Neural underpinnings of a two-phase memory suppression process in the neural response to self-related and observed perspective views

Xinwei Song [a,1], Qi Liu [a,1], Xiaodong Zhang [a], Can Liu [a], Chunmei Lan [b], Xiaolu Zhang [c], Ting Xu [a], Ran Zhang [d], Keith M. Kendrick [a], Benjamin Becker [e], Weihua Zhao [a,*]

[a] The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Laboratory for NeuroInformation, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, 611731, China
[b] School of Sport Training, Chengdu Sport University, Chengdu, 610041, China
[c] Anhui Children's Hospital, Pediatric Hospital Affiliated to Fudan University, Hefei, 230051, China
[d] Faculty of Psychology, Southwest University, Chongqing, 400715, China
[e] The State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong, 999077, China

ARTICLE INFO

ABSTRACT

Individuals often actively suppress intrusive memories to alleviate the distress they cause and maintain mental well-being. However, those with post-traumatic stress disorder (PTSD) often exhibit difficulties particularly in inhibiting or suppressing negative memories compared to individuals without PTSD. These memories can involve a physical threat either to the individual themselves or to others. Unfortunately, there is still limited understanding of the cognitive and neural mechanisms that underlie how suppression differs for self-related versus other-related memories. Here we capitalized on multivariate pattern analyses in combination with fMRI data acquired during a two-phase memory suppression paradigm where participants volitionally suppressed and subsequently recognized self-related and other-related stimuli. The results suggested that the recognition process following memory suppression demands more cognitive engagement for self-related stimuli than other-related stimuli, manifesting in increased activity in the dorsal anterior cingulate cortex (dACC). Furthermore, after memory suppression, we observed a stronger functional coupling between dACC identified during memory suppression, and both the middle frontal gyrus and the insula during self-related recognition compared to other-related recognition. An advanced multivariate pattern analysis substantiated that the limbic system and empathy network particularly contributed to accurately distinguishing between self-related and other-related recognition following memory suppression. Our findings demonstrated distinct neural representations of memory suppression related to self and others, suggesting that different strategies may be employed for suppressing intrusive memories originating from different sources.

Introduction

Emotion suppression in response to negative events or stressors plays a crucial role in flexible adaption and maintenance of our mental health. These stressors or events can involve a physical threat either directly to the individual themselves, for instance, through first-hand self-experience like enduring abuse, or indirectly to others, where the individual witnesses someone else's suffering (observed prospective), thereby triggering empathic distress (Baird et al., 2011; Decety et al., 2010; Pitman et al., 2012). However, the delicate balance of memory suppression is dysregulated in post-traumatic stress disorder (PTSD), a condition characterized by the involuntary re-experiencing of traumatic stressors or events, often manifesting as intrusive memories (Brewin et al., 2010). These intrusive memories are exceptionally vivid, filled with sensory and emotional details pertaining to specific traumatic occurrences. They are often triggered by a range of perceptually similar cues and tend to lack coherence, appearing disconnected from the rest of personal autobiographic memories (Ressler et al., 2022). Individuals with PTSD often exhibit deficits in emotion regulation, particularly in inhibiting or suppressing fear responses and overestimating threat

---

levels, compared to those without PTSD (Rabinak et al., 2017; Seligowski et al., 2019). However, studies investigating emotional suppression or memory suppression in PTSD have often failed to clearly distinguish between those exposed to self-experienced threats and those exposed to observed threats. This lack of distinction has hampered our understanding of the differences in suppression mechanisms between these two types of negative events or stressors. Therefore, the nuances of

suppressing responses to self-experienced and observed threats remain relatively underexplored.

Previous studies have suggested that the extreme stress caused by trauma is associated with alterations in memory suppression (Elzinga, 2002; Perl et al., 2023). Memory suppression, a control mechanism implemented by the frontoparietal network (Benoit & Anderson, 2012; Depue et al., 2007; Mary et al., 2020), critically relies on direct pathways
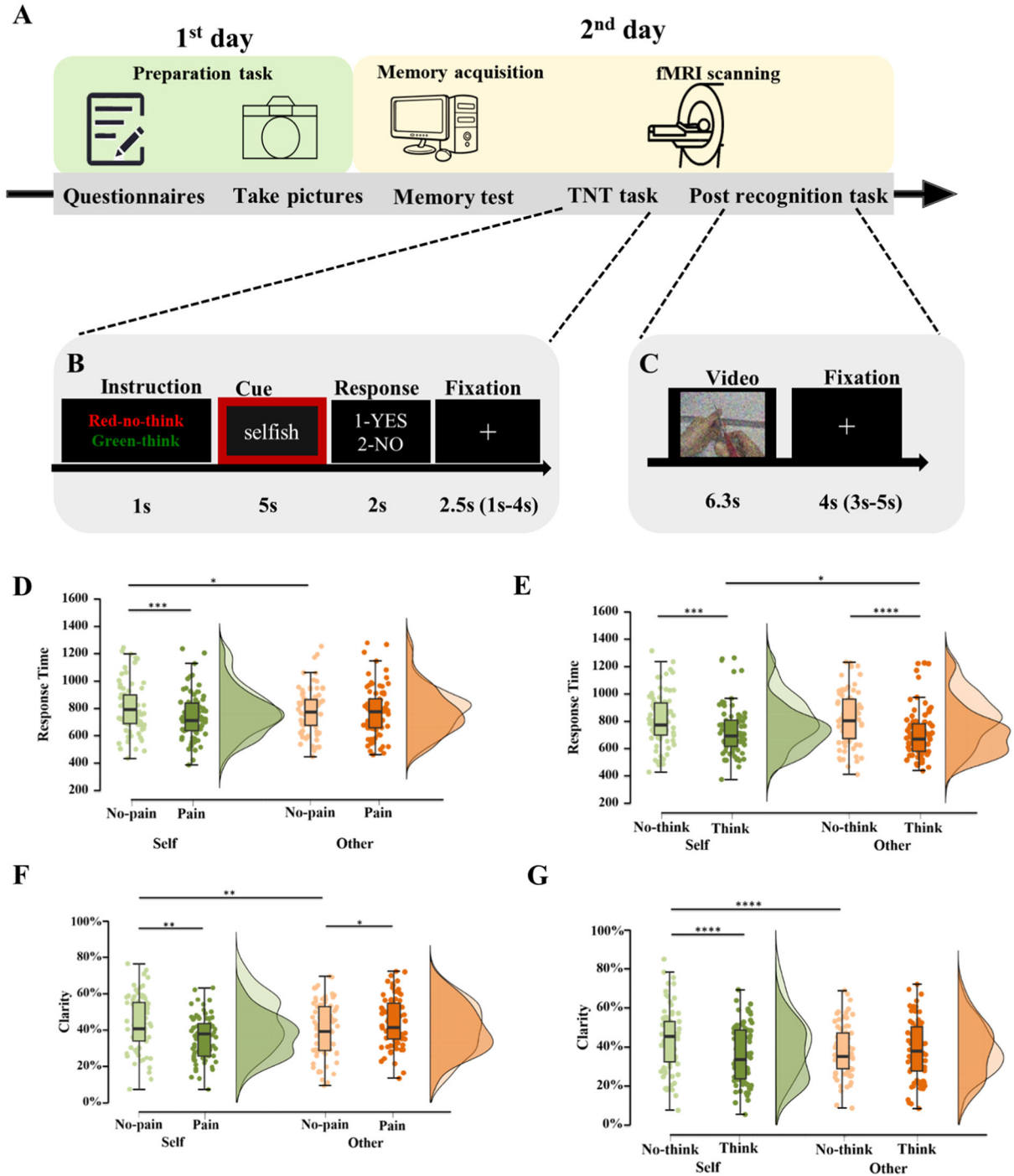

Fig. 1. Experimental protocol and behavioral result for think/no-think and following recognition task.

(A-C) Experimental flow chart. On the initial day, the subjects were asked to complete the questionnaires and the preparation task. On the following day, two fMRI scan sessions including Think/No-think (TNT) task (B) and post recognition task (C) were collected. (D-E) Post-hoc comparisons for the interaction effect between condition and pain (D), between condition and strategy (E) in terms of response time during TNT task. (F-G) Post-hoc comparisons for the interaction effect between condition and pain (F), between condition and strategy (G) in terms of clarity to recognize the content of the video clips during the following recognition task. Boxplots the center lines show the medians, box limits indicate the 25th and 75th percentiles, whiskers extend 1.5× the interquartile range from the 25th and 75th percentiles. *$p < 0.05$, **$p < 0.01$, ***$p < 0.005$, ****$p < 0.001$, post-hoc Bonferroni corrected tests for each comparison.

from the lateral prefrontal systems, including dorsolateral prefrontal cortex (dlPFC) and middle frontal gyrus (MFG), to the hippocampus (Gagnepain et al., 2017; Schmitz et al., 2017). This suggests that a higher level of dlPFC and MFG activation during the execution of emotion suppression is correlated with a decrease in hippocampal activation (Depue et al., 2007). This top-down suppression mechanism serves an adaptive purpose, aiming to counteract and regulate the involuntary emergence of intrusive memories into one's awareness (Gagnepain et al., 2017; Mary et al., 2020). These findings indicated that this ability to suppress emotions relies on inhibitory control mechanisms in the prefrontal cortex, which regulate the hippocampal retrieval process. Additionally, the neural circuitry encompassing the hippocampus, amygdala, anterior cingulate cortex (ACC), and prefrontal regions is essential for the regulation of emotional responses and facilitating the effective management of emotional states (Harnett et al., 2020; Fani et al., 2016; Yang et al., 2015; Milad et al., 2007; Zimmermann et al., 2017). However, previous studies have not explicitly delineated emotional suppression or memory suppression in response to self-experienced threats from those triggered by observed threats. This begs the question: are there distinct neural pathways for the trauma memories related to oneself versus those of others? If so, does the brain differentially suppress these two types of memories?

The think/no-think (TNT) paradigm, a suppression-induced or motivated forgetting paradigm, involves instructing subjects to either recall memories associated with a target cue (think condition) or employ a memory suppression strategy to inhibit such memories (no-think condition) (Herbert & Sütterlin, 2012). This paradigm can be used to assess the ability to prevent or suppress the re-emergence of intrusive memories (Gagnepain et al., 2017; Mary et al., 2020). Pain, being a prototypical unpleasant and highly aversive sensory and negative emotional experience (Agar-Wilson & Jackson, 2012; Koechlin et al., 2018), can be mitigated, to a certain extent, through effective emotion regulation strategies (Connelly et al., 2007; Montoya et al., 2004). Moreover, observing others in pain can trigger empathic distress (Baird et al., 2011; De Pascalis & Vecchio, 2022; Decety et al., 2010; Lamm et al., 2019). Therefore, painful stimuli can serve as suitable candidates for use in experiments exploring intrusive memories, particularly under self-related and other-related conditions.

Thus, in the present study, we aimed to utilize an updated TNT paradigm (Fig. 1A) in combination with functional magnetic resonance imaging (fMRI) utilizing both mass univariate analyses and multi-voxel pattern analyses to: (1) investigate the similarities and differences in behavioral responses (i.e. response time) and neural pathways during first-person pain experiences (self condition) and observing other's pain experience (other condition) while employing a memory suppression strategy; (2) examine the impact of suppression on subsequent memory retrieval; and (3) determine the associations between the first process of memory suppression and the second process of subsequent memory retrieval. To address these issues, subjects were required to memorize pairs of self/other-descriptive adjective words and painful/non-painful images. Upon instruction, they were required to recall the target painful stimulus associated with the adjective, with conditions including thinking about the stimulus and suppressing the memory through a "no-think" strategy. In the subsequent fMRI post-recognition task, subjects were required to identify the contents of dynamic video clips that displayed pictures undergoing gradual transformations from 0 % to 100 %, to assess the varying effects of suppression on memory retrieval in the context of self-experienced and other-based stimuli. We hypothesized that during the two-phase memory suppression process, there exist distinct differences between self-related and other-related stimuli, which are processed through different neural pathways.

