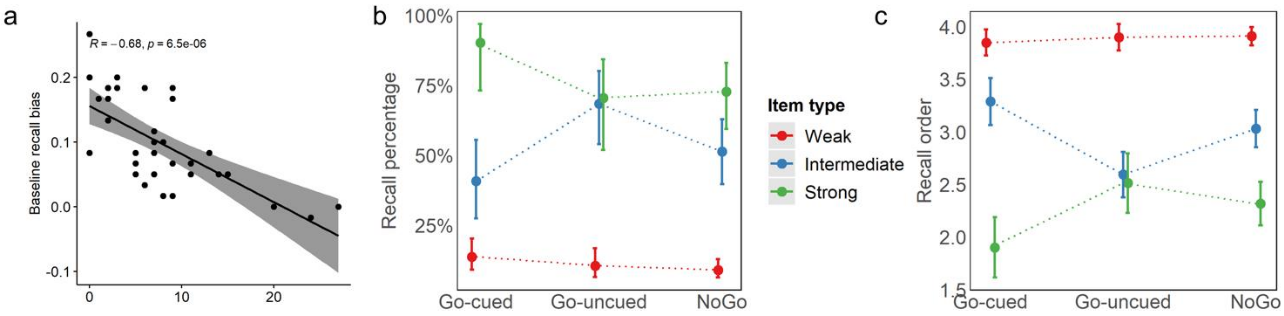
For recalled order, we found a significant interaction between TMR and pre-TMR memory strength, $F$ (4, 1065.7) = 7.22, $p < 0.001$. Follow-up analyses showed that TMR cueing prioritized the recall of strong traits, as evidenced by earlier recall order for Go-cued rather than Go-uncued traits, $\beta = -0.61$, $SE = 0.20$, $p < 0.01$, and also marginally than NoGo items, $\beta = -0.42$, $SE = 0.18$, $p = 0.053$ (see Figure. 2c). In contrast, for intermediate traits, TMR cueing rendered the Go-cued traits to be recalled later than Go-uncued traits, $\beta = 0.69$, $SE = 0.16$, $p < 0.001$, while Go-uncued traits were recalled earlier than NoGo traits, $\beta = -0.44$, $SE = 0.14$, $p < 0.01$. No other significant results were found (*p*s > 0.1). Thus, TMR prioritized the accessibility of strong traits at the cost of intermediate traits, as evidenced by their recall priority. Regarding the delay test, while it is worthwhile to examine TMR's long-term effect in the one-week delay tests, we were unable to perform the analyses given the number of traits that were recalled in the delay task was too few.

Lastly, to ensure that pre-TMR memory strength aligned with self-referential processing, we compared the baseline endorsement rating (obtained from the baseline SRET, see Figure. 1b) for weak, intermediate, and strong traits. We found that strong and intermediate traits had significantly higher endorsement ratings (i.e., more self-descriptive) than weak traits (all *ps* < 0.01), but no difference between intermediate and strong traits ($p = 0.596$). This suggests pre-

TMR memory strength reflected the extent to which participants would endorse traits as self-descriptive.

**Figure 2**

*The effect of TMR Conditions, Pre-TMR Memory Strength on Post-TMR Recall Performance*



*Note.* (a) Pearson Correlation Between BDI Scores and Baseline Self-Positivity Bias. Baseline recall bias was calculated by the number of endorsed and recalled positive words divided by the number of endorsed words during the baseline recall task. BDI scores were negatively correlated with baseline recall bias. (b) Recall percentage for positive traits. (c) Recall order (via a log transformation). Error bars indicate 95% confidence intervals (CIs).

### *Effects of CAT and TMR on Self-Evaluation Preferences*

We next assessed self-evaluation preferences towards positive traits using the probe task (see Figure. 1d). On each trial of the probe task, participants were presented with a pair of traits, each containing one Go trait and one NoGo trait (with Go and NoGo traits matched for baseline endorsement level) and were asked to select the trait which better described themselves. More specifically, we followed previous CAT research to sort Go–NoGo (GNG) pairs into low- and high-rating pairs [8,10]. We then used a series of GLMMs to analyze preference changes (see Methods). Immediately after CAT, participants were more likely to choose Go traits for low-rating GNG pairs (Go choices: 55.7%, SE = 2.2%, Z = 2.60, $p$ = 0.005) but not for high-rating GNG pairs (mean proportion = 51.5%, SE = 2.3%, Z = 0.66, $p$ = 0.254). The percentage of Go trait choices was not significantly different between high- versus low-rating pairs (Z = 1.55, *p* = 0.120).

We next examined the immediate effect of TMR in the post-TMR probe task. We found no TMR effects (i.e., choice proportion Go-cued > Go-uncued traits) for either low- or high-rating pairs. Participants were more likely to choose Go-cued and Go-uncued traits than NoGo traits for low-rating pairings: Go-cued: 55.5%, SE = 3.1%, Z = 1.81, *p* = 0.035; Go-uncued: 56.9%, SE = 3.1%, Z = 2.19, *p* = 0.014. This effect was not significant for high-rating pairings: Go-cued: 50.6%, SE = 3.3%, Z = 0.184, *p* = 0.427; Go-uncued 54.2%, SE = 4.0%, Z = 0.99, *p* = 0.162. There were no significant differences between high- and low-rating traits for Go-cued (Z = 1.31, *p* = 0.190) or Go-uncued items (Z = 0.74, *p* = 0.462).

After a one-week delay, participants significantly preferred Go-cued over NoGo traits in high-rating pairings, 56.2%, SE = 2.6%, Z = 2.30, *p* = 0.01, with a marginally significant effect in low-rating pairings, 55.3%, SE = 3.2%, Z = 1.61, *p* = 0.053. The choices of Go-cued traits were not different between high- versus low-rating pairings (Z = -0.24, *p* = 0.810). No choice preferences were found for Go-uncued over NoGo pairings nor between Go-cued and Go-uncued traits (all *p*s > 0.2). Together, these results suggest only Go-cued traits (i.e., TMR) showed a prolonged CAT effect on both high- and low-rating pairs comparably (Figure. 3a).

### *Effects of CAT and TMR on Self-Evaluative Judgements*

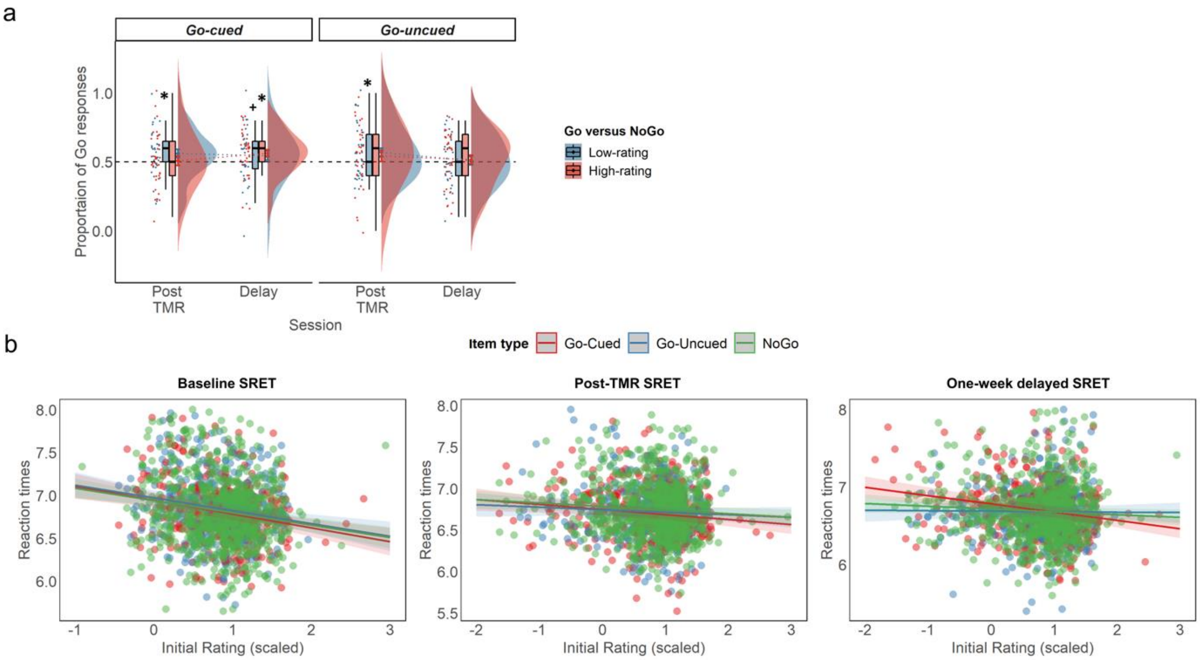
To analyze changes in positive self-evaluations, we ran a GLMM using TMR/CAT conditions (Go-cued, Go-uncued, and NoGo) and test sessions (pre-, post-TMR, and delay) as fixed effects to predict binary endorsement. We found no interactions between conditions and test sessions,

