Title

Dissociable brain mechanisms for long-term memory of disgust- and fear- related associations

Authors

Monika Riegel, Małgorzata Wierzba, Marek Wypych, Maureen Ritchey, Katrzyna Jednoróg, Anna Grabowska, Patrik Vuilleumier, Artur Marchewka

Abstract

Many associative memory traces are charged with emotion that can either impair or enhance subsequent retrieval. Various factors may influence such emotion effects, including the nature of associations and type of emotions.

We show that long-term recognition memory of very close verbal associations was enhanced for emotionally charged material, and these effects differed between two negative emotion categories matched for arousal level. Specifically, memory was better for word pairs related to disgust than word pairs related to fear. Furthermore, these two emotions distinctively modulated neural processes during encoding and its relationship with retrieval, regardless of arousal evoked by word pairs. Amygdala and perirhinal cortex activity were associated with enhanced memory for disgust-related unitizations, whereas better memory for fear-related unitizations engaged parahippocampal cortex and hippocampus. Finally, the magnitude of amygdala activation during encoding was related to the subsequent fidelity of hippocampal activity patterns during retrieval.

These results show that dissociable neural pathways and reinstatement mechanisms are involved in associative memory for different emotion categories and that arousal alone cannot fully explain emotional influences on associative memory.

Introduction

Much of the real-life information that we memorize has some emotional content. Emotion can *enhance* memory for individual items or particular events across various stimuli and paradigms [1,2]. However, a critical feature of human memory is the ability to form associations between different pieces of information during encoding and subsequently retrieve them based on these associations. The effects of emotion on associative memory have been less homogenous, including either *enhancement, impairment*, or *null effects* (reviews: [3–5]). This inconsistency might depend on at least two factors, namely, how close are the associations formed between items and what is the type of emotion.

First, when complex inter-item associations have to be formed, emotion was reported to impair associative memory [6–9]. However, emotion may enhance associative memory when the information is merged and processed as an intra-item association, or even as one single item, reflecting so-called unitization [4,10]. Unitization is a

continuous process, influenced by characteristics of the to-be-merged items and the encoding task. For instance, encoding instructions requiring integrative imagery trigger active unitization attempts more so than non-integrative encoding instructions [11,12].

Second, to date, memory modulation by emotion was explained mostly in relation to arousal and valence, even though behavioral differences were observed for distinct emotion categories, such as disgust and fear [13–16]. Emotion category-specific effects on memory for associations has not been investigated, but better recall of disgust-compared to fear-related stimuli was reported for single words [17] and images [14], regardless of arousal level. These differences might reflect dissociable components implicated in the memory of different emotion categories, perhaps evolutionary-based[18], and possibly supported by distinct brain mechanisms.

Modulation of emotional memories is known to involve the amygdala via its interaction with other medial temporal (MTL) regions [19,20]. Animal research has shown that the AMY shares strong reciprocal connections with anterior hippocampus (HC) and perirhinal cortex (PRC) [21]. Along with PRC receiving projections from the ventral visual 'what' stream (processing of items and objects in the environment), parahippocampal cortex (PHC) receives projections from the dorsal 'where' stream (processing of contextual information, for instance spatial and temporal) [22]. Neuroimaging studies demonstrated that during encoding and retrieval, unitized compared to non-unitized pairs may rely less on hippocampal binding, and more on familiarity-based processes mediated by PRC [23–25]. The arousal effect on memorizing unitized words was shown in negative correlation between the activation of left AMY, frontal areas, and HC [12]. On the other hand, an increase in functional connectivity between right HC and AMY was demonstrated for encoding of negatively valenced faces and identities [26] when compared to neutral ones. However, the role of AMY for memory was only investigated during encoding, tested after a relatively short time, and with reference to arousal and valence, while differences between basic emotion categories were generally disregarded.

Here we determined whether the emotional memory enhancement occurs for verbal unitizations over long-term delays, and if so, whether it is differentially sensitive to distinct emotion categories (disgust and fear) irrespective of their valence and arousal levels. We investigated brain mechanisms hypothesized to occur within medial temporal lobe (MTL) regions recruited during encoding and its relationship with retrieval, for both disgust- and fear- related unitizations. To do so, we combined a recognition memory task with functional magnetic resonance (fMRI) and three analysis approaches: univariate activation analysis, functional connectivity analysis and representational similarity analysis (RSA [27]). During encoding, participants mentally unitized and memorized semantically congruent word pairs (35 per condition) associated with disgust (e.g. swab – stinky), fear (e.g. throat – knife), or emotionally neutral (e.g. circle – axis). After 2-3 weeks, they performed a recognition test during which novel word pairs were also presented as lures (60 per condition). Both phases took place during fMRI scanning. Critically, in a final phase of the experiment, participants rated emotions evoked by the word pairs, so that we could

individually control for the effects of valence and arousal dimensions. Most importantly, this paradigm enabled us to compare the encoding-retrieval similarity of brain activation patterns elicited across two different emotion categories and determine which brain regions promoted this measure of neural reinstatement.

We found a memory enhancement effect due to emotion. Moreover, word pairs related to disgust were remembered better than word pairs related to fear, and this effect could not be explained by differences in arousal. We found distinctive neural pathways preferentially engaged in encoding modulation by disgust or fear. Specifically, the left AMY and left PRC were more activated during successful encoding of disgust-related unitized word pairs, whereas right PHC and HC were more activated for fear-related unitized word pairs, even when the level of arousal was regressed out from the analysis. Moreover, during successful encoding of disgust- compared to fear-related word pairs, we observed an increase in functional connectivity between AMY, PRC and PHC. Finally, for disgusting compared to fearful word pairs, the left AMY encoding activation was found to be critical for the level of trial-by-trial encoding-retrieval similarity of brain activation patterns.