Methods

Participants

A total of 85 (44 females, age range: 18–26 years) healthy Chinese University students were recruited in the current study. All subjects were right-handed, had normal or corrected-to-normal vision, no reported previous or current medical or psychiatric disorder, or current or regular use of medication. They were required to abstain from alcohol, caffeine or nicotine on the two days of the experiment. The research protocol was approved by the local ethics committee and was in accordance with the latest revision of the declaration of Helsinki. All subjects signed the written informed consent. Eight subjects were excluded due to 20 % of missing behavioral responses ($n = 1$), long reaction time during scanning task (z-score greater than 2.5, $n = 2$), excessive head motion larger than 3.8 mm (voxel size) for neuroimaging data ($n = 5$), resulting in 77 subjects (40 females, mean age ±SD: 20.13 ± 2.39 years) for the further analyses.

Experimental design

The study lasted two days as follows:

(1) On the first day subjects were required to complete the preparation task including taking photos and completing five mood and trait measuring questionnaires [Autism Spectrum Quotient (ASQ) (Ashwood et al., 2016), Becks Depression Inventory-II (BDI-II) (Bringmann et al., 2015), Trait Anxiety Inventory (TAI) (Knowles & Olatunji, 2020)] at the first day. (2) On the second day, two fMRI tasks including the TNT task and post recognition task were performed with the MRI scanner (Fig. 1A-C).

The preparation task

On the first day of arrival, each subject was required to choose 12 negative personality adjective words that were most relevant to describe himself or herself (self condition) from the validated Chinese word dataset (Chinese Affective Word System, CAWS, 60 words in total were selected) (Zhao et al., 2016). We also provided 12 negative adjective words from the same dataset to describe a same-gender stranger's personality. To prepare for the word-image pairs, we also required all subjects to take painful and non-painful pictures. The included materials for painful or non-painful pictures were divided into self-condition (scissors, stapler, hobby knife, safety pins, pins, bulldog clips) and other-condition (sewing needle, hammer, medical syringe with needle/cotton swab, small cactus potted plant, 1 kg dumbbell and binder clip). Specifically, each subject was instructed to imagine as vividly as possible that they were suffering from four of these injuries while taking the photos (scissors cutting the left thumb, a stapler piercing the left index finger, a sharp knife slicing the left index finger, a safety pin pricking the left index finger, a pin pricking the left index finger, a bulldog clip crushing the left index finger). The other-condition stimuli from two models were collected with these six events including a sewing needle injuring the index finger of the left hand, a hammer smashing the back of the left hand, a cactus pricking the index finger of the left hand, the tip of a syringe stabbing the index finger of the left hand, a binder clip pinching the index finger of the left hand, and a dumbbell smashing the left hand. Furthermore, we also collected non-painful pictures by asking subjects and models to touch the corresponding objects (see Fig. 1B for an example; the original materials are provided in Table S1). All pictures were gender matched between self and other condition and were consistent with same lamination and resolution using Photoshop CS6.

Think-No-Think fMRI task

On the second day of arrival, all subjects were instructed to memorize 24 words (12-self-related and 12 other-related items), that had been selected on the previous day. They were then required to undergo a memory test, in which they had to distinguish whether the presented

word described themselves or others. It was crucial that their memory accuracy reached 100 %; if not, they were required to repeat the test until they attained this benchmark. Following this, the 24 words were randomly paired with specific images, consisting of 12 pain and 12 no-pain-related pictures. Half of these pictures (6 pain pictures, 6 no-pain pictures) were personal photographs of each participant taken by themselves, while the remaining images were sourced from two independent experimenter models taken one female and one male, who were not part of the experiment. All participants were subsequently tasked with memorizing these 24 word-picture pairs, and it was imperative that they accurately recalled the matched image upon the presentation of the cue word, achieving 100 % accuracy.

Next, the formal fMRI scanning started with the TNT paradigm modified based on a previous study (Mary et al., 2020). Twenty-four prepared pairs were randomly assigned into three conditions (think-8 pairs, no-think-8 pairs, baseline-8 pairs) and only think and no-think conditions were included in the TNT scanning task. A total of 80 trials (2 runs, 40 trials per run, 16 pictures were repeated 5 times) incorporating eight conditions [2 condition (self vs. other) *2 pain (pain vs. no-pain) *2 strategy (think vs. no-think)] were presented to subjects. Each trial started with a cue instruction (i.e. green frame indicating think condition while red for no-think condition, 1 s). During the think session (green frame), subjects were asked to recall the matched picture as vividly as possible when the cue word was presented in the center of the screen (5 s). During no-think session (red frame), subjects were required to use a memory suppression strategy to keep the corresponding picture out of mind when they saw the cue word. Afterwards, they were asked to answer whether the corresponding picture came into mind via pressing buttons (2 s). Following a jitter (1s-4 s), a new trial started (Fig. 1B). The programming of current TNT task was presented in E-prime 2.0.

Post recognition fMRI task

The recognition task following the TNT task aimed to examine the ability of memory suppression between self and others. Individual 24 pictures, including baseline condition, were blurred using variable levels of morphing (from 100 % to 0 %, in 5 % steps) using a 'gaussian' transform with Matlab. Each trial started with each video tape comprising 21 gradually transformed pictures (0 %-100 %) lasting 6.3 s (21* 0.3 s). Between trials, there was a 4 s average inter-stimulus interval. All subjects were instructed to watch carefully and press the button as fast as possible when they recognized the content of the video. This task comprised three runs and each run consisted of 40 trials (24 videos were repeated 5 times) (Fig. 1C). Video stimuli were presented using Psychopy (V2021.2.3).

Brain imaging data acquisition and preprocessing

The brain imaging data were collected by a 3.0 T GE Discovery MR750 system (General Electric Medical System, Milwaukee, WI) which has an eight channel-phased array head coil. Functional MRI data were acquired using a T2*-weighted echo-planar imaging (EPI) pulse sequence (repetition time = 2 s, echo time = 30 ms, 36 slices, slice thickness = 3.8 mm, flip angle = 90°, voxel size = 3.13 × 3.13 × 3.8 mm). High-resolution T1-weighted anatomical images were acquired to improve the normalization of functional images, employing a spoiled gradient-recalled (SPGR) sequence with the parameters: TR = 6 ms, TE = 2 ms, flip angle = 9°, FOV = 256 × 256 mm, voxel size = 1 mm isotropic, consisting of 156 slices with a slice thickness of 1 mm. The functional MRI data were preprocessed by *fMRIPrep* 21.0.0.

The first 5 vol of each run were discarded to allow MRI T1 equilibration. Before the spatiotemporal filtering, head-motion was corrected (twenty-four motion-related nuisance, six head movement parameters and their squares, their derivatives and squared derivatives). The anatomical images were segmented into different tissues including grey matter, white matter and cerebrospinal fluid. The functional images

obtained following the segmentation procedure were subsequently normalized using Montreal Neurological Institute (MNI) space. Then the data were spatially smoothed with an 8 mm full-width at half maximum (FWHM) Gaussian kernel (more details please see Supplementary Material).

Behavioral data analyses

To maintain consistency across the two tasks, an identical repeated ANOVA with condition (self, other), pain (pain, no-pain) and strategy (think, no think) as within-subject factors was conducted for both TNT (i.e. reaction times) and recognition tasks (i.e. clarity percentage = recognition response time/ the total duration of each video). All data analyses were performed in R 4.1.3 (packages: tidyversse, ggpubr, rstatix). F/P values and pes (partial eta squared) values (effect size, 0.01-small size of effect, 0.06-medium size, 0.14- large size) are reported (Richardson, 2011).

Neuroimaging data analyses

Mass univariate general linear model (GLM) analyses

A two-level random effects GLM analysis was conducted on the fMRI signal using a mass univariate GLM analyses with SPM12. The first level model included ten regressors (self-pain-think, self-pain-nothink, self-nopain-think, self-nopain-nothink, other-pain-think, other -pain-nothink, other -nopain-think, other -nopain-nothink, cue and response) for the TNT task. For the recognition task, we modeled these regressors including self-pain-think, self-pain-nothink, self-nopain-think, self-nopain-nothink, other-pain-think, other-pain-nothink, other-nopain-think, other-nopain-nothink, and baseline (self-pain-baseline, self-nopain-baseline, other-pain-baseline, other-nopain-baseline) in the analysis. Response time was added as a covariate variable in the model. The boxcar regressors were convolved with the canonical hemodynamic response function (HRF) and processed by a high-pass filter of 128 s. Regressors of non-interest included six head movement parameters, their squares, their derivatives and squared derivatives resulting in a total of twenty-four motion-related nuisance regressors. Group-level models were performed to independently determine the behavior-based interaction effects for each task, specifically examining the interactions between condition and pain, condition and strategy. These analyses were conducted using ANOVA with family-wise error (FWE) correction to ensure the results were interpretable and explainable, with a significance level of $p < 0.05$, and a voxel threshold of 10 voxels was applied. To disentangle significant interaction effects, post-hoc analysis was performed on extracted parameter estimates using MarsBar toolbox (http://marsbar.sourceforge.net) to render the respective clusters as a mask.

Functional-connectivity analysis between TNT and recognition

Based on the whole-brain analysis results during the TNT task, the ROIs (dorsal mPFC/ dACC (x, y, z 8, 14, 50/−11, 28, 28/2, 20, 40)) that exhibited an interaction between condition and strategy were defined as seed regions. To further explore the relationship between TNT and post-recognition task-based functional connectivity, a generalized psycho-physiological interaction analyses (gPPI, https://www.nitrc.org/projects/gppi) approach (McLaren et al., 2012) was employed, focusing on the contrast between self and other during memory suppression processing with False Discovery Rate (FDR) $p < 0.05$ correction. Additionally, recognition response time and head motion were included as covariate variables into the model.

Multi-voxel pattern analysis (MVPA) across two tasks

To examine whether common brain representations underlie memory suppression and post recognition across conditions, we conducted a linear support vector machine (SVM, C = 1) to develop multivariate pattern classifiers between two conditions during TNT task and

recognition task separately from the whole-brain level (with a gray matter mask) based on Canlabcore toolbox (https://github.com/canlab/CanlabCore). The patterns of classifiers were trained on subject-wise univariate first level contrasts across two tasks. We assessed the accuracy of the SVMs classification with a leave-one-subject-out cross validation (LOOCV). Subsequently, we employed a whole brain searchlight-based approach (radius size of 3 voxels) to confirm the regions that could predict the corresponding two conditions across two tasks with a LOOCV procedure ($p < 0.05$ with FDR correction).

To explore the common neural basis across conditions, within-subject neural representation similarity score was calculated using Pearson correlation based on the validated searchlight-based pattern. First, single-trial estimates from the GLM each trial was run separately for the TNT task and following recognition task. The estimated beta value obtained for each trial and subject were then used to compute the neural representation similarity score. We defined masks based on the aforementioned searchlight patterns and calculated the Pearson correlation for each subject across all trials within specific conditions. Paired t tests were conducted to assess differences in neural representation similarity across conditions. Furthermore, to validate the consistency in neural processing between self-related pain and observing others' pain, we calculated the stimulus intensity-independent pain signature-1 (SIIPS1) pattern weights (Woo et al., 2017). Specifically, we calculated the pattern expression values for self-pain and other-pain conditions across both tasks using this developed pattern. Subsequently, Pearson correlation between self-pain and other-pain based on SIIPS1 was computed separately for the TNT task and post-recognition task. The P values obtained were corrected for the number of correlations performed (i.e., 2 correlations).