$\chi 2(4) = 1.89$, $p = 0.756$, indicating a null effect of CAT and TMR training on self-evaluative judgements.

Given that reaction times (RT) of choices could also infer preferences[34], we next ran a LMM using TMR/CAT conditions and test sessions to predict RTs. There was a significant interaction between the initial rating and TMR condition only in the delay session, $F$ (2, 1,644.2) = 5.06, $p < 0.01$ (Figure. 3b). Post-hoc comparisons showed a significantly greater negative correlation for Go-cued rather than Go-uncued positive traits, $\beta = -0.10$, $SE = 0.03$, $p < 0.01$, or NoGo traits, $\beta = -0.07$, $SE = 0.03$, $p = 0.034$, suggesting higher baseline endorsement levels would lead to faster RTs during the delay SRET for Go-cued traits. No differences were found between Go-uncued and NoGo traits, $\beta = 0.03$, $SE = 0.03$, $p = 0.585$. These negative associations were not found in previous behavioral patterns when only CAT was used (see SOM, Figure. S11), suggesting TMR uniquely promoted endorsement speed for high self-descriptive traits.

**Figure 3**

*Behavioral Results Across Test Sessions in Free-Recall, Probe and SRET Tasks*



*Note.* (a). Probe results across sessions for Go-cued and Go-uncued conditions. Distributions are shown in combinations of boxplots (left, annotated with medians and quartile ranges) and smoothed kernel density estimates (right, annotated with error bars indicating standard error of the mean)[35]. The dashed line reflects a chance level of 50%. Asterisks above each bar represent proportions higher than the chance level in a one-tailed generalized linear mixed-model analysis (log odds = 0; odds ratio = 1). $^{**}\ p < 0.01$, $^{*}\ p < 0.05$, $^{+}\ p < 0.1$.
(b). Interaction between initial Rating and TMR Condition on predicting reaction times (via a log transformation) during baseline SRET (left), post-sleep SRET (middle), and one-week delayed SRET task (right). Solid lines indicate linear regression fit to the trial-averaged data. The shaded area indicates two-sided parametric 95% CI.

## EEG results

### *Cue-Elicited Sigma Power During TMR Modulated by Pre-Memory Strength*

We analyzed cue-elicited electroencephalographic (EEG) responses during the TMR to investigate the neural mechanisms underlying reactivation of positive traits. Previous research showed that sleep spindles and spindle-related sigma power (12 to 16 Hz) are among the key factors driving TMR benefits [36–38]. For example, the cue-elicited spindle density and sigma power [37,39] positively correlated with post-sleep memory retention [40,41]. Moreover, research argues that following cue-elicited sigma power increase, there is a critical refractory period during which sigma power would decrease [42]. Importantly, theoretical accounts proposed that successful memory reactivation not only depends on the cue-elicited sigma power increase, but also on the following refractory periods wherein memory is re-processed with minimal interference [39,42,43].

Here, analyzing the TMR cue-elicited EEG power change at the central electrode (Cz [20,44]) revealed two significant positive clusters: the delta–theta–alpha band (0.75 to 11.75 Hz, 0.02 to 1.7 seconds), and the sigma–beta band (12 to 28 Hz, 0.3 to 1.66 seconds, Figure. 4a), and one negative cluster: sigma band (12 to 17 Hz, 1.9 to 2.44 seconds) during the 0 to 3000 milliseconds post-cue time window ($p < 0.001$, corrected). These clusters were consistent with previous findings on cue-elicited modulations during sleep [44]. Following previous research, this earlier positive cluster might represent sensory processing of the cue, whereas the following negative cluster corresponded to the refractory period, indicating a cue-specific memory reactivation process [45]. Table 2 presents the sleep staging results and Table 3 presents the mean number of trials for each -strength condition (i.e., weak, intermediate, strong) for all participants (N = 35).

**Table 2**

Sleep parameters (mean ± SEM, in minutes).

| Participants | Total time | Wake | N1 | N2 | N3 | REM |
|---|---|---|---|---|---|---|
| All participants (N = 35) | 90.10 ± 1.19 | 10.85 ± 1.14 | 2.88 ± 0.51 | 37.12 ± 1.50 | 28.68 ± 1.84 | 10.58 ± 1.28 |
| Participants contain all three cue conditions (N = 18) | 89.52 ± 1.33 | 11.48 ± 1.63 | 1.96 ± 0.49 | 37.21 ± 2.18 | 30.26 ± 2.64 | 8.62 ± 1.72 |

**Table 3**

Number of trials for each memory strength condition during TMR (mean ±SEM).

| Participants | Weak | Intermediate | Strong | Control |
|---|---|---|---|---|
| All participants (N = 35) | 149.37 ± 9.16 | 43.55 ± 4.36 | 25.33 ± 2.25 | 13.86 ± 0.81 |
| Participants contained in the EEG power analyses (N = 18) | 150.22 ± 12.21 | 43.94±5.27 | 25.39±2.88 | 14.61±1.04 |

To directly link TMR recall benefits and TMR-elicited EEG activity, we extracted the averaged power within the identified significant positive and negative clusters of spindle-related sigma band (12 to 16 Hz) for each participant and compared between pre-TMR memory-strength conditions (i.e., weak, intermediate, strong, and control traits) (Figure. 4b–e). Given that both TMR-elicited sigma power and the refractory period are important for memory reactivation, and these two sigma activity are often inversely correlated, we quantified the cueing-induced sigma activity as the difference between the earlier positive (0.58 to 1.64 seconds) and later negative sigma cluster (1.9 to 2.42 seconds). Specifically, we subtracted the negative cluster from the positive cluster, with higher values indicating stronger sigma- power changes and therefore memory reactivation.