Altogether, the results show that behavioural and neuronal mechanisms of long-term associative memory can be differently modulated by basic emotion categories, over and above of valence and arousal dimensions. For the first time, we revealed that due to AMY and PRC engagement, disgust as opposed to fear has a strong influence on reinstatement and recognition memory of unitized word pairs, which is in line with its distinctive and perhaps evolutionary-driven role in memorizing and avoiding possible sources of contamination. This result is an important complement to debates on theoretical frameworks for emotion and cognition research by showing basic emotion effects beyond affective dimensions.

Results

Emotion-specific effects on memory

Recognition memory was tested 15-19 days (M = 16.86, SD = 1.22) after encoding. Among old word pairs, both emotional categories were correctly recognized more often than neutral word pairs, and word pairs related to disgust were correctly recognized more often than those related to fear (Supplementary Fig. 1). To take into account possible responding strategies of the participants, we also computed the sensitivity index (d') derived from signal detection theory (SDT) and confirmed significant differences in d' between emotion categories ([$F(2,102) = 20.20$, $p < .001$, $\eta^2 = .28$]). Disgust-related word pairs were recognized with greater sensitivity than fear-related ($p < .001$) and neutral pairs ($p < .001$). There was no significant difference in d' between fear-related and neutral pairs (Fig. 1a).

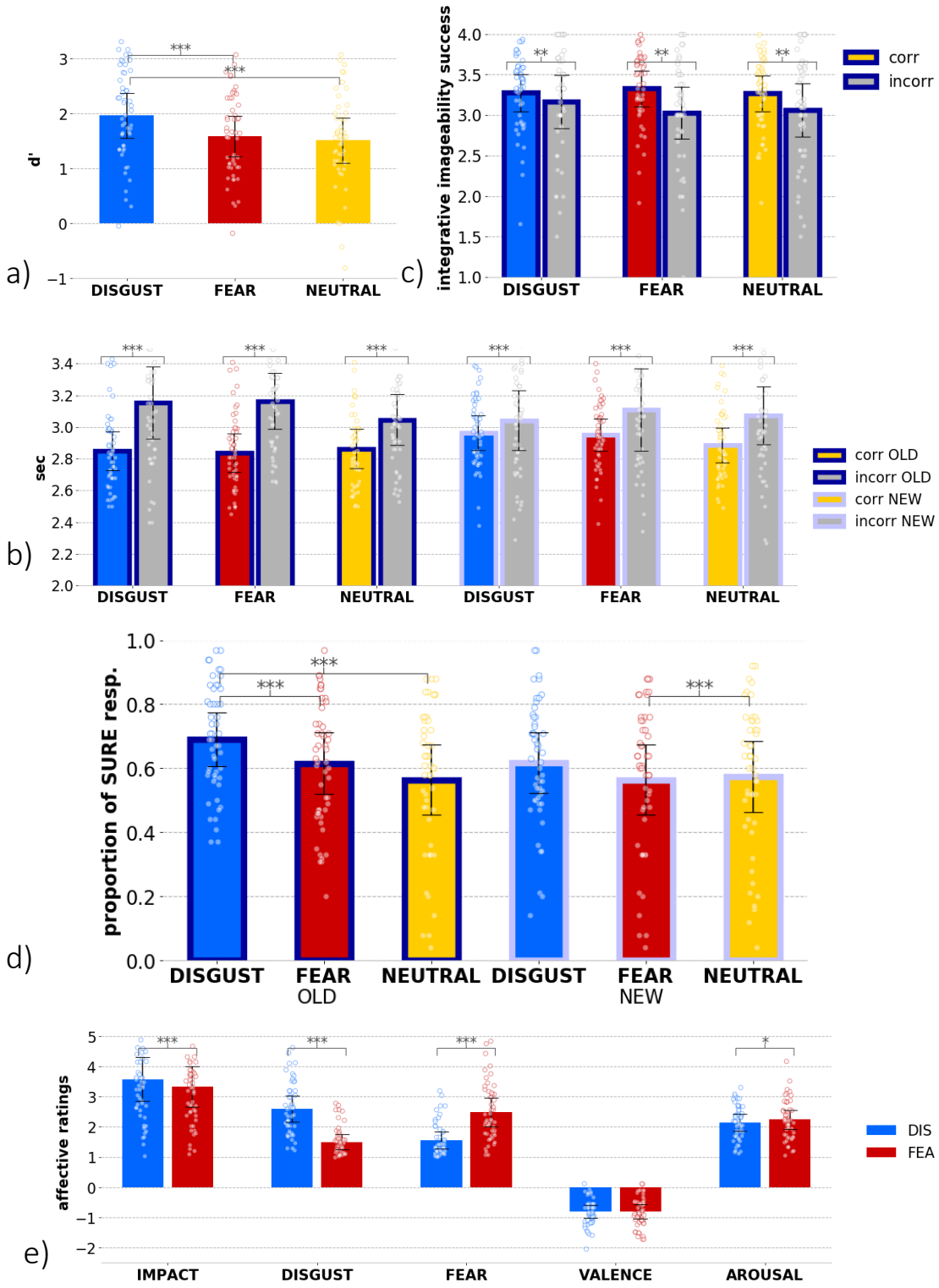

Fig. 1 a) Results of rm ANOVA on the sensitivity index (d') with emotion factor; b) Results of rm ANOVA on reaction times (RT) during recognition with emotion and correctness factors among old and new word pairs (y axis - time in seconds from the onset of a word until the response); c) Results of rm ANOVA on the ratings of integrative imageability success with emotion and correctness factors among old word pairs; d) Results of rm ANOVA on the sureness of correct recognition with emotion factor among old and new word pairs; e) Results of paired t-tests comparing

mean affective ratings on each scale between disgust- and fear- related word pairs; Error bars represent one standard deviation, dots represent individual subjects' scores; corr – correct, incorr – incorrect, DIS – disgust, FEA - fear, ** $p < .005$, *** $p < .001$.