Cross-task neural pattern similarity

Given the strong correlation between the two tasks, we also investigated the cross-task neural pattern similarity (NPS) between the TNT and post recognition tasks, examining both inter-subject and intra-subject patterns in the specific brain regions based on the validated results, such as the dACC. For the intra-subject cross-task neural pattern similarity, we calculated the similarity of neural representation between TNT task and recognition tasks within the same conditions respectively (self-think, self no-think, other think, other no-think, self-pain, self no-pain, other pain, other no-pain) with Pearson correlation in terms of single trials for each subject. Paired *t*-tests were used to determine if the intra-subject TNT-recognition neural similarity score differed significantly from 0 (indicating no relationship). Additionally, we examined if there were any differences in the intra-subject TNT-recognition NPS across conditions. For the inter-subject cross-task NPS, we calculated the correlation between the neural patterns of one subject ($i = 1$) and the remaining subjects ($n-i = 76$) in the same conditions between the TNT and recognition tasks. We then averaged these correlations to obtain the between-subject cross-task NPS score for the specific brain region. Finally, we conducted paired *t*-tests on the intra-subject cross-task pattern similarity scores for each subject to compare if the intra-subject TNT-recognition NPS was higher than the inter-subject TNT-recognition neural correlation.

Results

Behavioral results

Think-no-Think task

During the memory suppression phase (i.e. TNT task), the subjects accurately used the corresponding strategy during both think (mean $\pm$ SD, 89.2 % $\pm$ 10.2 %, responded "yes") and no-think conditions (mean $\pm$ SD, 74.1 % $\pm$18.3 %, responded "no"). Three-way repeated ANOVA results showed that there were significant main effects of pain [$F_{(1,76)}$ =9.31, $P = 0.003$, partial eta squared (pes)=0.109] and strategy [$F_{(1,76)}$ =30.90, $P < 0.001$, pes=0.289] but not for the condition [$F_{(1,76)}$ =0.20,

$P = 0.65$]. Moreover, two-way interactions between condition and pain [$F_{(1,76)}$ =6.86, $P = 0.011$, pes=0.083] and between condition and strategy [$F_{(1,76)}$ =4.66, $P = 0.034$, pes=0.058] were significant. Post-hoc comparisons with Bonferroni correction suggested that individuals had shorter RTs to painful rather than non-painful stimuli under the self-condition ($P_{bon} <0.001$) but not under the other condition (P_{bon}=0.48) indicating that individuals are more protective of self-related rather than other-related threatening stimuli (Fig. 1D). In addition, individuals responded faster for the think-relative to the no-think condition during both self-related stimuli ($P_{bon} <0.001$) and other-related stimuli ($P_{bon} <0.001$) suggesting that the suppression memory strategy was effective regardless of whether the stimuli were associated with the self or others (Fig. 1E).

The post-recognition task

Since we did not find any significant differences between think, no-think and baseline conditions, we excluded the baseline condition from further analysis (Ps >0.14). During the memory retrieval phase (i.e. post-recognition task), three-way repeated ANOVA results showed that there were significant main effects of strategy [$F_{(1,76)}$ =16.70, $P < 0.001$, pes=0.180] and significant two-way interactions (condition*pain: [$F_{(1,76)}$ =13.56, $P < 0.001$, pes=0.151]; condition*strategy: [$F_{(1,76)}$ =39.96, $P < 0.001$, pes=0.345]). Post-hoc comparisons of two-way interactions with Bonferroni correction suggested that individuals responded faster for self-related painful than non-painful stimuli (P_{bon}=0.002) but the opposite for the other-condition (P_{bon}=0.041, see Fig. 1F). Individuals responded slower using the no-think strategy relative to the think strategy (P_{bon}<0.001) for self-related video tapes, however no difference was found between no-think and think strategies for others (P_{bon}=0.816). Additionally, individuals retrieved self-related stimuli slower than other-related stimuli following the inhibitory strategy (P_{bon}<0.001, Fig. 1G). No main effects of condition [$F_{(1,76)}$ =1.32, $P = 0.255$], or pain [$F_{(1,76)}$=0.241, $P = 0.625$], or interactions between pain and strategy [$F_{(1,76)}$ =3.49, $P = 0.066$] or three-way interactions [$F_{(1,76)}$ =0.09, $P = 0.766$] were observed. Correlation analyses suggested that the response time during TNT under each condition was not associated with the subsequent recognition performance (Ps > 0.072).

Mass univariate general linear model (GLM) analyses

TNT GLM results

For the GLM results during memory suppression phase, the brain regions involved in the main effect of pain included supramarginal gyrus and inferior parietal lobule and the main effect of strategy involved the dlPFC, inferior parietal lobule (IPL), inferior frontal gyrus (IFG), insula, hippocampus and other regions (see Table 1, Figure S1). Most importantly, an interaction between condition and strategy was found in the dorsal medial prefrontal cortex (dmPFC)/ dorsal anterior cingulate cortex (dACC) (MNI coordinate: 8, 14, 50/−11, 28, 28/2, 20, 40, $k = 356$, $F = 17.20$, $P_{FWE} = 0.017$). Post-hoc comparisons with the extracted parameter estimates indicated greater responses to the suppression strategy relative to the think strategy in both self (P_{bon}=0.001) and other (P_{bon}<0.001) conditions, consistent with our behavioral results (Fig. 2 right).

Post-recognition GLM results

During the post memory retrieval phase, a main effect of pain was found in precuneus, posterior cingulate cortex (PCC), middle occipital gyrus (MOG) and superior occipital gyrus (SOG) and a main effect of strategy in right MOG, middle temporal gyrus (MTG) and fusiform gyrus (see Table 2, Figure S1). A two-way interaction between condition and pain was observed in postcentral gyrus (x, y, z, 2, −29, 74), precentral gyrus (x, y, z, 6, −21, 80) and precuneus (x, y, z, −9, −43, 76, $k = 351$, $F = 17.38$, $P_{FWE} = 0.018$) and post-hoc comparisons suggested reduced responses to self-related painful stimuli relative to non-painful stimuli (P_{bon}= 0.039), however the pattern was opposite in other condition

Table 1
Univariate general linear model results for TNT task.

	Hemisphere	Region	Cluster p (FWE)	K	Peak F/T	MNI x	y	z
Main effect of Pain	Right	Supramarginal Gyrus/Inferior Parietal Lobule	0.030	309	19.75	58	−29	40
Main effect of Strategy		Supramarginal Gyrus	0.001	41,772	240.80	62	−45	34
		Dorsal Lateral Prefrontal Cortex			235.05	48	4	54
		Superior Frontal Gyrus/ Supplementary motor area			230.60	10	14	66
	Left	Cuneus	0.001	9748	106.57	−15	−59	24
		Inferior Parietal Lobule			92.11	−35	−79	40
	Right	Precuneus			89.68	14	−57	24
	Left	Supramarginal Gyrus	0.001	1099	84.38	−61	−51	34
		Inferior Frontal Gyrus	0.001	4774	82.40	−47	14	−7
		Inferior Frontal Gyrus/Orbitofrontal Cortex			75.31	−39	20	−13
		Insula			59.88	−39	10	2
		Lingual Gyrus			31.26	−9	−91	−13
		Inferior Frontal Gyrus	0.001	1798	58.94	−45	30	16
		Inferior Frontal Gyrus/ Precentral Gyrus			35.91	−57	14	38
		Inferior Frontal Gyrus			24.34	−47	46	4
		Middle Temporal Gyrus	0.001	861	39.04	−57	−51	−9
		Hippocampus			21.12	−37	−29	−13
	Right	Pallidum	0.022	333	35.29	22	−1	4
		Putamen			24.53	26	−9	6
		Postcentral Gyrus	0.042	281	27.60	58	−17	52
		Precentral Gyrus	0.026	322	24.51	34	−17	72
		Precentral Gyrus			17.07	36	−27	72
		Postcentral Gyrus			13.02	34	−29	48
Interaction effect of Condition*Strategy	Left	Dorsal Medial Prefrontal Cortex(dmPFC)	0.017	356	17.20	−11	28	28
	Right	Dorsal Medial Prefrontal Cortex(dmPFC)			16.48	8	14	50
		Dorsal Anterior Cingulate Cortex(dACC)			15.89	2	20	40
Interaction effect of Condition*Pain								
Self No-pain vs. Self Pain	Right	Inferior Parietal Lobule	0.016	467	4.34	38	−41	38
		Inferior Parietal Lobule			4.07	38	−49	46
		Supramarginal Gyrus			3.66	66	−31	40
Self No-pain vs. Other No-pain	Left	Inferior Parietal Lobule	0.001	1029	4.81	−43	−49	52
		Inferior Parietal Lobule			4.30	−45	−47	40
		Inferior Parietal Lobule			3.93	−41	−57	62

(P$_{bon}$=0.029, Fig. 2 left). Additionally, based on behavioral findings, exploratory analysis suggested that greater activation was found in cuneus/posterior cingulate cortex (x, y, z, 10, −73, 6) and lingual gyrus (x, y, z, 10, −67, −3, k = 1127, T = 5.99, P$_{FWE}$ = 0.001) in the other no-think condition compared to the self no-think one (more details see Table 2).

Functional-connectivity analysis between TNT and recognition

Given that an interaction between condition and strategy was found in dmPFC and dACC during the memory suppression phase, we defined these ROIs as seed regions (dmPFC: 8, 14, 50/−11, 28, 28; dACC: 2, 20, 40) respectively. The results indicated that specifically stronger dACC-insula (x, y, z, −39, 12, 12, P$_{FDR}$ = 0.0045, k = 255, T = 4.77), dACC- left MFG (x, y, z, −39, 36, 30, P$_{FDR}$ = 0.043, k = 272, T = 6.06) and dACC-right MFG (x, y, z, 36, 42, 34, P$_{FDR}$ = 0.033, k = 317, T = 5.70) connectivity was found in self condition compared to other condition for the inhibited stimuli during the subsequent recognition task (Fig. 3).

Multi-voxel pattern analysis, searchlight and pattern similarity results

Multi-voxel and pattern similarity results of TNT task

We firstly used the whole gray mask to distinguish the whole-brain patterns between using think and suppression strategies under self and other conditions during the TNT session.

The results suggested that the patterns accurately classified the think vs. no-think under the self-condition [accuracy ± SE= 84 % ± 2.9 %, P < 0.001, sensitivity= 86 %, CI (77 %−93 %), specificity =83 % CI (74 %−91 %)] and other-condition [accuracy ± SE= 83 % ±3.0 %, P < 0.001, sensitivity = 84 % CI (76 %−92 %), specificity =82 % CI (73 %− 90 %)]. More specifically, searchlight-based analyses suggested that this pattern including left MTG, orbitofrontal cortex (OFC), ACC, right IPL, hippocampus, IFG and putamen can successfully classify self think vs. self no-think (accuracy range 58 %−79 %, P$_{FDR}$<0.05, Fig. 4A). On the other hand, a distinct pattern including MTG, IFG, insula, ACC, IPL, supramarginal gyrus, hippocampus, PCC, precuneus and medial prefrontal cortex (mPFC) can successfully classify other think and other no-think (accuracy range 59 %−80 %) (P$_{FDR}$<0.05, Fig. 4B, more details see Table S2).