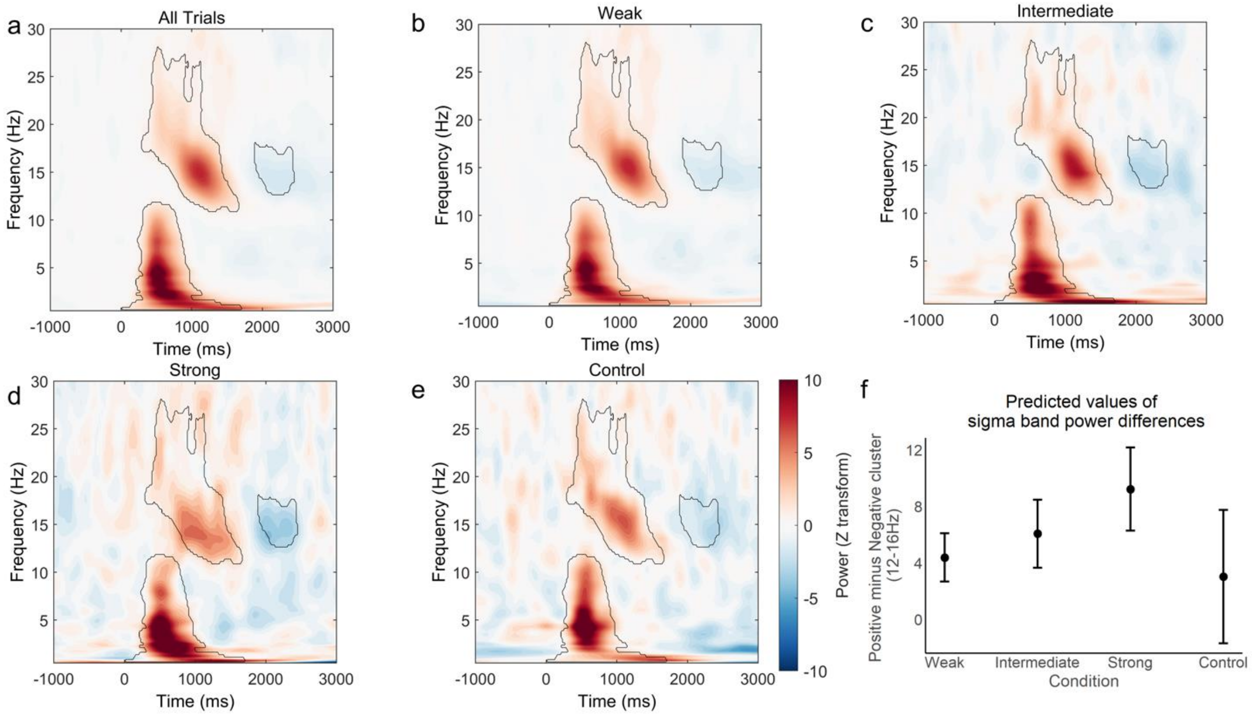
We ran a LMM using pre-TMR memory strength (i.e., weak, intermediate, strong, control) as fixed effects, and by participant and trait as random effects, to predict post-cue sigma power changes (12 to 16 Hz) for each trait word. The included participants were the same as in post-TMR recall analysis ($N = 18$). We found a significant main effect of pre-memory-strength condition, $F$ (3, 21.85) = 4.84, $p = 0.01$. Pairwise comparisons showed that strong traits elicited significantly larger sigma power changes than the other three conditions ($p < 0.001$, $p = 0.067$, $p = 0.057$, compared with weak, intermediate, control, respectively, see Figure. 4f), while no differences were found among other conditions, all $p$s > 0.08. This result might indicate preferential reactivation and reprocessing for strongly encoded traits during sleep.

Next, when separately testing sigma power in the early positive and late negative cluster, we only found a main effect of pre-memory strength on the negative cluster ($F$ (3, 9.7) = 4.96, $p = 0.024$) but not the positive cluster ($F$ (3, 6.1) = 1.82, $p = 0.243$). Post-hoc comparison showed that strong traits had been associated with the lowest sigma power as compared to weak and intermediate traits (all $p$s <0.05), but no difference from the control trait word ($p = 0.1$). These results indicate that the difference of post-cue sigma power changes among weak, intermediate, strong, and control traits were most likely driven by the differences of the later negative sigma cluster.

Finally, no significant main effect of pre-memory strength was found when predicting post-cue delta (0.5 to 4 Hz), theta (4 to 8 Hz), alpha (8 to 12 Hz), or beta (16 to 25 Hz) band power increases (all *p*s > 0.068).

**Figure 4**

*Time-frequency Representations Following Cue Presentation*



*Note.* (a) Time-frequency power spectrum averaging all traits played during sleep at channel Cz (Z transformed). Time zero indicates the onset of spoken traits. (b–e) Time-frequency power spectrum for specific trait-word types. (f) Predicted values of sigma power differences (the positive minus the negative cluster at sigma band, 12 to 16 Hz). Black outlines indicate the same cluster for all pre-TMR memory-strength type traits.

### *Relationship Between Sigma Power Change and Self-Evaluation*

To explore the relationship between TMR-elicited sigma activity and post-TMR self-evaluative memories and self-evaluations, we ran a series of LMM analyses using averaged sigma power changes for each trait as the fixed effect, and by participant and/or trait as random effects to predict self-evaluation memories (i.e., recall task), self-evaluation preferences (i.e., probe task), and self-evaluation endorsement (i.e., SRET task).

When predicting post-TMR recall order, we found a negative association $F$ (1, 244.5) = 5.4, $p = 0.021$, such that greater sigma power changes correlated with earlier recall (degree of association = -0.02). Although there was a trend toward greater post-cue sigma power change predicted higher post-TMR recall percentage, the association did not reach statistical significance ($\chi 2(1) = 3.05$, $p = 0.081$). Next, regarding post-TMR self-evaluation preferences (i.e., binary choice from Go versus NoGo pairs in the probe task), we found that higher sigma power changes positively predicted Go choice, $F$ (1, 422.03) = 4.84, $p = 0.028$. Finally, when predicting self-referential endorsement, we did not find any associations between sigma power changes and endorsement proportion nor endorsement speed after TMR. No predictions were found for the above models in the one-week delay session (all *p*s > 0.2). Therefore, while sigma-band power changes might index self-evaluative memory reactivation as evidenced by recall, it remains unclear whether they bear relationships with self-evaluation preference changes (Figure. 3a) and self-referential endorsement speed (Figure. 3b).

## Discussion

We asked whether we could enhance ones' positive self-evaluative memories via wakeful cue-approach training (CAT) and sleep-based targeted memory reactivation (TMR). Extending prior CAT research, CAT reliably induced preferences for trained self-evaluative positive-traits (behavioral replications are reported in SOM). Following CAT, we reactivated a subset of previously trained positive-traits during participants' NREM sleep to enhance their memory accessibility. Results showed that TMR prioritized the retention of strong self-evaluative trait-word memories while weakening the retention of intermediate trait-word memories in the short term. In terms of its possible long-term benefits, participants preferred Go-cued traits as self-descriptive over NoGo traits, and were faster to endorse Go-cued traits when they were highly endorsed at the baseline assessment. We also found that re-playing spoken traits to participants during their sleep enhanced spindle-related sigma power change difference for strongly encoded traits than for other cue or control traits, which might indicate memory reactivation for highly self-referential traits.

Behaviorally, when assessing CAT and TMR benefits on self-evaluative memories, participants who were presented with all three types of personality traits (i.e., strong, intermediate, weak) during sleep consistently showed earlier recall order for strong traits during the post-TMR free-recall task. This finding indicates that when cueing is performed during sleep, it enhances the accessibility of traits that were strongly memorized before sleep. Contrary to our findings in SOM where we found a CAT effect on enhancing recall order for intermediate-type traits after a one-week delay, TMR weakened intermediate-type traits immediately after cueing, indicating that the memory benefits for strong traits came at the expense of intermediate traits. Surprisingly, we did not find any behavioral evidence to support TMR benefits for weak traits. Note that previous TMR and sleep memory research often reports that weak memories are preferentially consolidated [16,18,24,46]; such a discrepancy could be due to the materials used. Whereas prior research has mostly focused on self-irrelevant materials, such as images and words, we examined self-evaluative memories. Particularly, participants were less likely to endorse weak traits as self-descriptive and were unable to recall them in either of the pre-sleep tests. The significantly low level of endorsement might make weak traits less salient and further prevent them from being consolidated. Note, the definitions of, and testing protocols for, memory strengths also differ across studies, making direct between-study comparisons difficult, if not impossible [47]. Hence, previous findings that showed weak memories were preferentially consolidated during sleep may not generalize to self-evaluative memories. By contrast, we postulate that for reactivating traits during sleep, strongly encoded, highly self-relevant memories might be better preserved given their higher salience.