To confirm emotion category-specific effects (i.e. better recognition of disgust- vs. fear-related word pairs), we controlled for the level of arousal associated with each condition by including the difference in arousal ratings (see: Fig. 1e) for all word pairs as a covariate in the analysis described above. The results still showed the main effect of emotion ([$F(1,50) = 19.79$, $p < .001$, $\eta^2 = .28$]) with higher sensitivity index (d') for disgust than fear ($p < .001$). This finding emphasizes that differences in memory performance observed for these two basic emotion categories were not driven by differences in emotional arousal.

Reaction times (RT) during recognition were also analyzed to control for the possible differences in task difficulty among emotion conditions. Only interaction between correctness and novelty (old, new) of word pairs was found [$F(1,45) = 3.97$, $p = .05$, $\eta^2 = .08$], such that correct recognition of old word pairs was significantly faster than new word pairs, but there was no difference in the case of incorrect recognition, as depicted in Fig. 1b. These effects were independent of emotion categories (disgust, fear, neutral) and we can infer that recognition task was not harder for any emotion category.

To determine whether emotion effects on memory could be related to the strength of associations created between words, we then analyzed the ratings of integrative imageability success given by the participants during encoding. The results showed that correctly recognized word pairs were better integrated during encoding compared to those recognized incorrectly as false alarms ([$F(1,45) = 11.086$, $p = .002$, $\eta^2 = .198$]), again regardless of emotion category (disgust, fear, neutral). Thus, the differential memory modulation by emotion categories did not result from the strength of associations created during encoding (Fig. 1c).

Next, we analyzed how the memory and emotion effects were related to subjective confidence, or sureness, rated by the participants for each recognition response. We found that the proportion of sure responses was significantly modulated by emotion category (disgust, fear, neutral) and by novelty (old, new) of the word pairs (interaction effect [$F(2,78) = 4.309$, $p = .017$, $\eta^2 = .099$]), with more frequent sure responses to disgust- than fear-related ($p = .004$) and neutral word pairs ($p = .001$) among old pairs (Fig 4). Among new word pairs, it was lower for fear-related than neutral ($p = .038$) ones (Fig. 1d). In general, the emotion and novelty effects on subjective confidence of recognition showed a similar pattern to their effect on the recognition rate.

Finally, affective ratings collected from the participants were analyzed to determine any potential differences in subjective ratings collected for different emotion categories (disgust, fear, neutral). As expected, we found that all the affective parameters were higher ($p < .001$) for both disgust- and fear-related word pairs than

for neutral pairs. The ratings of disgust were significantly higher for disgusting than fearful ($t = 12.53$, $p < .001$) pairs, whereas ratings of fear were lower for disgust- than fear-related ($t = -12.42$, $p < .001$) pairs, which demonstrates the validity of our experimental manipulation. There was no significant difference in valence ratings between them. Unexpectedly, however, significant differences in the arousal and impact (for definition see: Methods) ratings were found. Specifically, the level of arousal was rated as lower for disgusting then fearful ($t = -2.73$, $p = .009$) word pairs, whereas impact was rated higher for disgust- than fear-related ($t = 3.78$, $p < .001$) pairs (Fig. 1e). Therefore, these differences in subjective emotional experience between different categories of word pairs were taken into account in our subsequent analyses.

Emotion-specific effects on encoding-related brain activity

To elucidate the neural mechanisms of memory modulation by emotion, we first identified brain regions involved in the successful encoding of word pairs related to different emotion categories: disgust (DIS), fear (FEA), neutral (NEU). We focused our initial analyses on the activity within key anatomical regions of interest (ROIs) implicated in memory and emotion interactions, including the bilateral amygdala (AMY), perirhinal cortex (PRC), parahippocampal cortex (PHC), and hippocampus (HC). For each ROI and each emotion condition separately, we extracted activity during successful (later correctly recognized) encoding of word pairs. We found that these regions contributed differently to successful encoding [$F(7,357) = 4.780$, $p = .001$, $\eta^2 = .086$] and that their engagement was dependent on emotion condition (interaction effect [$F(14,714) = 6.463$, $p < .001$, $\eta^2 = .112$]). To unpack these effects, we run a post hoc analysis of simple effects restricted to the two emotion categories (DIS, FEA) which showed that successful encoding produced significantly higher activity for disgust than fear in left AMY ($p < .001$) and left PRC ($p = .01$). On the other hand, successful encoding activity was significantly higher for fear than disgust in right PHC ($p = .007$) and marginally in right HC ($p = .057$).

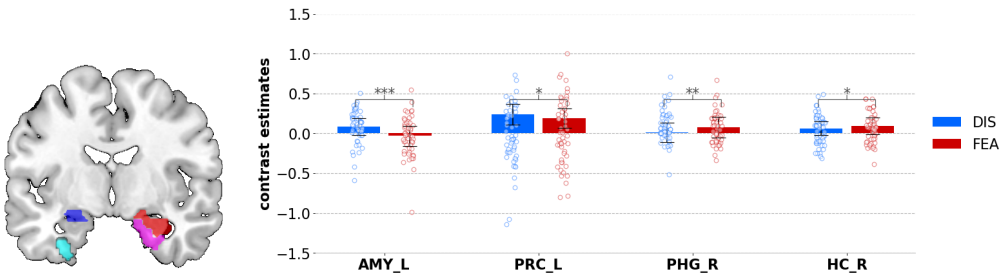

Fig. 2 Results of rm ANOVA on contrast estimates for successful encoding activity from main anatomical ROIs (marked on the brain template: left AMY in blue, left PRC in cyan; right PHC in violet and right HC in red), split according to the emotion factor (disgust- and fear-related word pairs). Error bars represent a standard deviation; *** $p < .001$, ** $p < .005$, * $p < .01$.