We also calculated the neural representation similarity scores based on searchlight results across conditions using single trial analysis and the higher the classification accuracy was the lower was the similarity score. The results indicated that the neural pattern similarities of within-think [self: mean ± SD, 0.277 ± 0.088, other: mean ± SD, 0.249 ± 0.074] and within-no-think [self: mean ± SD, 0.323 ± 0.099, other: mean ± SD, 0.323±0.106] were significantly higher than between think and no think under the self [mean ± SD, 0.193 ± 0.074, Ps <0.001, Fig. 4A] and other conditions [mean ± SD, 0.181 ± 0.067, Ps <0.001, Fig. 4B], consistently indicating the different neural representations between think and no-think independent of self or other conditions.

Multi-voxel and pattern similarity results for the recognition task

Moreover, during the recognition task, the pattern can successfully classify the difference between pain vs. no-pain under the self condition [accuracy ± SE= 66 % ± 3.8 %, P < 0.001, sensitivity = 69 % CI (58 %− 79 %), specificity =64 % CI (53 %−73 %)] and other condition [accuracy ± SE= 60 % ± 4 %, P = 0.02, sensitivity = 61 % CI (50 %−72 %), specificity =58 % CI (48 %−70 %)]. More specifically, searchlight-based analyses suggested that this pattern including left and right para-hippocampus gyrus, left hippocampus, left IFG, right insula, left mPFC, ACC and PCC can classify self pain and self no-pain (accuracy
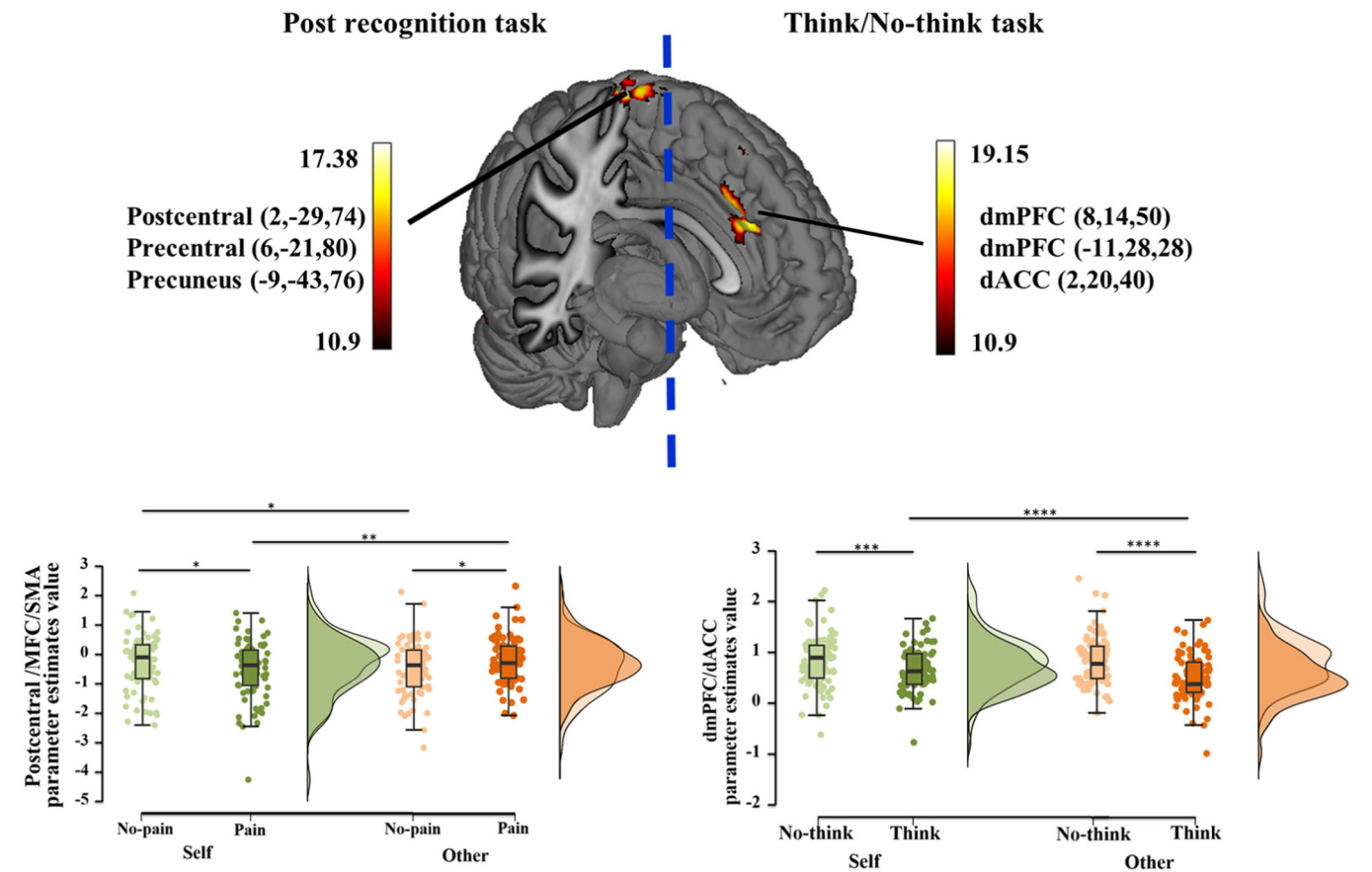
Fig. 2. Univariate general linear model (GLM) and functional-connectivity between TNT and post recognition results.
GLM results of interaction effect between condition and strategy, along with a post-hoc analysis during TNT (right), and between condition and pain, along with a post-hoc analysis during post recognition task (left). $*p < 0.05$, $**p < 0.01$, $***p < 0.005$, $****p < 0.001$ post-hoc Bonferroni corrected tests for each comparison.

Table 2
Univariate general linear model results for post recognition task.

	Hemisphere	Region	Cluster p(FWE)	K	Peak F	MNI x	y	z
Main effect of Pain	Left	Precuneus/ Cuneus	0.029	311	33.12	−19	−59	18
		Posterior Cingulate Cortex			14.75	−13	−51	8
	Right	Cuneus/Middle Occipital Gyrus	0.007	434	21.12	16	−97	8
		Superior Occipital Gyrus			14.42	18	−99	22
Main effect of Strategy								
No-think vs. Think	Left	Fusiform Gyrus	0.002	725	4.99	−39	−59	−19
	Right	Fusiform Gyrus	0.001	3376	4.63	44	−71	−19
		Fusiform Gyrus			4.47	34	−47	−17
		Middle Occipital Gyrus			4.32	22	−105	−1
		Middle Temporal Gyrus	0.015	469	3.72	40	−79	10
		Middle Temporal Gyrus/Middle Occipital Gyrus			3.72	36	−79	18
		Middle Occipital Gyrus			3.66	30	−81	30
Interaction effect of Condition*Pain								
	Left	Postcentral Gyrus	0.018	351	17.38	2	−29	74
		Precentral Gyrus			14.8	6	−21	80
		Precuneus			11.37	−9	−43	76
Interaction effect of Condition*Strategy								
Self Think vs. Other Think		Cuneus	0.043	346	5.03	−5	−85	12
Self No-think vs. Other No-think		Cuneus	0.024	418	5.21	−11	−85	8
		Cuneus			4.5	−5	−89	22
Other No-think vs. Self No-think	Right	Cuneus/Posterior Cingulate Cortex	0.001	1127	5.99	10	−73	6
		Lingual Gyrus			5.76	10	−67	−3

range 54 %−63 % $P_{FDR}<0.05$, Fig. 4C), on the other hand, the pattern including right para-hippocampus, inferior temporal gyrus (ITG), insula, PCC and medial frontal gyrus (MFG) can classify other pain and other no-pain (accuracy range 54 %−62 %, $P_{FDR}<0.05$, Fig. 4D, more details

see Table S3). The neural representation similarity scores of within-pain [self: mean ± SD, 0.210±0.097, other: mean ± SD, 0.132±0.060] and within-no pain trials [self: mean ± SD, 0.189±0.084, other: mean ± SD, 0.139±0.061] were greater than the between pain and no pain trials
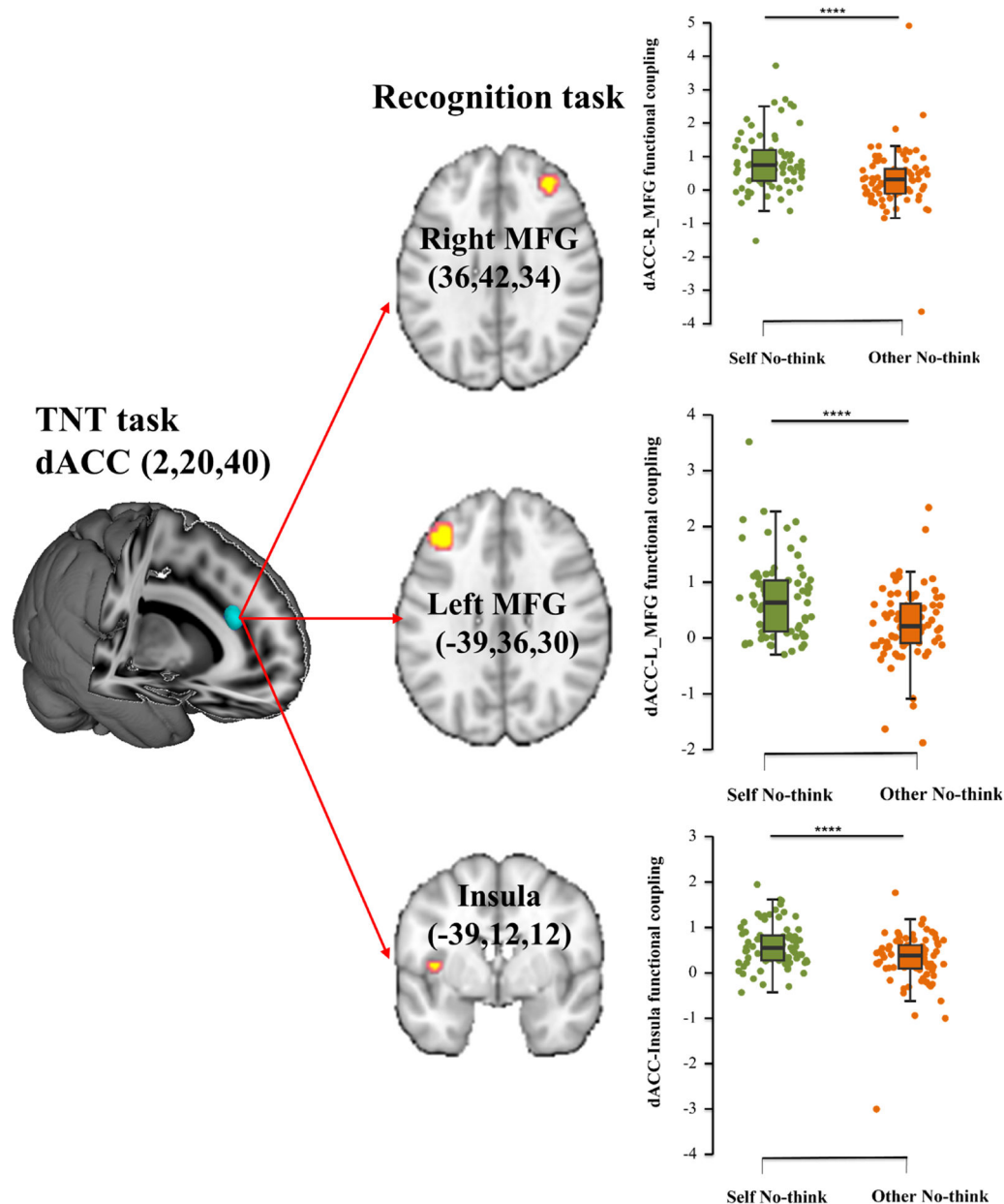
Fig. 3. Functional-connectivity between TNT and post recognition task.
Functional connectivity between a specific seed region in the dorsal anterior cingulate cortex (dACC), identified from TNT task, and bilateral middle frontal gyrus (MFG) as well as the insula to self no-think vs. other no-think conditions with post-hoc analyses during post recognition task. ****$p < 0.001$ post-hoc Bonferroni corrected tests for each comparison.