In examining how TMR-related neural activity influenced subsequent self-evaluative memory recall, we found that spoken traits induced a transient spindle-related sigma power increase shortly after the cue onset, as well as a late power decrease around 2–3 seconds after cue presentation. Although previous TMR studies have reported a positive link between the transient post-cue sigma power increase and promoted memory [48,49], our findings did not reveal such an association. Instead, when we closely examined post-cue sigma power, the difference between the early positive and late negative sigma cluster (which we defined as sigma power change), as well as the negative cluster, proved statistically significant among different cues. Importantly, trait adjectives that were strongly memorized prior to sleep elicited greater spindle-related sigma power change than other traits (albeit marginally greater than control and intermediate traits), indicating a preferential memory reactivation for strong traits [50]. This result also suggests that early sensory processing and a late refractory period following a spoken trait word are necessary for its successful memory reactivation [45]. Additionally, in this study we had unbalanced trial numbers for the cueing traits that were assigned to the three memory conditions; future studies could include more traits and trials to examine the replicability of cueing-induced spindle activity for weak, intermediate, or strong self-evaluative memories. Nevertheless, the present findings extend prior research on vocabulary reactivation, indicating that the sleeping brain not only processes newly acquired vocabularies or semantic words [17,51], but also self-evaluative trait memories.

In addition to self-evaluative memories, we also assessed participants' self-evaluation preferences in the binary-probe task. We found that after CAT, participants were more likely to choose Go versus NoGo traits as self-descriptive. This result extends prior CAT research [10], demonstrating that CAT is effective for influencing change in not only food choices but also self-evaluations. We subsequently examined whether reactivating positive traits during sleep could further influence participants' self-evaluations. Immediately after sleep, TMR cueing did not shift participants' preference choices toward Go-cued than Go-uncued traits, although participants preferred both Go-cued and Go-uncued traits over NoGo traits (i.e., the CAT effect). In other words, the TMR did not further improve CAT effects in self-evaluation preferences. Notably, TMR benefits may emerge over a longer term one week later: participants were more likely to choose Go-cued than NoGo traits as self-descriptive, although the differences between CAT-cued and CAT-uncued traits did not reach significance. These results suggest TMR may preserve the CAT effect (Go-cued > NoGo) in the binary-preference choice task after a long-term delay. However, no clear associations were identified between TMR induced EEG power changes and choice preference in the delay session. This might indicate that long-term preference change was not driven by sigma power change difference.

When assessing participants' binary endorsement change using SRET, we did not find significant main effects of CAT or TMR. We postulated that, because the endorsement proportion of positive traits was already high (i.e., > 80%), such a ceiling would leave little space for CAT or TMR to increase participants' positive trait endorsement. In addition to endorsement judgement, reaction times (RT) of a certain choice could also infer preferences [34]. By examining the association between endorsement ratings and RT across different memory-strength conditions, we found a stronger negative association for Go-cued rather than Go-uncued traits after a one-week delay, which suggests cueing strongly-endorsed traits during sleep leads to the faster endorsement in the long term. Thus, TMR selectively promoted endorsement speed for initially high-endorsed traits. Note, we also failed to observe any associations between cue-elicited EEG power changes and endorsement proportion or speed. This could be due to the ceiling effect reported in SRET (> 80% endorsement rates for positive traits), which may prevent us from detecting the possible TMR benefits on endorsements. To better understand memory reactivation and positive self-endorsement changes, future studies could carefully match the number of endorsed traits and not-endorsed traits prior to TMR.

Limitations shall be discussed. First, because our study only included self-relevant traits, to which extent the observed effects are self-specific remain an open question. For example, can CAT and TMR be used to increase people’s endorsements for traits that are considered as relevant to other people? Can the impact of pre-sleep memory strength on TMR effects be found for non-personality, neutral words? Second, due to the limited amount of time participants spent in REM sleep during nap, our current dataset did not allow for further testing of the role of REM sleep in consolidating self-evaluative memories. Given that REM sleep plays an important role in emotional memory and vocabulary learning [52,53], future research shall test the role of REM-related processing in self-evaluative memories.

To summarize, we show that reactivating positive personality traits during NREM sleep impacts the consolidation of self-evaluative memories as a function of pre-sleep memory strength. TMR prioritized recall of strong self-evaluative memories, which also elicited enhanced sigma power during sleep. In contrast, intermediate self-evaluative memories were weakened, and no change was identified for weak self-evaluative memories. After one week, TMR not only preserved preferences for cued traits, but also facilitated endorsement of traits when they were highly endorsed at baseline, prior to CAT and TMR manipulations. These findings highlight the importance of internal (i.e., baseline endorsement of traits) and external factors (e.g., training, sleep, self-evaluative tasks) in generating the joint effect of CAT and TMR on self-evaluation. The present study contributes to the understanding of self-evaluative memories via cognitive training and sleep-based memory-reactivation paradigms. If positive self-evaluations and memories can be enhanced, future research can test whether this approach can help people with low self-evaluations and self-doubts.

## Methods

### Participants

Our final sample included 35 participants with valid behavioral and EEG data (8 males, *M*age ± *SD* = 20.83 ± 2.20 years), which is comparable to recent TMR studies (e.g., Schechtman et al. 2021). 10 participants were excluded prior to this due to insufficient slow-wave sleep (SWS), and thus very few numbers of cues were played (< 2 rounds). To facilitate sleep in the lab, we asked participants to wake up one hour earlier than their usual waking time and to avoid consuming caffeinated drinks on the day prior to – and of – the experiment. Participants were pre-screened regarding any current or history of sleep, psychiatric, or neurological disorders and had normal or corrected-to-normal vision. Participants received monetary compensation for their participation (250 RMB, or roughly 38 USD) and gave written consent prior to the experiment.

The study was approved by the Human Research Ethics Committee of the University of Hong Kong.

## Materials

All experimental procedures were implemented in E-Prime® 3.0 (Psychology Software Tools, Inc., Sharpsburg, Pennsylvania, USA). We selected 60 positive personality trait adjectives (e.g., ‘clever’) and 60 negative personality trait adjectives (e.g., ‘lazy’; see SOM for the complete list of personality traits). Each trait word was presented verbally for a duration of 400 to 600 milliseconds. In addition, we intermixed a novel adjective that was only played during TMR with the other cue words as a control.

## Experimental Tasks

### *Task Overview*

Participants attended two lab sessions, scheduled approximately one week apart (Figure. 1a). In the first session, participants arrived to the lab at approximately 12:00 pm (exact arrival times ranged between 11:30 am to12:30 pm), where they read and signed consent forms and were set up with EEGs. Subsequently, a series of four task phases began in which participants completed a number of tests, beginning with baseline tests in the first phase, followed by CAT and post-CAT/pre-TMR tests in the second phase, sleep-based TMR in the third phase, and post-TMR tests in the fourth phase. In the preliminary baseline phase, participants completed computer-based personality questionnaires, serving as a cover story for the personality trait words (hereafter, traits) presented to them in the following SRET. During the SRET, participants rated the extent to which specific traits described themselves. Participants then completed a self-evaluative memory free-recall test. In the second phase, participants manually responded to positive traits (i.e., Go traits), prompted by visual and aural cues presented on screen and from a nearby loudspeaker (CAT). Participants then completed a free-recall test and a probe test, in which they were presented with Go and NoGo trait word pairs and asked to select the trait word that was more self-descriptive. In the third phase, half of the positive traits were aurally re-played to sleeping participants during slow-wave sleep (SWS). Then, in the fourth phase, participants completed the same free-recall test, probe test, and SRET. In the second lab visit (~ 7 days later), participants completed the same free-recall test, probe test and SRET as previously completed in the final phase of the first visit to examine the possible long-term TMR effects. Thus, they completed four self-evaluative memory free-recall tests (baseline, post-CAT, post-TMR, delay), three SRETs (baseline, post-TMR, delay), and three probe tasks (post-CAT, post-TMR, delay).