To complement the ROI analyses based on a priori hypotheses about their role in emotional memory formation, we also tested for any difference in large-scale cortical

networks associated with successful encoding of different emotion conditions using a whole-brain random-effect analysis. We compared emotional to neutral conditions (EMO > NEU), and both emotion categories with one another (DIS > FEA and FEA > DIS). Critically, mean ratings of arousal were included as covariates to rule out any confound with emotion category effects (for a model before regressing out arousal, see Supplementary Fig. XX). We found that in general, successful encoding of emotional pairs (EMO > NEU) engaged medial prefrontal and parietal regions. More specifically, however, successful encoding of disgust-related unitizations (DIS > FEA) was related to activations in left AMY, left INS, and left dmPFC, as well as other extensive clusters of frontal, temporal and posterior parieto-occipital regions, related for instance to precision of episodic memory [28]. Successful encoding of fear-related unitizations (FEA > DIS) showed activations in bilateral PHC, temporal and parietal regions (see Fig. 3 and Supplementary Table XX).

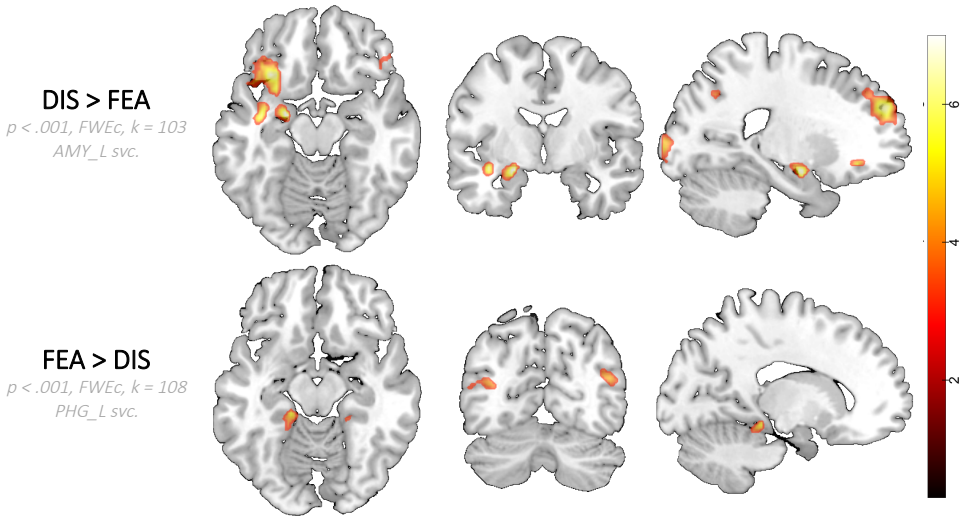

Fig. 3 Differences in brain activation during successful encoding between basic emotion conditions (disgust and fear), after regressing out the effects of arousal; colour bar represents a scale of t-values. FWEc – cluster-level FWE–corrected; k – cluster extent, svc. – small volume correction; colour bar represents a scale of t-values; xyz coordinates in MNI space given above each brain slice.

We also examined how these successful encoding effects were modulated by individual differences in experienced emotion using the affective ratings of word pairs collected in the third experimental session. Remarkably, brain activity during encoding was modulated by the disgust intensity in left dmPFC and bilateral AMY, whereas it was modulated by the fear intensity in bilateral dmPFC, and medial parietal regions, related for instance to threat contextualization [29] and vividness of episodic memory [28] (Supplementary Fig. XX and Suppementary Table XX). No modulation of MTL regions was found for other affective parameters (impact, valence, and arousal).

Finally, the functional connectivity analysis was performed to determine any temporal synchronization of the MTL regions, beyond the modulation of their activation level. To this aim, we used a gPPI analysis to analyze the interaction between functional connectivity of each source and target ROIs (ROI-to-ROI analysis) and experimental

conditions [30]. Thus, our gPPI analysis treated these regions as seeds and determined how they were functionally coupled in pre-defined contrasts of interest. Critically, when directly comparing disgust- and fear-related word pairs (DIS > FEA), we observed an increase in connectivity of the left AMY seed with left PRC [t(51) = 3.05; punc. = .004; pFDR = .026], and right AMY seed with left PRC [t(51) = 3.46; punc. = .001; pFDR = .008] as well as right AMY seed with left PHC [t(51) = 2.59; punc. = .012; pFDR = .043]. There were no results for a negative contrast, meaning that no significant increase in connectivity was found for the selected ROIs in the FEA > DIS contrast.

These results indicate that correct encoding of disgust-related word pairs was related to an increase in functional connectivity between brain regions typically related to affective processing, context memory, and items or unitized associations.

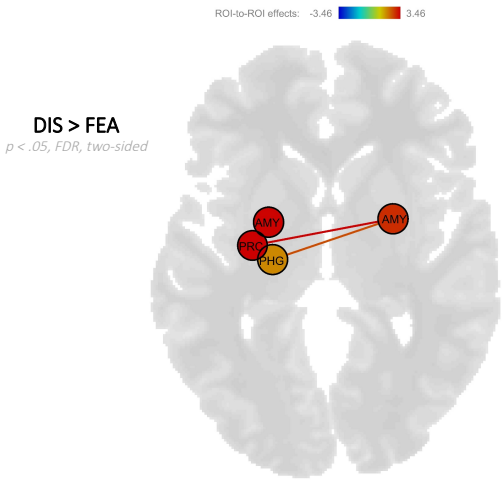

Fig. 4 Increase in functional connectivity during successful encoding of old word pairs related to disgust when compared to fear. DIS – disgust, FEA – fear, pos. – positive, AMY – amygdala, PRC – perirhinal cortex, PHC – parahippocampal cortex; colour bar represents the scale of effect sizes (regression coefficients).