(Ps< 0.001) regardless of self (mean $\pm$ SD, 0.164 $\pm$ 0.086, Fig. 4C) or other conditions (mean $\pm$ SD, 0.102 $\pm$ 0.059, Fig. 4D).

The pattern can successfully classify the difference between the self vs. other under the pain [accuracy $\pm$ SE= 71 % $\pm$ 3.7 %, $P < 0.001$, sensitivity = 70 % CI (59 %−80 %), specificity =71 % CI (61 %−81 %)] and no-pain conditions [accuracy $\pm$ SE= 69 % $\pm$ 3.7 %, $P < 0.001$, sensitivity = 69 % CI (58 %−79 %), specificity =70 % CI (60 %−80 %)]. Searchlight-based analyses suggested that this pattern including parahippocampus, superior temporal gyrus (STG), MFG, MTG, cuneus, superior parietal lobule (SPL) can classify self pain vs. other pain (accuracy range 50 %−60 % P$_{FDR}$<0.05, Fig. 5A). The searchlight-based pattern including parahippocampus, fusiform, ITG, pallidum, hippocampus, amygdala, MTG, insula, cuneus, precuneus, postcentral, mPFC, PCC can classify self no-pain vs. other no-pain (accuracy range 50 %−58 % P$_{FDR}$<0.05, Figure S2A, more details see Table S3). The neural pattern

similarity of within-self [pain: mean $\pm$ SD, 0.446 $\pm$ 0.188, no pain: mean $\pm$ SD, 0.410 $\pm$ 0.172] and within-other [pain: mean $\pm$ SD, 0.427 $\pm$ 0.181, no pain: mean $\pm$ SD, 0.406 $\pm$ 0.183] was larger than between self and other no matter for pain (mean $\pm$ SD, 0.398 $\pm$ 0.185, Fig. 5A) or no pain (mean $\pm$ SD, 0.364 $\pm$ 0.181, Figure S2A) (Ps< 0.001).

To validate the consistency of our results, we calculated the Pearson correlation between the self-pain and other-pain weights in the TNT and post recognition tasks separately based on the SIIPS1 signature including PCC, mPFC, precentral gyrus, supplementary motor area (SMA), putamen, parahippocampus (Fig. 5B). The results indicated that self-pain was highly correlated with other-pain of SIIPS1 in the TNT ($r =$ 0.89, $p < 0.001$, corrected for 2 correlations, Fig. 5C) and post recognition tasks ($r =$ 0.73, $p < 0.001$ corrected for 2 correlations, Fig. 5D), suggesting a common basis between self-pain and other pain processing during memory retrieval.
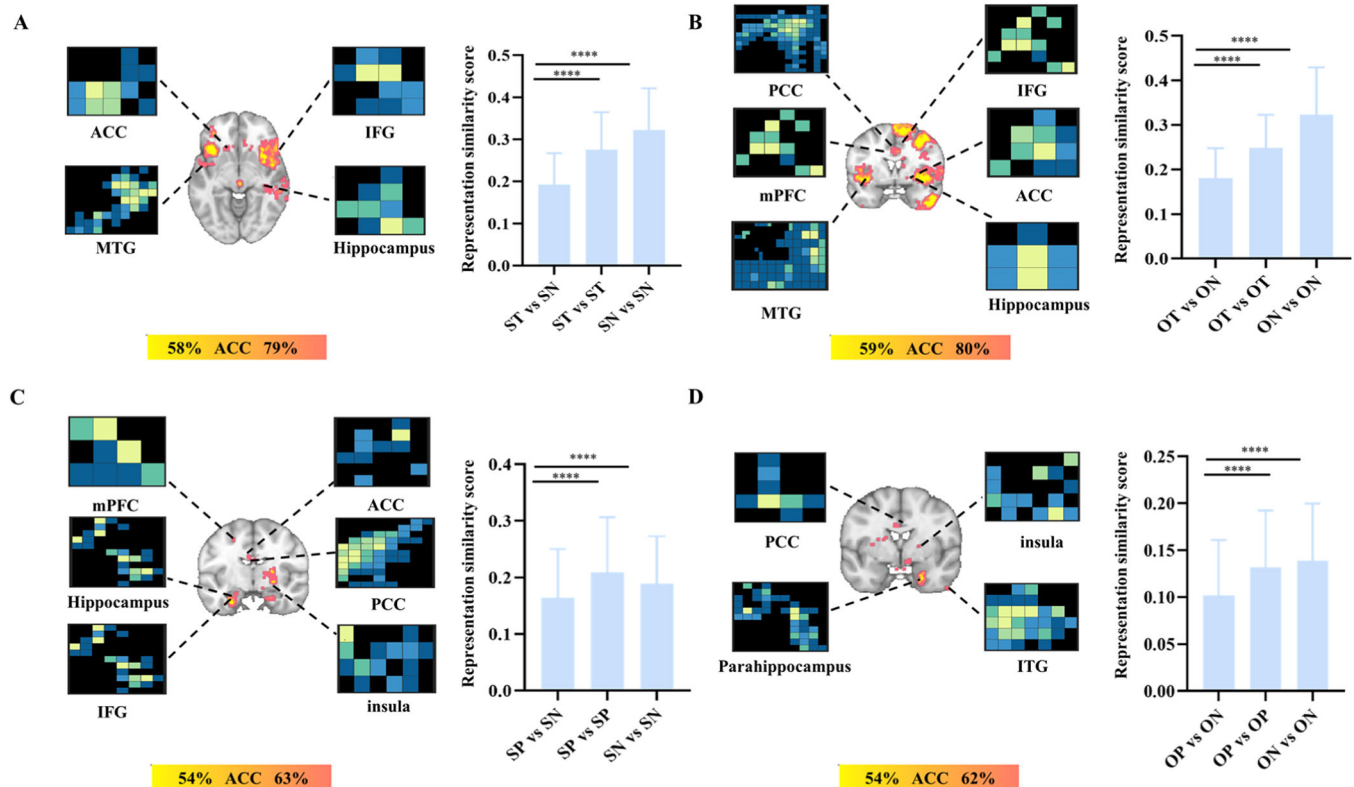
Fig. 4. The searchlight and neural representation similarity results.

The searchlight and neural representation similarity results for TNT task (A. self-think (ST) vs. self no-think (SN); B. other think (OT) vs. other no-think (ON)) and post recognition task (C. self-pain (SP) vs. self no-pain (SN); D. other pain (OP) vs. other no-pain (ON)). Note, ACC, anterior cingulate cortex; IFG, inferior frontal gyrus; MTG, middle temporal gyrus; PCC, posterior cingulate cortex; mPFC, medial prefrontal cortex; ITG, inferior temporal gyrus. FDR correction pFDR < 0.05 for searchlight analyses. ****p < 0.001 Bonferroni corrected tests for each comparison in terms of representation similarity score.

In addition, the pattern can classify the self no-think vs. other no-think [accuracy ± SE= 73 % ± 3.6 %, P < 0.001, sensitivity= 74 %, CI (64 %−84 %), specificity =71 % CI (61 %−81 %)]. The contributing regions could classify self no-think vs. other no-think and included parahippocampus, STG, PCC, amygdala, IFG, MFG, hippocampus and IPL (accuracy range 52 %−68 %, P_{FDR}<0.05, Fig. 6A, more details see Table S3). The similarity score between self no-think and other no-think from searchlight-based regions is 0.418 (mean ± SD, 0.418±0.172), which was lower than within self no-think [mean ± SD, 0.452±0.176] and within other no-think [mean ± SD, 0.466±0.169, Fig. 6B] similarity (Ps < 0.001).

The pattern can classify the self think vs. other think [accuracy ± SE= 62 % ± 3.9 %, P < 0.005, sensitivity= 61 %, CI (50 %−72 %), specificity =62 % CI (51 %−73 %)]. The searchlight-based analyses suggested that ITG, parahippocampus, MFG, ACC, precuneus, cuneus, insula, postcentral, SMA can classify self think vs. other think (accuracy range 53 %−61 %, P_{FDR}<0.05, Figure S2B, more details see Table S3). The similarity score between self think and other think from the above regions is 0.201 (mean ± SD, 0.201 ± 0.129) and was smaller than within self think [mean ± SD, 0.273±0.127] and within other think [mean ± SD, 0.252±0.124] similarity (Ps <0.001) (Figure S2B).

Cross-task neural pattern similarity

We calculated the cross-task NPS (Fig. 6C) between TNT-recognition in dACC in these conditions (self-think, self no-think, other think, other no-think, self pain, self no-pain, other pain, other no-pain) and compared whether there were differences between inter- and intra-subjects' similarity scores. The results suggested that the inter-subject similarity scores in these conditions were larger than the intra-subject similarity (Ps <0.001) showing that each individual had a partially

unique set of similarity representations for TNT and recognition tasks in the dACC to produce the painful and no-painful in think and no-think conditions (Fig. 6D and Table 3). There were consistently observed differences in cross-task NPS between self-think and self-no-think conditions (self-think < self no-think, intra-subject: t = 2.672, P_{FDR} =0.024; inter-subject: T = 5.878, P_{FDR}=0.0008), as well as between other think and other no-think (other think < other no-think, intra-subject: T = 2.213, P_{FDR} =0.06; inter-subject: T = 4.398, P_{FDR}=0.0008) conditions. These findings, analyzed both intra- and inter-subject levels and indicate a higher consistency between the TNT task and subsequent memory retrieval when using a memory suppression strategy. Additionally, given the inter-subject cross-task NPS, a higher consistency was found in the self no-think condition compared to the other no-think condition (T = 3.36, P_{FDR} =0.005), suggesting a greater efficiency in suppressing self-related memory. In turn, these findings provide evidence for the behavioral results, indicating that it is more challenging to retrieve self-related memories (i.e. a higher clarity of the video stimuli for retrieval) than other-related memories.