### *Baseline*

Participants completed preliminary computer-based personality questionnaires, including the Rosenberg Self-Esteem Scale (RSES[54]), Narcissistic Personality Inventory (NPI[55]) , Big Five Inventory (BFI[56]), Beck Depression Inventory-II (BDI-II[57]), State-Trait Anxiety Inventory (STAI state and STAI trait [58]), and Barratt Impulsiveness Scale (BIS-11[59]). This served as a cover story to promote the reliability of the personality trait adjectives that would be presented to them in the following self-referential encoding task (SRET). Descriptions of questionnaire scores are presented in Table S2.

In SRET (see Figure. 1b), a cross symbol was presented on a computer screen at the beginning of each trial for 0.5 seconds, followed by the presentation of the sentence ‘I think this word is applicable to me’ in the center of the screen for another 0.5 seconds. After 1.2 to 1.4 seconds, participants were presented with a random word, given visually in written form and aurally from a speaker, from a selection of120 adjectives for 0.8 seconds. After, participants were shown a blank screen for another 0.8 seconds and then were prompted to select if a trait word applied to them within 2.5 seconds by moving the mouse cursor continuously. The spatial location of 'Yes' and 'No' responses were counterbalanced (upper left/upper right or upper right/upper left). Following a 'Yes' response, participants were asked to rate the extent to which a trait word applied to them on a scale ranging from "slightly accurate" to "extremely accurate"; following a 'No' response, participants were asked to rate the extent to which a trait word did not apply to them on a scale ranging from "slightly inaccurate" to "extremely inaccurate".

Within the three minutes that followed, participants were asked to complete a self-evaluative free-recall task. Unlike previous recall tasks where participants must write down as many traits as possible, here, participants were asked to recall only the traits they had been presented with and endorsed during the previous SRET, typing each recalled word on a computer one at a time. Therefore, performance during this version of the recall task reflected self-evaluative memories.

### *CAT and Post-CAT Tests*

Following baseline assessments, participants completed a CAT task, followed by a 5-minute working memory task for distraction purposes. In the CAT (see Figure. 1c), we ranked all 60 positive traits from 1 (least accurate) to 60 (most accurate) based on the methods of Schonberg et al. (2014) and the SRET endorsement ratings. During CAT, 30 traits (across the entire list of trait adjectives) were paired with a visual Go cue that required participants to press a button as quickly as possible before the current trial's offset. The Go traits were presented aurally (< 0.8 seconds) and visually (1.2 seconds), followed by Go cues. We used an adaptive response window to keep participants attentive to Go cues. Specifically, initial Go cues were presented with a GSD (the delay between trait-word onset and Go-cue onset) of approximately 0.9 seconds.

If the participants successfully pressed a button before the offset of the trial, the GSD was increased by 17 milliseconds to increase task difficulty. If participants failed to make a button press before the offset of the trial, the GSD was reduced by 50 milliseconds to reduce task difficulty. Participants then completed a 3-minute post-CAT self-evaluative memory-recall task, as used in the baseline assessment.

Next, participants completed a post-CAT probe task so we could assess CAT effects. In the probe task (see Figure. 1d), a fixation cross appeared in the center of the screen for 1 second at the beginning of each trial. Following this, participants viewed two traits that were presented side-by-side and were instructed to select the trait word that best described them within 4 seconds. We excluded trials where response times exceeded 5 seconds, allowing a short window of time to account for potential delays caused by using the mouse. Participants confirmed their selection by clicking a push button, lasting for 0.5 seconds, which appeared under their chosen trait word. If participants took longer than 4 seconds to respond, they received a prompt in the confirmation phase to remind them to respond as quickly as possible. To balance the positions of Go and NoGo traits, we carried out a total of two blocks. The positions of the Go/NoGo traits per pair were randomly assigned to the upper-left/right or upper-right/left sides of the monitor in the first block. In the second block, each pair's left/right positions were swapped accordingly.

### *Nap Targeted Memory Reactivation (TMR)*

Participants took a 90-minute nap in a quiet, darkened sleep chamber. Background white noise (at ~ 38 dB) was played to participants throughout the duration of the nap via a loudspeaker placed near the bed. We used EEG recording devices to monitor participants' brain and physiological activities continuously while they napped. Once they entered SWS, we presented spoken, positive traits (the same spoken traits presented during the SRET and CAT tasks) at approximately 40 dB. Note, we set the volume of white noise to approximately 38 dB to ensure that the spoken traits (played at ~ 40 dB) could be heard above it, yet on a subtle level to avoid arousing participants from their nap.

Each spoken trait lasted no longer than 1 second, with an interstimulus interval of 5 seconds and an additional randomized interval of 0 to 1 second. Half of the positive Go traits were played during the TMR (i.e., 15 traits in total; see the previous section on CAT/TMR trait word-groupings for more details about trait selection). Additionally, we included a neutral trait word to the playlist, lasting a duration of 600 milliseconds, intermixed into each play cycle. Presenting the neutral trait word enabled us to monitor participants' brain activity specific to cued-memory reactivation versus nonspecific auditory processing of semantic stimulation. The minimal number of repetitions of a sequence (i.e., 16 traits) was 3.

Specifically, spoken traits were played to participants until they showed sustained SWS (i.e., SWS lasting at least 2 minutes). We then played the novel control spoken traits three times at the beginning of the TMR to ensure that the auditory stimulation did not wake participants. We started presenting the spoken traits to participants when they exhibited no more signs of arousal or changes in NREM sleep stage. If participants awoke, we stopped playing the cues and waited for them to re-enter SWS before repeating the above procedures. TMR was terminated immediately after 30 minutes except when SWS was not observed, or EEG recordings indicated micro-arousal or full awakening. If we observed no sign of SWS after 40 minutes, we started presenting the spoken traits to participants when they entered the N2 stage of sleep. We woke participants when their EEG recordings indicated that they had entered the N1 or N2 sleep stage (after around 90 minutes). After waking, participants were given a break of approximately 5 minutes to reduce the effects of sleep inertia.

### *Post-TMR Tests*

Participants completed the free-recall task, probe task, and SRET task. Here, the SRET was similar to the baseline SRET except participants only made a 'Yes'/'No' binary response.

### *One-Week Delayed Tests*

Participants returned to the lab around one week later to complete the delayed tests. They completed the following tasks in order: (1) a 3-minute self-evaluative memory-recall task; (2) a probe task; (3) a SRET 'Yes'/'No' self-evaluative task; and (4) post-experiment questionnaires. Unless otherwise stated, the tasks were identical to the ones used in the behavioral study in SOM. Participants were not informed of the tasks involved during the entire experiment. After the visit, participants were debriefed and paid.

**Details of the Stimuli used in the CAT, Probe, and TMR Tasks.** For the selection of Go and NoGo traits with matched ratings during the probe task, we sorted all ratings from lowest (1) to highest (60), then grouped Go and NoGo traits as follows: one Go and one NoGo item from the same rank for matched Go and NoGo trait word pairings, generating 30 pairs in total. Using the same sorting strategy as before, we grouped each pair into low- and high-rating pairs based on the order (see Figure. S12a for details of how Go/NoGo traits were assigned and counterbalanced across participants). We used this pairing strategy to obtain similar initial endorsement ratings between Go and NoGo traits. We also paired one low-value Go item from ranks 5 to 24 with one high-value NoGo item from ranks 37 to 56 (see Figure. S12b) and one low-value NoGo item from ranks 5 to 24 with one high-value NoGo item from ranks 37 to 56 (see Figure. S12c); additionally, we paired one low-value Go item with one high-value Go item (see Figure. S12d) and one low-value NoGo item with one high-value Go item (see Figure.

S12e), but these pairings served as controls to rule out mere exposure effects and were therefore not analyzed.

Finally, we assigned Go traits to Go-cued and Go-uncued categories, following the same order per participant (see Figure. S12 for the selection and grouping of positive traits during the CAT, probe, and TMR).