Emotion-specific effects on recognition-related brain activity

Although it was not the main focus of our study, we also investigated brain activity during recognition and found that it was partly consistent with the findings from encoding phase. Again, we first focused our analyses on key anatomical regions of interest (ROIs) (bilateral AMY, PRC, PHC, and HC). We found that these regions contributed differently to correct recognition (main effect of ROI [$F(7,357) = 6.952$, $p < .001$, $\eta^2 = .120$]) and that their engagement was dependent on emotion condition (interaction effect [$F(14,714) = 2.331$, $p = .019$, $\eta^2 = .044$]). A post hoc analysis restricted to emotion categories (DIS, FEA) revealed that activity during correct recognition was marginally higher for disgust than fear in left AMY ($p = .062$) and left PRC ($p = .062$). On the other hand, activity was significantly higher for fear in right PHC ($p = .022$). No difference was observed in hippocampus.

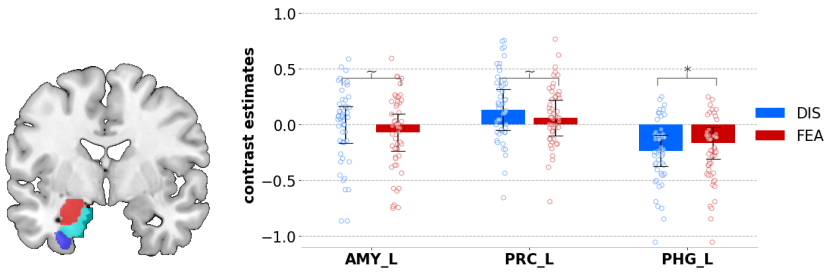

Fig. 5 Results of rm ANOVA on contrast estimates for correct recognition activity from main anatomical ROIs (marked on the brain template: left AMY in blue, left PRC in cyan; right PHC in violet and right HC in red), split according to the emotion factor (disgust- and fear-related old word pairs). Error bars represent one standard deviation; * p < .01.

Similar to encoding, we tested for concomitant differences in large-scale cortical networks using whole-brain random analysis to compare recognition activity between emotional conditions (EMO > NEU, as well as DIS > FEA and FEA > DIS). These comparisons were performed for old word pairs after excluding effects of the same contrasts for new word pairs, in order to isolate activity related to memory from perception of emotion-related words [31]. Again, we also regressed out arousal ratings in this analysis (Fig. 6), for a model without arousal regressed out, see Supplementary Fig. XX and Supplementary Table XX. A direct comparison testing for disgust effect on memory (old DIS > FEA masked exclusively by new DIS > FEA) revealed activations in frontal and parietal regions, typically associated with attentional control and semantic processing. For fear (old FEA > DIS masked exclusively by new FEA > DIS), we observed activations in parietal regions previously implicated in spatial orientation and visual imagery (Supplementary Table XX).

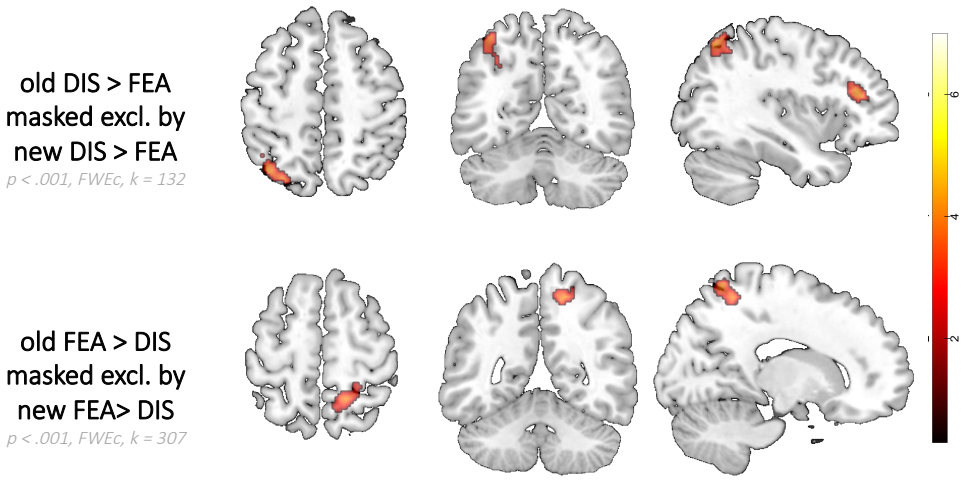

Fig. 6 Differences in brain activation during successful recognition between basic emotion conditions (disgust and fear), after regressing out the effects of arousal. Effects of old word pairs were exclusively masked by the same effects of new word pairs. DIS > FEA corr – disgust > fear correct, FEA > DIS corr – fear > disgust correct; FWEc – cluster-level FWE–corrected; k – cluster extent; colour bar represents a scale of t-values; xyz coordinates in MNI space given above each brain slice.

Additional results concerning the recognition activity as a function of individual ratings of experienced emotion are presented in Supplementary Fig. XX, and ROI-to-ROI functional connectivity in Supplementary Fig. XX.