Discussion

In the current study, we explored whether there are differences between one's own pain experience and observing others' pain while utilizing a memory suppression strategy and its modulation of subsequent memory retrieval. Overall, our findings indicated that the recognition process following memory suppression requires greater cognitive engagement for self-related stimuli compared to other-related stimuli. This increased engagement associated with greater activity in the dorsal anterior cingulate cortex. Additionally, we observed a more robust functional coupling between the dACC, which was found during memory suppression, and both the middle frontal gyrus and the insula during
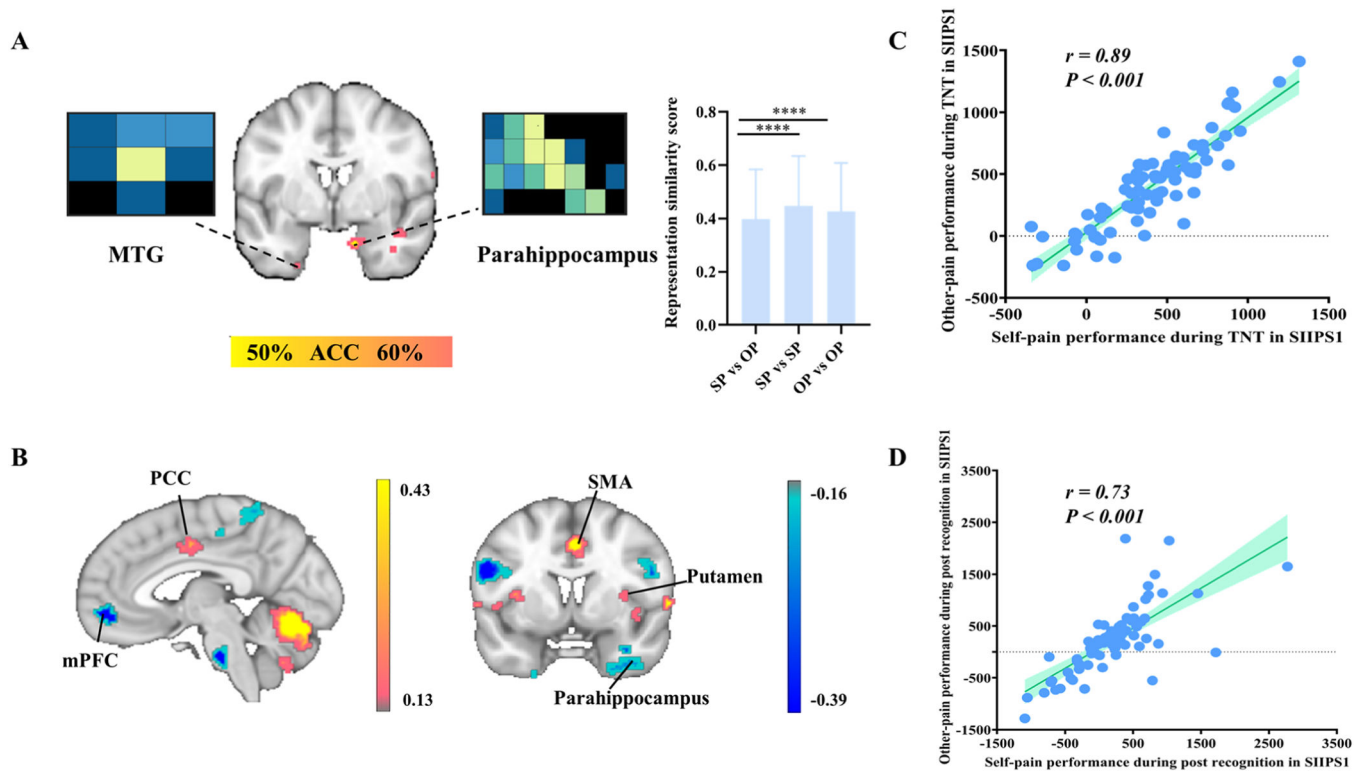
Fig. 5. The searchlight and neural representation similarity results for recognition task and validation results from SIIPS-1.
(A) The searchlight and neural representation similarity results between self pain (SP) and other pain (OP) during post recognition task. (B) Signature of stimulus intensity-independent pain signature-1(SIIPS-1). (C-D) The Pearson correlation between the performance in SIIPS-1 of self pain and other pain during TNT task (C) and the post recognition task (D). Note, MTG, middle temporal gyrus; PCC, posterior cingulate cortex; mPFC, medial prefrontal cortex; SMA, supplementary motor area. FDR correction pFDR < 0.05 for searchlight analyses. ****$p < 0.001$ Bonferroni corrected tests for each comparison in terms of representation similarity score. The P values of Pearson correlation were corrected for the number of correlations ($n = 2$).

self-related recognition, as opposed to other-related recognition. An advanced multivariate pattern analysis further confirmed that neural responses from the limbic system (i.e. amygdala, hippocampus) and empathy network (i.e. IPL, STG, IFG) can accurately distinguish between self-related and other-related recognition after memory suppression. Our results demonstrate distinct neural representations of memory suppression associated with one's self and others.

The current study utilized the think/no-think paradigm, which assesses inhibitory control of a newly acquired memory by evaluating the ability to recall it at a later time (Anderson & Green, 2001), to investigate the mechanisms underlying memory suppression for self- and other-related stimuli during two sessions (memory suppression and memory retrieval). At the behavioral level, during the first memory suppression task (i.e. TNT task), individuals responded more quickly to self-related painful relative to non-painful images, suggesting self-protection under threat. Additionally, they responded faster in the think condition than in the memory suppression condition, regardless of whether the stimuli were related to the self or others, indicating the successful use of memory suppression. During the second post-recognition task, individuals spent less time on recognizing the content of self-related painful videos compared to non-painful stimuli. On the other hand, it was more difficult to retrieve self-related memory following the memory suppression strategy, compared to when using a thinking strategy. However, this trend was not consistent with the response to stimuli related to other's pain, suggesting that the memory suppression strategy was more effective in suppressing self-related memories rather than those pertaining to others. In summary, the two task sessions consistently showed that individuals exhibit heightened sensitivity to threats directed towards themselves, and that first-hand pain experiences are comparable to witnessing other's pain (i.e. pain empathy). However, during the memory retrieval period, recalling

self-related memories is more challenging compared to other-related memories after memory suppression, contrary to the typical self-referential advantage (Symons & Johnson, 1997).

Furthermore, mass univariate GLM analysis of the TNT task revealed an interaction effect between condition and strategy in the dmPFC and dACC, indicating greater neural activity in these regions when using the memory suppression strategy compared to the think strategy in response to self or other related images. Consistent with the behavioral findings, these two regions, especially dACC play a more significant role in the demand for memory control (Clairis & Lopez-Persem, 2023). They respond both proactively to early warning signals about unwelcome content and reactively to the emergence of negative thoughts themselves (Crespo-García et al., 2022), indicating higher demand for memory control than memory consolidation (i.e. think condition) in both self and other-related memory. The dACC identified in the interaction effect between condition and strategy during memory suppression, was defined as the seed region. The functional connectivity results suggested that stronger coupling between dACC and insula, as well as bilateral MFG in the self-related compared to other condition for the memory retrieval of the suppressed stimuli during the subsequent recognition task. In line with the behavioral evidence for a disadvantage of memory retrieval for self- rather than other-related memory, stronger dACC-insula and dACC-MFG functional coupling was found in self-relative to other-related memory retrieval. Notably, the insula has been proposed to be crucial in the conscious representation of emotional bodily states, and both dACC and insula are included as part of salience network that modulates cognitive control (Molnar-Szakacs & Uddin, 2022). The MFG is critical for inhibitory control across various cognitive task contexts as well as inhibitory regulation of conscious awareness for unwanted memories (Anderson & Green, 2001; Anderson & Hulbert, 2021; Benoit & Anderson, 2012; Gagnepain et al., 2017). These results
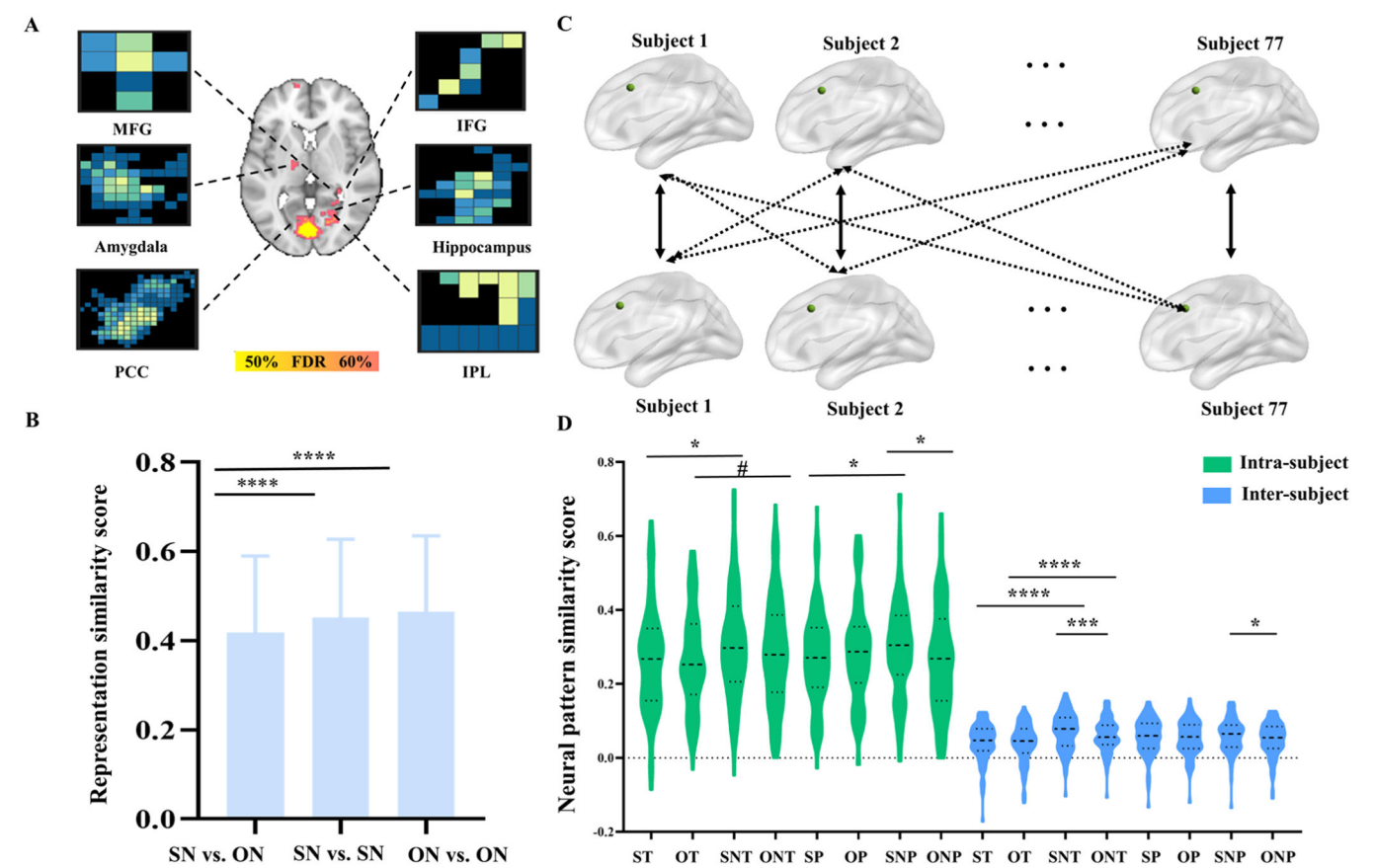
Fig. 6. The searchlight and neural representation similarity results for recognition task and cross-task neural pattern similarity between TNT and recognition tasks. The searchlight (A) and neural representation similarity (B) results between self no-think (SN) and other no-think (ON) during post recognition task. (C)The pipeline for the inter-subject and intra-subject cross-task neural pattern similarity calculation between TNT and post recognition. The solid lines represent intra-subject and the dash lines represent the inter-subject neural pattern similarity. (D)The paired *t*-test between inter-subject pattern similarity scores or intra-subject pattern similarity scores across all conditions. Note, MFG, middle frontal gyrus; IFG, inferior frontal gyrus; PCC, posterior cingulate cortex; IPL, inferior parietal lobule. ST, self-think; OT, other-think; SNT, self no-think; ONT, other no-think; SP, self pain; OP, other pain; SNP, self no-pain; ONP, other no-pain. $^{\#}p = 0.06$, $^{*}p < 0.05$, $^{***}p < 0.005$, $^{****}p < 0.001$ for *t*-tests FDR correction.