## EEG Data Acquisition

Continuous EEGs were recorded using a 63-channel customized cap with passive Ag/AgCl electrodes via a BrainAmp amplifier with 1000 Hz sampling rate (Brain Products, Gilching, Germany). Electrodes were positioned according to the International 10–10 system. The ground electrode was located at AFz, with FCz as the on-line reference electrode. The impedances were kept below 20 kΩ. We placed one electro-oculography (EOG) electrode under participants' left eyes and bipolar electromyography (EMG) electrodes on their chins to monitor eye movements and muscle activity during sleep.

## Behavioral Data Analysis

Statistical analyses were carried out using R (Version 4.2.1.[60]). Behavioral analysis methods were like those used in the behavioral experiments in SOM, except the TMR factor was added where applicable. We performed (G)LMMs to analyze the CAT- and TMR-induced behavioral changes.

For statistical significance testing, we used Type III Analysis of Variance with the Satterthwaite approximation method for the LMM, and Type III Wald Chi-Square tests for the GLMM. We followed up significant effects with post-hoc comparisons in *emmeans* [61] to derive the estimated marginal means from each model. Unless otherwise stated, we used the FDR method to adjust for multiple comparisons to control for false-positive results.

### *Behavioral Replication Analysis in the Baseline Session*

To test whether we replicated the classic pattern of valence-dependent memory bias, we ran two paired t tests with valence as the independent variable, endorsement proportion (via SRET), and processing bias score (via free-recall task and the number of positive/negative traits words that were recalled and endorsed divided by number of traits endorsed) as the dependent variable, respectively. Next, we ran a Pearson correlation test to examine the relationship between BDI scores (depressive symptoms) and the self-evaluative memory bias from the recall task at baseline session. The self-evaluative memory bias was defined as the difference between the positive and negative processing score, with a higher score indicating a larger self-positive processing bias.

### *Self-Evaluative Memories in the Free-Recall Task*

To better depict memory changes across multiple times, we took preceding recall performance into account in the (G)LMMs. Specifically, when analyzing post-CAT recall performance, we added the baseline recall (‘Yes’/‘No’) as a fixed effect. When analyzing post-TMR recall performance, we grouped traits into weak, intermediate, and strong categories depending on baseline and post-CAT recall tasks. We defined weak-condition traits as traits that were not recalled at all in both baseline and post-CAT tests, intermediate-condition traits as traits that were only recalled once in either baseline or post-CAT tests, and strong-condition traits as traits that were recalled in both baseline and post-CAT tests. Only participants with all three types of traits for cued and uncued conditions were included in our analysis.

Next, we performed the following four (G)LMMs across post-CAT and post-TMR for three outcome variables: (1) recall or not recall; and (2) recall order. These were as follows:

(1) Recall (binary, ‘Yes’ or ‘No’):

The GLMM for the immediate post-CAT recall task was defined as:

Post-CAT_Yes/No ~ 1 + CAT Condition × Baseline recall + (1+CAT|Subject ID) + (1|Trait ID).

The GLMM for the post-TMR recall task was defined as:

Post-TMR_Yes/No ~ 1 + TMR Condition × Memory Strength + (1+TMR|Subject ID) + (1|Trait ID).

(2) Recall order of recalled traits:

Note, we removed the random factor for participant because the model encountered a singular fitting issue due to the limited number of traits per CAT/TMR condition. The LMM for the for post-CAT recall task was defined as:

Post-CAT_Recall Order ~ 1 + CAT Condition × Baseline Recall + (1|Trait ID).

The LMM for the post-TMR recall task was defined as:

Post-TMR_Recall Order ~ 1 + TMR Condition × Memory Strength + (1|Trait ID).

### ***Evaluation Preference in the Probe Task***

Following previous CAT research, we ran GLMMs separately for low- and high-rating subsets across three test sessions (i.e., post-CAT, post-TMR, and delay) on binary choice outcome in the probe task. To investigate the effect of TMR on binary choices during post-TMR and delay, we separated trait-word pairs into Go-cued and Go-uncued categories after the nap TMR and tested them by adding a TMR condition (i.e., Go-cued versus Go-uncued) as the fixed effect.

Specifically:

(1) We ran a GLMM with the choice (‘Yes’ or ‘No’) as the dependent variable and participants as the random effect to evaluate the CAT effect on self-evaluative judgment change. The simplest GLMM was defined as:

Choice ~ 1 + (1|Subject ID).

(2) Next, we added high- versus low-rating Go-NoGo (GNG) pairs as a fixed effect and participants as the random effects to explore the difference in probe across low- versus high-rating traits. This line of GLMMs were defined as:

Choice ~ 1 + High Low + (1|SubjectID).

(3) We ran the GLMM with the TMR as a fixed effect to test the TMR benefits. This GLMM was defined as:

Choice ~ 1 + TMR + (1|SubjectID).

Note, we applied a one-sided test for the simplest GLMMs but a two-sided test for the other comparisons to be consistent with CAT literature.

### ***Self-Evaluative Judgments in the SRET***

To analyze the changes in positive self-evaluative judgments, we ran a GLMM for the positive traits only. The model is specified as:

Choice ~ 1 + Session × TMR + (1+TMR |Subject ID) + (1|Trait ID).

Finally, to examine potential associations between initial rating strength, TMR, and RTs of ‘Yes’ answers, we analyzed trial-level RTs using a linear mixed model with initial rating, TMR condition as the fixed effect, and participant as the random effect, separately across post-CAT, post-TMR, and delay. Since participants’ responses could change between test sessions, the relationship between initial rating strength and RT was examined separately for each test session. Note, due to the limited trait numbers assigned to each memory-strength condition (i.e., weak, intermediate, strong), we removed the random slope for participant and random effect for trait IDs in the (G)LMMs to avoid a singular fitting problem.

The LMM was defined as:

Log RT ~ 1 + TMR × Initial rating + (1 |Subject ID).

Unless otherwise stated, all tests were two-sided, with alpha level set at .05.

## EEG Data Analysis

### *EEG Data Pre-Processing*

EEG data were pre-processed using custom-written scripts and the MATLAB Toolbox EEGLAB [62]. First, nap EEG data were down-sampled to 250 Hz, notch-filtered at 50 Hz, and then re-referenced to the averaged mastoids. Second, EEG data were band-pass filtered at 0.1 to 40 Hz. While EOG and EMG data were used for sleep staging, these data were not used in the time-frequency analysis and phase analysis.

### *Offline Sleep Stage Scoring*

In sleep offline analyses, we scored sleep stages (N1, N2, SWS, and rapid-eye-movement (REM)) based on EEG (Channel Cz), EOG, and EMG patterns, using algorithms implemented in the YASA open-source Python Toolbox [63].

### *Time-Frequency Power Analysis*

Before analyzing cue-elicited time-frequency power change, the cue-elicited EEG data were epoched into -4 to 6 second segments, relative to the onset of each cued trait word. This long epoch ensured that we had enough edge artifact-free segments for each clean epoch to assess TMR benefits ( -1 to 3 seconds). Epochs with artefact were visually inspected and removed. Time-frequency decomposition was performed in the Fieldtrip open-source MATLAB toolbox [64]. We used 3 to 15 cycles in a step of 0.25 Hz Morlet wavelet and baseline corrected using z-transformation of all trials from -1000 to -100 milliseconds relative to the cue onset. Following previous TMR studies [20,65], we selected data from channel Cz for further analyses. The calculated time-frequency decompositions were then down-sampled to 50 Hz.

To identify significant cue-elicited EEG activity, we conducted non-parametric permutation tests to test the significance of cue-elicited power changes. We used a cluster-based permutation test to find the common region across all participants in the time-frequency domain [66]. Specifically, time frequency representations were shuffled within each participant 1000 times. The shuffled data were used to identify null distribution effects. Then we established temporally- and spectrally adjacent significant clusters (threshold *p* at 0.001). Next, we calculated the sum of t-values for each identified cluster in the original and permutated data. If no t-values reached significance in the permutations, the cluster was classified as 0. We then calculated the rank of cluster t-values in the distribution of random data. The cluster was marked as significant if an absolutely higher t-value was found in less than 0.1% of the random permutations.