Encoding-retrieval reinstatement modulated by emotion

All analyses above considered memory and emotion effects for encoding and retrieval phases, separately. However, by collecting data both during encoding and recognition, we were able to directly compare both phases and determine trial-specific reinstatement of brain activity patterns. Further, we could identify how the similarity of neural activation patterns between encoding and recognition [32] was related to the correctness of recognition and emotion category.

To this aim, we performed a split-half correlation analysis based on representational similarity analysis (RSA [27]), using both the whole-brain searchlight and ROI methods. The searchlight analysis was performed with a sphere of interest moving from one voxel to the next one through the whole brain volume [33], whereas ROIs related to emotion and memory were defined anatomically. Using this voxel-wise multivariate approach to compute encoding-retrieval similarity (ERS) [34], we found a significantly higher ERS index for correctly than incorrectly (CORR > INCORR) recognized stimuli in left AMY and right PHC, and for disgust- than fear-related word pairs (DIS > NEU) in right PHC (both $p < .01$, $k = 5$, svc.).

However, a recent study [35] demonstrated that the relationship between encoding and retrieval might be more complex than a faithful replay of past events and involve additional constructive, transformation processes. Therefore, we complemented the previous analysis with reciprocal comparisons between encoding and retrieval brain activity patterns based on [36]. Specifically, we first queried the *retrieval* data for the regions in which a trial-by-trial univariate activity was predicted by (correlated with) encoding activity in the critical ROIs. Second, we queried the *encoding* data for regions whose activation predicted further trial-by-trial reinstatement of activation patterns during recognition, as indicated by the ERS index in masks specific to emotion categories. We compared the results across emotion categories.

In the first analysis, we found that activation of the left AMY during encoding significantly modulated neural activity observed during recognition in the right HC, left HC and right AMY (see Fig. 7 and Supplementary Table XX) regardless of emotion category.

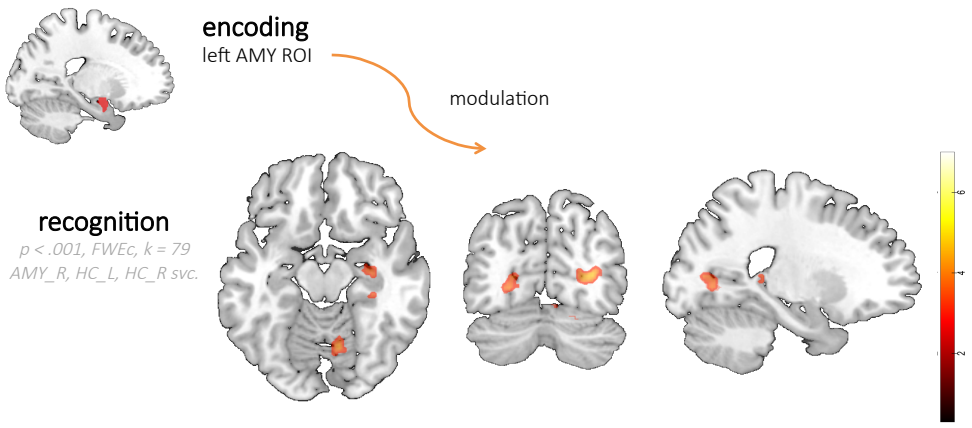

Fig. 7 Brain activation during recognition, modulated by trial-by-trial activation in the left AMY ROI at encoding. HC – hippocampus; FWEc – cluster-level FWE–corrected; k – cluster extent, svc. – small volume correction; colour bar represents a scale of t-values; xyz coordinates in MNI space given above each brain slice.

In the second analysis, we found that during encoding of disgust as compared to fear (DIS > FEA), the activation in left AMY as well as other regions (Fig. 8 and Supplementary Table XX) predicted further ERS in emotion-specific ROIs.

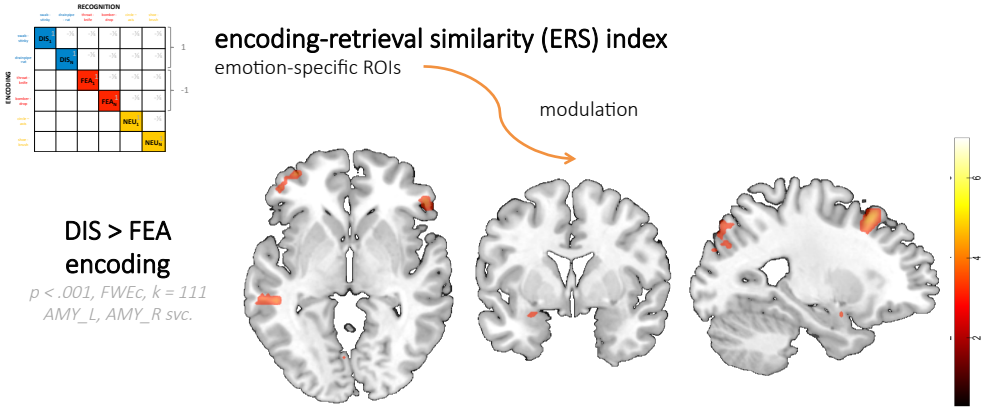

Fig. 8 Brain activation during encoding modulated by trial-by-trial ERS. AMY – amygdala, IFG – inferior frontal gyrus; FWEc – cluster-level FWE–corrected; k – cluster extent, svc. – small volume correction; colour bar represents a scale of t-values; xyz coordinates in MNI space given above each brain slice.

In total, these univariate and multivariate measures show that AMY played an important role in the encoding-retrieval reinstatement of brain activity. Specifically, there was a positive correlation between the activity of left AMY during encoding and regions engaged in mnemonic and attentional processes during recognition. We also found a correlation between ERS index computed for each trial and left AMY activation on the corresponding encoding event for disgust compared to fear. This suggests that the magnitude of initial AMY activation is related to the fidelity of subsequent reinstatement of cortical activity patterns during recognition, particularly of disgust-associated stimuli. Consistent with current theories of episodic memory,

these findings demonstrate a critical link between AMY activation during an emotional experience and subsequent cortical reinstatement.