Table 3
Cross-task neural pattern similarity between TNT and recognition tasks from intra-subject and inter-subject approach.

Cross-task NPS	Intra-subject		Inter-subject	
Pairs	t	P_FDR	t	P_FDR
ST vs. OT	0.138	0.891	0.515	0.608
SP VS. OP	−0.728	0.469	0.876	0.384
SNT vs. ONT	1.268	0.209	3.36	0.0027
SNP vs. ONP	2.681	0.024	2.875	0.01
ST vs. SNT	−2.672	0.024	−5.878	<0.001
OT vs. ONT	−2.213	0.06	−4.398	<0.001
SP vs. SNP	−2.552	0.029	−0.602	0.549
OP VS. ONP	1.197	0.235	1.244	0.217

NOTE, ST, self-think; OT, other-think; SNT, self no-think; ONT, other no-think; SP, self pain; OP, other pain; SNP, self no-pain; ONP, other no-pain. NPS, neural pattern similarity.

suggested that during the recognition phase, suppressing self-related memories was more effective compared to suppressed other-related memories. The differences observed may be attributed to the initial memory encoding phase, which likely had a different impact on self-related versus other-related memories (Symons & Johnson, 1997).

During the post-recognition task, a notable interaction was observed between condition and pain, suggesting a reduced activity in the post-central/precentral gyrus and precuneus in response to self-related painful stimuli compared to non-painful stimuli. Conversely, the pattern was reversed under other conditions. The precuneus plays a crucial role in various cognitive processes that necessitate the physical point of reference of the self (Lyu et al., 2023). Additionally, the exploratory analysis indicated that greater activity in posterior cingulate cortex during the suppression of memories related to others, as compared to those pertaining to oneself. Generally, the posterior cingulate cortex has been implicated in episodic memory, self-referential cognition, and various other higher-order, abstract thought related to contextual and semantic processing (Foster et al., 2023). Specially, PCC activity is notably augmented during successful episodic memory retrieval (Kahn et al., 2004; Wagner et al., 2005), consistent with behavioral findings that suggest memory retrieval is less effective for other-related compared to self-related memory following memory suppression. Unfortunately, we did not observe any significant differences in the hippocampus, a crucial region involved in memory suppression and retrieval (Gagnepain et al., 2017). This may be attributed to the lack of association with the experimental condition (self vs. other).

Our findings are particularly noteworthy as they deviate from conventional univariate results. Searchlight-based classifiers within the limbic system, specifically the para-/hippocampus and amygdala, as well as the empathy network encompassing STG, IFG, MFG, IPL, PCC, were found to contribute to retrieving self- and other-related memories following memory suppression. This goes beyond the sole involvement of the PCC. The proposed adaptive Bayes process model suggests that ongoing memory consolidation facilitates the integration of new

evidence with pre-existing representations. This dynamic process is further propelled by the emotional and motivational significance of expectations, reflecting the comprehensive limbic control of memory. Specifically, fundamental representations of the limbic system's motivational and memory control functions are integrated with abstract, multi-level cognitive representations and this integration allows for a dynamic and adaptive memory system that responds to changing environments and experiences (Tucker & Luu, 2021). In parallel, empathy is a multidimensional construct that preserves the distinction between oneself and others (Kogler et al., 2020; Zhou et al., 2020) and the neural network associated with empathy comprising regions such as the STG, IFG, MFG, IPL, and PCC (Fan et al., 2011; Kogler et al., 2020). This involvement of limbic and empathy networks, including distributed regions, suggests the complexity of the underlying neural mechanisms and their dynamic interplay in the process of self- and other-related memories following memory suppression.

Our findings highlighted that the dACC serves as a key region, exhibiting shared neural representations across memory suppression and the subsequent memory retrieval. Notably, a greater degree of similarity in representation was observed in suppressing self-related memory, indicating its potential role in the generation of subjective experiences and adaptive responses to both actual and predicted states in the self and others (Bernhardt & Singer, 2012). However, contrary to self-referential effect, which suggests improved integration of bodily and external environment information when processing self-relevant stimuli (Qin et al., 2020), our study revealed challenges in retrieving self-related memories. This finding suggests a more effective emotion regulation strategy when suppressing self-related memories in turn, a process that exhibits consistent similarity with memory suppression and the subsequent memory retrieval.

A number of fMRI studies have highlighted similarities in the anterior insula and ACC during the processing of first-hand experience of pain and third-hand experience of observing others' pain (Decety et al., 2008; Morrison et al., 2004). One study proposed that empathic distress is rooted in the neural representation of first-hand pain (Rütgen et al., 2015) and that it could influence the perception of first-hand experience of pain (Vachon-Presseau et al., 2011). However, searchlight-based analyses suggested that, rather than being involved in the perception of painful stimuli, distributed regions such as the parahippocampus, STG, MFG, MTG and SPL are able to distinguish between self-pain and other-pain during memory retrieval. Concurrently, we also identified some common neural basis encompassing mPFC, PCC, SMA and putamen in the processing of both self-pain and other's pain during memory suppression and the following memory retrieval. Strikingly, these distinct and common neural patterns do not overlap. The former pattern appears to be more closely linked to memory processing (Wagner et al., 2005), while the latter one aligns more with the empathy network (Bernhardt & Singer, 2012; Singer et al., 2004), indicating that while there were shared neural bases underlying experienced pain and pain empathy during memory retrieval, they also have their unique neural bases.

Some studies have proposed that during peri-traumatic or post-traumatic periods, the ascending information transfer of sensory stimuli at lower levels of the neural hierarchy becomes disconnected from self-referential perceptual processes at high levels. This disconnection is believed to result in a lack of integration of sensory input into the updating of self-representation leading to in depersonalization and dissociation (Kube et al., 2020). The current study helped us understand how the self-other distinction modulated the memory suppression and the associated neural networks.

There were several potential limitations in our study. Firstly, the use of personalized and individualized stimuli introduces a degree of variability that could affect the overall results. Secondly, despite our analysis, we did not observe any significant three-way interaction between condition, strategy, and pain at both behavioral and neural levels, particularly for painful stimuli. This may suggest that the induced painful sensations were not sufficiently intense to yield significant differences. Finally, our current findings were based on healthy controls and further clinical studies, particularly with patients suffering from PTSD, should be explored to confirm the reliability of these distinct suppression processes and subsequent memory retrieval in self-related and other-related memories.

Collectively, the current study demonstrated that distinct patterns between self-related and other-related condition during memory suppression and subsequent memory retrieval. This finding may provide evidence that interventions based on memory suppression should clarify whether they target self-experienced suffering or witnessing other's suffering.

Data and code availability

The data and all scripts that support the findings of this study are available on request from the corresponding author.

CRediT authorship contribution statement

Xinwei Song: Conceptualization, Methodology, Investigation, Data curation, Software, Formal analysis, Writing – original draft, Writing – review & editing. **Qi Liu:** Methodology, Investigation, Data curation, Software, Formal analysis, Writing – review & editing. **Xiaodong Zhang:** Investigation, Data curation, Software, Formal analysis. **Can Liu:** Investigation, Data curation, Software, Formal analysis. **Chunmei Lan:** Investigation, Data curation, Software, Formal analysis. **Xiaolu Zhang:** Investigation, Data curation. **Ting Xu:** Data curation, Software. **Ran Zhang:** Investigation, Data curation. **Keith M. Kendrick:** Funding acquisition, Resources, Writing – review & editing. **Benjamin Becker:** Methodology, Writing – review & editing. **Weihua Zhao:** Conceptualization, Methodology, Resources, Supervision, Project administration, Funding acquisition, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgments

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijchp.2024.100509.

References

Agar-Wilson, M., & Jackson, T. (2012). Are emotion regulation skills related to adjustment among people with chronic pain, independent of pain coping? *European Journal of Pain (London, England), 16*, 105–114. https://doi.org/10.1016/j.ejpain.2011.05.011

Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature, 410*, 366–369. https://doi.org/10.1038/35066572

Anderson, M. C., & Hulbert, J. C. (2021). Active forgetting: Adaptation of memory by prefrontal control. *Annual Review of Psychology, 72*, 1–36. https://doi.org/10.1146/annurev-psych-072720-094140

Ashwood, K. L., Gillan, N., Horder, J., Hayward, H., Woodhouse, E., McEwen, F. S., Findon, J., Eklund, H., Spain, D., Wilson, C. E., Cadman, T., Young, S., Stoencheva, V., Murphy, C. M., Robertson, D., Charman, T., Bolton, P., Glaser, K., Asherson, P., Simonoff, E., & Murphy, D. G. (2016). Predicting the diagnosis of autism in adults using the Autism-Spectrum Quotient (AQ) questionnaire. *Psychological Medicine, 46*, 2595–2604. https://doi.org/10.1017/S0033291716001082

Baird, A. D., Scheffer, I. E., & Wilson, S. J. (2011). Mirror neuron system involvement in empathy: A critical look at the evidence. *Social Neuroscience, 6*, 327–335. https://doi.org/10.1080/17470919.2010.547085

Benoit, R. G., & Anderson, M. C. (2012). Opposing mechanisms support the voluntary forgetting of unwanted memories. *Neuron, 76*, 450–460. https://doi.org/10.1016/j.neuron.2012.07.025

Bernhardt, B. C., & Singer, T. (2012). The neural basis of empathy. *Annual Review of Neuroscience, 35*, 1–23. https://doi.org/10.1146/annurev-neuro-062111-150536

Brewin, C. R., Gregory, J. D., Lipton, M., & Burgess, N. (2010). Intrusive images in psychological disorders: Characteristics, neural mechanisms, and treatment implications. *Psychological Review, 117*, 210–232. https://doi.org/10.1037/a0018113

Bringmann, L. F., Lemmens, L. H. J. M., Huibers, M. J. H., Borsboom, D., & Tuerlinckx, F. (2015). Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychological Medicine, 45*, 747–757. https://doi.org/10.1017/S0033291714001809

Clairis, N., & Lopez-Persem, A. (2023). Debates on the dorsomedial prefrontal/dorsal anterior cingulate cortex: Insights for future research. *Brain : A Journal of Neurology, 146*, 4826–4844. https://doi.org/10.1093/brain/awad263

Connelly, M., Keefe, F. J., Affleck, G., Lumley, M. A., Anderson, T., & Waters, S. (2007). Effects of day-to-day affect regulation on the pain experience of patients with rheumatoid arthritis. *Pain, 131*, 162–170. https://doi.org/10.1016/j.pain.2007.01.002

Crespo-García, M., Wang, Y., Jiang, M., Anderson, M. C., & Lei, X. (2022). Anterior cingulate cortex signals the need to control intrusive thoughts during motivated forgetting. *The Journal of Neuroscience : The official journal of the Society for Neuroscience, 42*, 4342–4359. https://doi.org/10.1523/JNEUROSCI.1711-21.2022

De Pascalis, V., & Vecchio, A. (2022). The influence of EEG oscillations, heart rate variability changes, and personality on self-pain and empathy for pain under placebo analgesia. *Scientific Reports, 12*, 6041. https://doi.org/10.1038/s41598-022-10071-9

Decety, J., Michalska, K. J., & Akitsuki, Y. (2008). Who caused the pain? An fMRI investigation of empathy and intentionality in children. *Neuropsychologia, 46*, 2607–2614. https://doi.org/10.1016/j.neuropsychologia.2008.05.026

Decety, J., Yang, C.-Y., & Cheng, Y. (2010). Physicians down-regulate their pain empathy response: An event-related brain potential study. *Neuroimage, 50*, 1676–1682. https://doi.org/10.1016/j.neuroimage.2010.01.025