Note, traits were grouped into three conditions based on baseline recall and post-CAT recall (i.e., weak, intermediate, strong). We then performed a linear mixed model with the four conditions as the fixed effect and by participant and trait word as random effects to predict the identified post-cue power change on single trait word level, including (1) the post-cue sigma power change differences between the positive and negative cluster (12 to 16 Hz), (2) delta-theta-alpha (0.5 to 4 Hz, 4 to 8 Hz, and 8 to 12 Hz) and beta (16 to 25 Hz) power changes. We used a boxplot method to detect potential outliers on the item level prior to performing statistical analysis for each power band, respectively. No multiple comparison corrections were applied during post-hoc tests due to a prior hypothesis that Go-cued traits would elicit greater power change than control traits. To match the findings observed in the post-TMR recall task, we ran this analysis on the same 18 participants. The model was defined as:

Sigma power change difference ~ 1 + Condition + (1|Subject ID) + (1|Trait ID).

To explore the potential difference of positive and negative post-cue sigma power, we ran the LMM on these two clusters separately. The model was defined as:

Positive cluster/Negative cluster ~ 1 + Condition + (1|Subject ID) + (1|Trait ID).

### ***Brain-Behavior Association Analysis***

To explore participants' EEG activity and behavioral performance after TMR, we also ran several (G)LMMs using post-cue sigma power difference (i.e., difference between the positive and negative cluster in sigma band, 12 to 16 Hz) as the fixed effect and by participant and/or trait word as random effects. We ran this analysis on behavioral measurements obtained from the post-TMR recall task, probe task, and SRET during post-TMR and delay sessions, respectively.

We removed the by-participant or trait-word random effect if the model encountered a singular fitting issue.

In the free-recall task, the models were defined as:

(1) Recall outcome (Yes/No) ~ 1 + Sigma power change difference + (1|Subject ID) + (1|Trait ID).

(2) Recall Order (Logged) ~ 1 + Sigma power change difference + (1|Trait ID).

In the probe task, the model was defined as:

Choice (Yes/No) ~ 1 + Sigma power change difference + (1|Subject ID) + (1|Trait ID).

In the SRET task, the models were defined as：

(1) Choice (Yes/No) ~ 1 + Sigma power change difference + (1|Subject ID) + (1|Trait ID).

(2) Reaction times (Logged) ~ 1 + Sigma power change difference + (1|Subject ID) + (1|Trait ID).

## Data and code availability

The data and analytical code supporting the study’s findings are available at the Open Science Framework repository: https://osf.io/h2967/?view_only=1b259a176db245b4b9ef0e1b64ca74f9.

## References

1. Fields, E. C. & Kuperberg, G. R. Loving yourself more than your neighbor: ERPs reveal online effects of a self-positivity bias. *Soc. Cogn. Affect. Neurosci.* **10**, 1202–1209 (2015).

2. Guenther, C. L. & Alicke, M. D. Deconstructing the better-than-average effect. *J. Pers. Soc. Psychol.* **99**, 755–770 (2010).

3. Taylor, S. E. & Brown, J. D. Illusion and well-being: A social psychological perspective on mental health. *Psychol. Bull.* **103**, 193–210 (1988).

4. Watson, L. A., Dritschel, B., Obonsawin, M. C. & Jentzsch, I. Seeing yourself in a positive light: Brain correlates of the self-positivity bias. *Brain Res.* **1152**, 106–110 (2007).

5. Dainer-Best, J., Trujillo, L. T., Schnyer, D. M. & Beevers, C. G. Sustained engagement of attention is associated with increased negative self-referent processing in major depressive disorder. *Biol. Psychol.* **129**, 231–241 (2017).

6. Sowislo, J. F., Link to external site, this link will open in a new window, Orth, U. & Link to external site, this link will open in a new window. Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychol. Bull.* **139**, 213–240 (2013).

7. Orth, U., Link to external site, this link will open in a new window, Robins, R. W. & Link to external site, this link will open in a new window. Is high self-esteem beneficial? Revisiting a classic question. *Am. Psychol.* **77**, 5–17 (2022).

8. Schonberg, T. *et al.* Changing value through cued approach: an automatic mechanism of behavior change. *Nat. Neurosci.* **17**, 625–630 (2014).

9. Oudiette, D. & Paller, K. A. Upgrading the sleeping brain with targeted memory reactivation. *Trends Cogn. Sci.* **17**, 142–149 (2013).

10. Salomon, T. *et al.* The Cue-Approach Task as a General Mechanism for Long-Term Non-Reinforced Behavioral Change. *Sci. Rep.* **8**, 3614 (2018).

11. Itzkovitch, A., Bar Or, M. & Schonberg, T. Cue-approach training for food behavior. *Curr. Opin. Behav. Sci.* **47**, 101202 (2022).

12. Diekelmann, S. & Born, J. The memory function of sleep. *Nat. Rev. Neurosci.* **11**, 114–126 (2010).

13. Rasch, B. & Born, J. About Sleep's Role in Memory. *Physiol. Rev.* **93**, 681–766 (2013).

14. MacDonald, K. J. & Cote, K. A. Contributions of post-learning REM and NREM sleep to memory retrieval. *Sleep Med. Rev.* **59**, 101453 (2021).

15. Petzka, M., Chatburn, A., Charest, I., Balanos, G. M. & Staresina, B. P. Sleep spindles track cortical learning patterns for memory consolidation. *Curr. Biol.* **32**, 2349-2356.e4 (2022).

16. Creery, J. D., Oudiette, D., Antony, J. W. & Paller, K. A. Targeted Memory Reactivation during Sleep Depends on Prior Learning. *Sleep* **38**, 755–763 (2015).

17. Schreiner, T. & Rasch, B. Boosting Vocabulary Learning by Verbal Cueing During Sleep. *Cereb. Cortex* **25**, 4169–4179 (2015).

18. Cairney, S. A., Lindsay, S., Sobczak, J. M., Paller, K. A. & Gaskell, M. G. The Benefits of Targeted Memory Reactivation for Consolidation in Sleep are Contingent on Memory Accuracy and Direct Cue-Memory Associations. *Sleep* **39**, 1139–1150 (2016).

19. Schapiro, A. C. *et al.* Sleep Benefits Memory for Semantic Category Structure While Preserving Exemplar-Specific Information. *Sci. Rep.* **7**, 14869 (2017).

20. Whitmore, N. W., Bassard, A. M. & Paller, K. A. Targeted memory reactivation of face-name learning depends on ample and undisturbed slow-wave sleep. *Npj Sci. Learn.* **7**, 1–6 (2022).

21. Abdellahi, M. E. A., Koopman, A. C. M., Treder, M. S. & Lewis, P. A. Targeting targeted memory reactivation: characteristics of cued reactivation in sleep. 2021.12.09.471945 Preprint at https://doi.org/10.1101/2021.12.09.471945 (2021).

22. Memory reactivation in slow wave sleep enhances relational learning | bioRxiv. https://www.biorxiv.org/content/10.1101/2022.03.29.486197v1.abstract.

23. Hu, X., Link to external site, this link will open in a new window, Cheng, L. Y., Chiu, M. H. & Paller, K. A. Promoting memory consolidation during sleep: A meta-analysis of targeted memory reactivation. *Psychol. Bull.* **146**, 218–244 (2020).

24. Denis, D. *et al.* Sleep spindles preferentially consolidate weakly encoded memories. *J. Neurosci.* (2021) doi:10.1523/JNEUROSCI.0818-20.2021.

25. Oyarzún, J. P., Morís, J., Luque, D., Diego-Balaguer, R. de & Fuentemilla, L. Targeted Memory Reactivation during Sleep Adaptively Promotes the Strengthening or Weakening of Overlapping Memories. *J. Neurosci.* **37**, 7748–7758 (2017).