Discussion

How emotion influences associative memory is a fundamental question for affective and cognitive neuroscience. While previous studies focused mainly on showing how mnemonic processes are modulated by the arousal evoked by to-be-remembered information, here for the first time we uncover brain mechanisms of differential encoding and reinstatement modulation by discrete emotion categories, independent of valence and arousal. Moreover, unlike previous studies using visual material and testing memory shortly after encoding, we used emotionally charged verbal material and tested memory after a long-term delay. Our results revealed that after 2-3 weeks, unitized word-pairs evoking disgust were remembered better than word pairs evoking fear, an effect related to different brain activity patterns. Specifically, AMY and PRC supported better memory for disgust-related word pairs, whereas PHC and HC supported better memory of fear-related word pairs. We also found that AMY was temporally synchronized with other MTL regions during successful encoding of disgust more than fear. Finally, AMY activity during encoding was shown to modulate brain activation in HC during recognition, as well as to predict the subsequent fidelity of brain activity pattern reinstatement in cortical areas during recognition. These results unveil how different emotion categories can modulate memory formation in different ways, despite similar negative valence and arousal level, through the engagement of distinct neural pathways.

Differences in memory modulation by disgust and fear

Our first major finding was that disgust and fear-related information produced behaviourally distinct effects on long-term memory of unitizations. Even though these two emotions are negative, arousing, and associated with motivational avoidance [37,38], they are known to induce distinct physiological responses [39,40] and different effects on cognitive processes including item memory. Better recall of disgust- compared to fear-related stimuli was previously demonstrated for single words [17] and images [14]. Here we show better memory for unitized word pairs evoking disgust compared to fear, an effect unexplained by differences in arousal levels. Based on previous literature, we assume that this effect is unlikely to reflect a differential influence of disgust and fear on attentional processes [15,16,40,41,13]. Also, we did not find any differences in cognitive effort (measured as RT) during recognition between these two basic emotion categories. Given that we investigated the influence of disgust and fear on memory of verbal associations, we additionally controlled for other possibly related factors: imageability success during encoding, and semantic coherence of word pairs (Supplementary Fig. XX). No differences between disgust- and fear- related word pairs were found for these factors. A recent study attributed the higher level of free recall for disgust-related than neutral and fear-related words to deeper elaborative encoding [42]. Since the nature of our task (creating common mental representations of word pairs) forced elaborative processing for all the experimental conditions, it should eliminate also such differences. Altogether, disgust appears to

have a special salience in memory for verbal associations regardless of possible confounding factors.

Extending our behavioural effects, analyzed neuroimaging data revealed that disgust and fear engaged distinct neural pathways of mnemonic processes. First and foremost, we observed a distinctive involvement of AMY in encoding and reinstatement of disgust-related verbal unitizations. In animals, AMY activity was shown to initiate autonomic responses to salient stimuli (LeDoux et al., 1998), consistent with human neuroimaging studies [43]. Despite a traditional focus on the processing of fear and threat in AMY [44], recent studies demonstrate that AMY may activate to any stimuli characterized by a high level of personal impact, independent of intrinsic emotional properties [45] , relevance detection [46] and evaluative processing of goals [47]. Here we found increased AMY activation for the memory of word pairs related to disgust, which were indeed characterized by higher subjective ratings of impact, but not arousal. Also, we show that individual ratings of disgust, but not ratings of arousal or impact, modulated activity in AMY.

Our results are in line with previous studies showing that AMY activity at encoding increases the likelihood of remembering emotional, but not neutral items [48,49], even without a conscious emotional experience [50]. However, the current study is the first to show that AMY differentially contributes to memory for two distinct negative emotions, even after controlling for arousal level [51]. Most importantly, we also found that left AMY activation during encoding critically influenced the fidelity of brain activity patterns reinstatement during recognition, as shown by our univariate and multivariate analyses. Previous studies employing multi-voxel analyses have only investigated the relationship between encoding and retrieval brain activation pattern [32], either for emotional items [52] or neutral associations [53], but to our knowledge emotion modulation of associative memory has never been tested this way.

Another brain structure playing a crucial role in memory of disgust was PRC. Other findings support complex multi-modal item representations in this region [54,55], such as integration of novel odors with visual category information [56], or integration of visual with conceptual object features [57]. PRC was also shown to reflect semantic similarity between words [58], yet here we found no difference in the semantic coherence of word pairs between disgust and fear. This region has direct connections and a privileged access to signals from AMY [21,59], which may be important for coding item salience. Greater PRC activity during emotional compared to neutral memory encoding was observed in previous research [60], but only if it was accompanied by better recollection compared to neutral items, similar to our study. Also, the engagement of PRC might result from the multisensory nature of disgust processing, although this factor was not manipulated nor quantified in our experiment.

Differences possibly related to unitization process

In contrast, we found that left PHC and right HC played a crucial role in successful encoding of word pairs evoking fear, again regardless of arousal levels. The ROI analyses showed that the right PHC and right HC activated more during encoding of

fear- than disgust- related word pairs, whereas left PHC activated during recognition of fear- more than disgust-related word pairs. Given that PHC is typically related to context processing [4,5], these results may suggest that fear-related word pairs might evoke pre-existing associations related to the meaning of words within pairs, but be less efficiently associated together by unitization. However, our measure of integrative imageability success during encoding did not reveal any difference between emotions. In any case, the current results clearly converge to indicate that different unitization and memory mechanisms were engaged when modulated by disgust- and fear- related word pairs.