Depue, B. E., Curran, T., & Banich, M. T. (2007). Prefrontal regions orchestrate suppression of emotional memories via a two-phase process. *Science (New York, N. Y.), 317*, 215–219. https://doi.org/10.1126/science.1139560

Elzinga, B. (2002). Are the neural substrates of memory the final common pathway in posttraumatic stress disorder (PTSD)? *Journal of Affective Disorders, 70*, 1–17. https://doi.org/10.1016/S0165-0327(01)00351-2

Fan, Y., Duncan, N. W., de Greck, M., & Northoff, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience and Biobehavioral Reviews, 35*, 903–911. https://doi.org/10.1016/j.neubiorev.2010.10.009

Fani, N., King, T. Z., Shin, J., Srivastava, A., Brewster, R. C., Jovanovic, T., Bradley, B., & Ressler, K. J. (2016). Structural and functional connectivity in posttraumatic stress disorder: associations with fkbp5: 2015 donald f klein investigator award finalist: connectivity, PTSD, and FKBP5. *Depression and Anxiety, 33*, 300–307. https://doi.org/10.1002/da.22483

Foster, B. L., Koslov, S. R., Aponik-Gremillion, L., Monko, M. E., Hayden, B. Y., & Heilbronner, S. R. (2023). A tripartite view of the posterior cingulate cortex. *Nature reviews. Neuroscience, 24*, 173–189. https://doi.org/10.1038/s41583-022-00661-x

Gagnepain, P., Hulbert, J., & Anderson, M. C. (2017). Parallel regulation of memory and emotion supports the suppression of intrusive memories. *The Journal of neuroscience : The Official Journal of the Society for Neuroscience, 37*, 6423–6441. https://doi.org/10.1523/JNEUROSCI.2732-16.2017

Harnett, N. G., Goodman, A. M., & Knight, D. C. (2020). PTSD-related neuroimaging abnormalities in brain function, structure, and biochemistry. *Experimental Neurology, 330*, Article 113331. https://doi.org/10.1016/j.expneurol.2020.113331

Herbert, C., & Sütterlin, S. (2012). Do Not Respond! doing the Think/No-Think and Go/No-Go tasks concurrently leads to memory impairment of unpleasant items during later recall. *Front. Psychol., 3*. https://doi.org/10.3389/fpsyg.2012.00269

Kahn, I., Davachi, L., & Wagner, A. D. (2004). Functional-Neuroanatomic Correlates of Recollection: Implications for Models of Recognition Memory. *The Journal of neuroscience : The Official Journal of the Society for Neuroscience, 24*, 4172–4180. https://doi.org/10.1523/JNEUROSCI.0624-04.2004

Knowles, K. A., & Olatunji, B. O. (2020). Specificity of trait anxiety in anxiety and depression: Meta-analysis of the State-Trait Anxiety Inventory. *Clinical Psychology Review, 82*, Article 101928. https://doi.org/10.1016/j.cpr.2020.101928

Koechlin, H., Coakley, R., Schechter, N., Werner, C., & Kossowsky, J. (2018). The role of emotion regulation in chronic pain: A systematic literature review. *Journal of Psychosomatic Research, 107*, 38–45. https://doi.org/10.1016/j.jpsychores.2018.02.002

Kogler, L., Müller, V. I., Werminghausen, E., Eickhoff, S. B., & Derntl, B. (2020). Do I feel or do I know? Neuroimaging meta-analyses on the multiple facets of empathy. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior, 129*, 341–355. https://doi.org/10.1016/j.cortex.2020.04.031

Kube, T., Berg, M., Kleim, B., & Herzog, P. (2020). Rethinking post-traumatic stress disorder – A predictive processing perspective. *Neuroscience and Biobehavioral Reviews, 113*, 448–460. https://doi.org/10.1016/j.neubiorev.2020.04.014

Lamm, C., Rütgen, M., & Wagner, I. C. (2019). Imaging empathy and prosocial emotions. *Neuroscience Letters, 693*, 49–53. https://doi.org/10.1016/j.neulet.2017.06.054

Lyu, D., Stieger, J. R., Xin, C., Ma, E., Lusk, Z., Aparicio, M. K., Werbaneth, K., Perry, C. M., Deisseroth, K., Buch, V., & Parvizi, J. (2023). Causal evidence for the

processing of bodily self in the anterior precuneus. *Neuron, 111*, 2502–2512. https://doi.org/10.1016/j.neuron.2023.05.013. e4.

Mary, A., Dayan, J., Leone, G., Postel, C., Fraisse, F., Malle, C., Vallée, T., Klein-Peschanski, C., Viader, F., de la Sayette, V., Peschanski, D., Eustache, F., & Gagnepain, P. (2020). Resilience after trauma: The role of memory suppression. *Science (New York, N.Y.), 367*, eaay8477. https://doi.org/10.1126/science.aay8477

McLaren, D. G., Ries, M. L., Xu, G., & Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *Neuroimage, 61*, 1277–1286. https://doi.org/10.1016/j.neuroimage.2012.03.068

Milad, M. R., Wright, C. I., Orr, S. P., Pitman, R. K., Quirk, G. J., & Rauch, S. L. (2007). Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biological Psychiatry, 62*, 446–454. https://doi.org/10.1016/j.biopsych.2006.10.011

Molnar-Szakacs, I., & Uddin, L. Q. (2022). Anterior insula as a gatekeeper of executive control. *Neuroscience and Biobehavioral Reviews, 139*, Article 104736. https://doi.org/10.1016/j.neubiorev.2022.104736

Montoya, P., Larbig, W., Braun, C., Preissl, H., & Birbaumer, N. (2004). Influence of social support and emotional context on pain processing and magnetic brain responses in fibromyalgia. *Arthritis and Rheumatism, 50*, 4035–4044. https://doi.org/10.1002/art.20660

Morrison, I., Lloyd, D., Di Pellegrino, G., & Roberts, N. (2004). Vicarious responses to pain in anterior cingulate cortex: Is empathy a multisensory issue? *Cognitive, Affective & Behavioral Neuroscience, 4*, 270–278. https://doi.org/10.3758/CABN.4.2.270

Perl, O., Duek, O., Kulkarni, K. R., Gordon, C., Krystal, J. H., Levy, I., Harpaz-Rotem, I., & Schiller, D. (2023). Neural patterns differentiate traumatic from sad autobiographical memories in PTSD. *Nature Neuroscience, 26*, 2226–2236. https://doi.org/10.1038/s41593-023-01483-5

Pitman, R. K., Rasmusson, A. M., Koenen, K. C., Shin, L. M., Orr, S. P., Gilbertson, M. W., Milad, M. R., & Liberzon, I. (2012). Biological studies of post-traumatic stress disorder. *Nature reviews. Neuroscience, 13*, 769–787. https://doi.org/10.1038/nrn3339

Qin, P., Wang, M., & Northoff, G. (2020). Linking bodily, environmental and mental states in the self—A three-level model based on a meta-analysis. *Neuroscience and Biobehavioral Reviews, 115*, 77–95. https://doi.org/10.1016/j.neubiorev.2020.05.004

Rabinak, C. A., Mori, S., Lyons, M., Milad, M. R., & Phan, K. L. (2017). Acquisition of CS-US contingencies during Pavlovian fear conditioning and extinction in social anxiety disorder and posttraumatic stress disorder. *Journal of Affective Disorders, 207*, 76–85. https://doi.org/10.1016/j.jad.2016.09.018

Ressler, K. J., Berretta, S., Bolshakov, V. Y., Rosso, I. M., Meloni, E. G., Rauch, S. L., & Carlezon, W. A. (2022). Post-traumatic stress disorder: Clinical and translational neuroscience from cells to circuits. *Nature reviews Neurology, 18*, 273–288. https://doi.org/10.1038/s41582-022-00635-8

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review, 6*, 135–147. https://doi.org/10.1016/j.edurev.2010.12.001

Rütgen, M., Seidel, E.-M., Silani, G., Riečanský, I., Hummer, A., Windischberger, C., Petrovic, P., & Lamm, C. (2015). Placebo analgesia and its opioidergic regulation suggest that empathy for pain is grounded in self pain. *Proceedings of the National Academy of Sciences, 112*. https://doi.org/10.1073/pnas.1511269112

Schmitz, T. W., Correia, M. M., Ferreira, C. S., Prescot, A. P., & Anderson, M. C. (2017). Hippocampal GABA enables inhibitory control over unwanted thoughts. *Nature Communications, 8*, 1311. https://doi.org/10.1038/s41467-017-00956-z

Seligowski, A. V., Lebois, L. A. M., Hill, S. B., Kahhale, I., Wolff, J. D., Jovanovic, T., Winternitz, S. R., Kaufman, M. L., & Ressler, K. J. (2019). Autonomic responses to fear conditioning among women with PTSD and dissociation. *Depression and Anxiety, 36*, 625–634. https://doi.org/10.1002/da.22903

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science (New York, N.Y.), 303*, 1157–1162. https://doi.org/10.1126/science.1093535

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin, 121*, 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Tucker, D. M., & Luu, P. (2021). Motive control of unconscious inference: The limbic base of adaptive Bayes. *Neuroscience and Biobehavioral Reviews, 128*, 328–345. https://doi.org/10.1016/j.neubiorev.2021.05.029

Vachon-Presseau, E., Martel, M. O., Roy, M., Caron, E., Jackson, P. L., & Rainville, P. (2011). The multilevel organization of vicarious pain responses: Effects of pain cues and empathy traits on spinal nociception and acute pain. *Pain, 152*, 1525. https://doi.org/10.1016/j.pain.2011.02.039

Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences, 9*, 445–453. https://doi.org/10.1016/j.tics.2005.07.001

Woo, C.-W., Schmidt, L., Krishnan, A., Jepma, M., Roy, M., Lindquist, M. A., Atlas, L. Y., & Wager, T. D. (2017). Quantifying cerebral contributions to pain beyond nociception. *Nature Communications, 8*, 14211. https://doi.org/10.1038/ncomms14211

Yang, Z., Quan, H., Peng, Z., Zhong, Y., Tan, Z., & Gong, Q. (2015). Proton magnetic resonance spectroscopy revealed differences in the glutamate + glutamine/creatine ratio of the anterior cingulate cortex between healthy and pediatric post-traumatic stress disorder patients diagnosed after 2008 W enchuan earthquake. *Psychiatry and Clinical Neurosciences, 69*, 782–790. https://doi.org/10.1111/pcn.12332

Zhao, W., Yao, S., Li, Q., Geng, Y., Ma, X., Luo, L., Xu, L., & Kendrick, K. M. (2016). Oxytocin blurs the self-other distinction during trait judgments and reduces medial

prefrontal cortex responses: Oxytocin Blurs Self-Other Distinction. Hum. *Brain Mapp, 37*, 2512–2527. https://doi.org/10.1002/hbm.23190

Zhou, F., Li, J., Zhao, W., Xu, L., Zheng, X., Fu, M., Yao, S., Kendrick, K. M., Wager, T. D., & Becker, B. (2020). Empathic pain evoked by sensory and emotional-communicative cues share common and process-specific neural representations. *Elife, 9*, e56929. https://doi.org/10.7554/eLife.56929

Zimmermann, K., Walz, C., Derckx, R. T., Kendrick, K. M., Weber, B., Dore, B., Ochsner, K. N., Hurlemann, R., & Becker, B. (2017). Emotion regulation deficits in regular marijuana users. *Human Brain Mapping, 38*, 4270–4279. https://doi.org/10.1002/hbm.23671