26. Bennion, K. A., Payne, J. D., Kensinger, E. A. & Link to external site, this link will open in a new window. The impact of napping on memory for future-relevant stimuli: Prioritization among multiple salience cues. *Behav. Neurosci.* **130**, 281–289 (2016).

27. Alger, S. E., Kensinger, E. A. & Payne, J. D. Preferential consolidation of emotionally salient information during a nap is preserved in middle age. *Neurobiol. Aging* **68**, 34–47 (2018).

28. Tamaki, M. *et al.* Reward does not facilitate visual perceptual learning until sleep occurs. *Proc. Natl. Acad. Sci.* **117**, 959–968 (2020).

29. Sterpenich, V. *et al.* Reward biases spontaneous neural reactivation during sleep. *Nat. Commun.* **12**, 4162 (2021).

30. Kuiper, N. A. & Derry, P. A. Depressed and nondepressed content self-reference in mild depressives. *J. Pers.* **50**, 67–80 (1982).

31. Symons, C. S. & Johnson, B. T. The self-reference effect in memory: A meta-analysis. *Psychol. Bull.* **121**, 371–394 (1997).

32. Auerbach, R. P. *et al.* Self-referential processing in adolescents: Stability of behavioral and ERP markers. *Psychophysiology* **53**, 1398–1406 (2016).

33. Disner, S. G., Shumake, J. D. & Beevers, C. G. Self-referential schemas and attentional bias predict severity and naturalistic course of depression symptoms. *Cogn. Emot.* **31**, 632–644 (2017).

34. Konovalov, A. & Krajbich, I. Revealed strength of preference: Inference from response times. *Judgm. Decis. Mak.* **14**, 381–394 (2019).

35. Allen, M. *et al.* Raincloud plots: a multi-platform tool for robust data visualization. Preprint at https://doi.org/10.12688/wellcomeopenres.15191.2 (2021).

36. Carbone, J. *et al.* The effect of zolpidem on targeted memory reactivation during sleep. *Learn. Mem.* **28**, 307–318 (2021).

37. Rakowska, M., Abdellahi, M. E. A., Bagrowska, P., Navarrete, M. & Lewis, P. A. Long term effects of cueing procedural memory reactivation during NREM sleep. *NeuroImage* **244**, 118573 (2021).

38. Schreiner, T., Petzka, M., Staudigl, T. & Staresina, B. P. Endogenous memory reactivation during sleep in humans is clocked by slow oscillation-spindle complexes. *Nat. Commun.* **12**, 3112 (2021).

39. Cairney, S. A., Guttesen, A. á V., El Marj, N. & Staresina, B. P. Memory Consolidation Is Linked to Spindle-Mediated Information Processing during Sleep. *Curr. Biol.* **28**, 948-954.e4 (2018).

40. Fogel, S. M. & Smith, C. T. The function of the sleep spindle: A physiological index of intelligence and a mechanism for sleep-dependent memory consolidation. *Neurosci. Biobehav. Rev.* **35**, 1154–1165 (2011).

41. Kuula, L. *et al.* Higher sleep spindle activity is associated with fewer false memories in adolescent girls. *Neurobiol. Learn. Mem.* **157**, 96–105 (2019).

42. Antony, J. W. *et al.* Sleep Spindle Refractoriness Segregates Periods of Memory Reactivation. *Curr. Biol.* **28**, 1736-1743.e4 (2018).

43. Lewis, P. A. & Bendor, D. How Targeted Memory Reactivation Promotes the Selective Strengthening of Memories in Sleep. *Curr. Biol.* **29**, R906–R912 (2019).

44. Schechtman, E. *et al.* Multiple memories can be simultaneously reactivated during sleep as effectively as a single memory. *Commun. Biol.* **4**, 1–13 (2021).

45. Antony, J. W. *et al.* Sleep Spindle Refractoriness Segregates Periods of Memory Reactivation. *Curr. Biol.* **28**, 1736-1743.e4 (2018).

46. Drosopoulos, S., Schulze, C., Fischer, S. & Born, J. Sleep's function in the spontaneous recovery and consolidation of memories. *J. Exp. Psychol. Gen.* **136**, 169–183 (2007).

47. Petzka, M., Charest, I., Balanos, G. M. & Staresina, B. P. Does sleep-dependent consolidation favour weak memories? *Cortex* **134**, 65–75 (2021).

48. Antony, J. W., Cheng, L. Y., Brooks, P. P., Paller, K. A. & Norman, K. A. Competitive learning modulates memory consolidation during sleep. *Neurobiol. Learn. Mem.* **155**, 216–230 (2018).

49. Emotional arousal modulates oscillatory correlates of targeted memory reactivation during NREM, but not REM sleep | Scientific Reports. https://www.nature.com/articles/srep39229.

50. Fernandez, L. M. J. & Lüthi, A. Sleep Spindles: Mechanisms and Functions. *Physiol. Rev.* **100**, 805–868 (2020).

51. Beck, J., Loretz, E. & Rasch, B. Exposure to relaxing words during sleep promotes slow-wave sleep and subjective sleep quality. *Sleep* **44**, zsab148 (2021).

52. Hutchison, I. C. *et al.* Targeted memory reactivation in REM but not SWS selectively reduces arousal responses. *Commun. Biol.* **4**, 1–6 (2021).

53. Batterink, L. J., Westerberg, C. E. & Paller, K. A. Vocabulary learning benefits from REM after slow-wave sleep. *Neurobiol. Learn. Mem.* **144**, 102–113 (2017).

54. Rosenberg, M. Society and the adolescent self-image. in (Princeton, NJ: Princeton University Press, 1965).

55. Raskin, R. & Hall, C. S. The Narcissistic Personality Inventory: Alternative Form Reliability and Further Evidence of Construct Validity. *J. Pers. Assess.* **45**, 159 (1981).

56. John, O. P., Donahue, E. M. & Kentle, R. L. The Big Five Inventory—Versions 4a and 54. in vol. 10 (Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research, 1991).

57. Beck, A. T., Steer, R. A., Ball, R. & Ranieri, W. F. Comparison of Beck Depression Inventories-IA and-II in Psychiatric Outpatients. *J. Pers. Assess.* **67**, 588–597 (1996).

58. Spielberger, C. D. State-Trait Anxiety Inventory, Form Y (STAI). in (Palo Alto, CA: Consulting Psychologicals Press, 1983).

59. Patton, J. H., Stanford, M. S. & Barratt, E. S. Factor structure of the Barratt Impulsiveness Scale. *J. Clin. Psychol.* **51**, 768–774 (1995).

60. R Core Team (2020). R: a language and environment for statistical computing. R Foundation for Statistical Computing.

61. Lenth, R. V. *et al.* emmeans: Estimated Marginal Means, aka Least-Squares Means. (2022).

62. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004).

63. Vallat, R. & Walker, M. P. An open-source, high-performance tool for automated sleep staging. *eLife* **10**, e70092 (2021).

64. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput. Intell. Neurosci.* **2011**, 1–9 (2011).

65. Schechtman, E. *et al.* Sleep reactivation did not boost suppression-induced forgetting. *Sci. Rep.* **11**, 1383 (2021).

66. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).

## Acknowledgements

This research was supported by the National Natural Science Foundation of China (No. 31922089, No.32171056) and General Research Fund (No. 17601318) of Hong Kong Research Grants Council to X.H.

## Author contributions

Ziqing Yao: Conceptualization, Methodology, Data curation, Writing- Original draft preparation. Software, Project administration. Tao Xia: Writing - Review & Editing, Visualization. Zhiguo Zhang: Writing - Review & Editing. Xuanyi Lin: Writing - Review & Editing. Pengmin Qin: Writing - Review & Editing. Yina Ma: Writing - Review & Editing. Xiaoqing Hu: Conceptualization, Methodology, Writing- Original draft preparation. Writing- Reviewing and Editing, Funding acquisition, Resources, Supervision.

## Conflict of interest statement

None