Previous studies have indicated a double dissociation between PRC and PHC for encoding of object vs. scene context [61,62], paralleling the dissociation found here between disgust and fear. One possibility is that fear-related unitized mental representations were processed in a scene-like fashion of spatial exploration, in line with an effect of threat signals on the monitoring of the environment. Another possibility is that emotions may generally enhance context processing, particularly in the case of fear. PHC could therefore mediate this strong connection between contextual processing and emotional information, facilitating emotion understanding but also subsequent episodic memory [63].

The involvement of HC in emotional memory has been also intensely investigated. One important insight to this issue was provided by a study [64] where superior memory for emotional material came at the cost of reduced memory for contextual details. This was reflected by an increase in functional connectivity between the left HC and right AMY for negative compared to neutral photographs. Another study[11] reported that associative memory for negative information was consistently impaired, but accompanied by increased activity in AMY, as well as in the HC. The authors suggested that when sufficiently arousing information precludes unitization-based or within-domain associative encoding supported by MTL-cortex regions, an alternative, relational HC-dependent encoding strategy may be engaged. Our results showed that AMY activity during encoding modulated the activity of right HC during recognition, shedding a new light on this relationship also across the two key stages of mnemonic processes, i.e., encoding and recognition.

It is important to note that our results conflict with previous reports of a disruptive role of AMY during encoding of arousing associations [11,12,65,66]. However, they align with recent finding that the engagement of both AMY and PRC can support the recollection advantage for emotional items [67], which reflect how unitized word pairs were memorized and recognized. Our results showed an increase in functional connectivity between AMY and PRC for disgust compared to fear. Despite no differences in the strength of associations created during encoding at the behavioural level, the interplay between AMY and PRC may be promoted by a more successful unitization [11]. Also, our data converge with the *emotional binding* account [5], according to which AMY mediates the recollection of item-emotion bindings that are forgotten more slowly than item-context bindings supported by the HC [5]. This account is in line with our results whereby more AMY activation related to encoding and retrieval of disgust (as compared to fear-related and neutral word pairs) induced

better memory. A question remains, however, if item-emotion associations can be unitized, as instructed in our experiment. If the current model is correct, the unitized item-emotion associations should be supported by AMY and PRC, as well as later recognized based solely on familiarity. However, we tested only recognition memory and not associative memory, so this prediction will require further studies.

Possible explanations of distinct brain mechanisms

Although disgust and fear modulated memory of unitized associations through partly distinct neural mechanisms, we do not imply that there exist separate neural substrates for basic emotion categories, which has long been debated [68–70]. Rather, our study was designed to determine whether, and how, encoding and reinstatement of verbal associations may be differently modulated by disgust and fear regardless of overall negative valence and arousal. Above all, we do not imply that AMY is not involved in fear processing [71–73] or that PHC cannot be involved in associative memory of different emotion categories [63]. However, we surmise that fear and disgust might recruit partly distinct emotion appraisal components that activate different kind of associations in memory and thus recruit partly distinct neural systems holding such associations.

Notably, distress associated with exposure to disgusting cues has been shown to be more cognitively penetrable than distress associated with exposure to fear cues due to more effective reappraisal and counterconditioning [74,75]. Likewise, it is known that disgust is a relatively complex, multifaceted emotion originating in distaste but extending to social domains and moral disgust, resulting in multisensory features specific to this emotion [76,77]. In contrast, fear may take different forms characterized as reactive or cognitive, and in case of slow escape decisions engage brain areas related to memory or threat anticipation, such as posterior cingulate cortex, HC, and ventromedial prefrontal cortex [78]. Moreover, both disgust and fear imply avoidance, but for different reasons [37]. More research is also needed to clarify the role of semantic memory in the effects observed in our experimental design, and better dissociate emotional modulation from the semantic elaboration [79–81], although global semantic processing effects should be cancelled with our direct contrasts between emotion categories.

We should however acknowledge some limitations to these conclusions. It is possible that this dissociation is specific to verbal material and that the latter is more likely to evoke disgust than fear. Thus, a different pattern might be observed for other modalities, for instance with visual or auditory material [82–84]. AMY activation can influence both visual and semantic processing [85], but how these influences differ for visual and verbal unitizations has not been explored. It is also possible that our effects were specific for this kind of memory task, with unitization instruction during encoding and recognition after a long-term delay, which may require complex cognitive operations [24,25,86,87]. Moreover, our experimental design does not allow us to draw any conclusions about emotional modulation of associative source memory [67], so we cannot determine whether our participants recollected information based

on recollection or familiarity and how continuous is the unitization process across these conditions [5,88].

Also, our experimental design was limited to comparing only disgust, fear and neutral content, preventing us to compare our findings with other basic emotion categories or positive information [89]. Finally, it must be underscored that studies on memory and emotion exhibit a wide variety of experimental procedures, with most manipulating only valence and arousal, such that it is difficult to directly compare them [90]. Last but not least, there might be individual and sex differences in emotion processing that influence emotional memory task performance [91,92], but we did not assess these factors.

Conclusions

In sum, our results reveal remarkable differences between memory modulation by disgust and fear, at both the behavioural and neuronal levels. We find not only that long-term recognition memory for unitized word pairs is modulated by emotion, but this enhancement appears stronger for disgust- than fear-evoking material, an effect not caused by higher arousal. Successful memory of disgust word pairs was related to greater AMY and PRC activations, whereas PHC and HC supported better memory for fear. Moreover, amygdala activation during encoding was critical for later encoding-retrieval similarity of activation patterns in several ROIs within memory pathways. These differences might result from distinct associative processes at play and particular motivational and physiological responses [93] evoked by disgust and fear. For instance, fear-motivated avoidance may protect a person from direct danger, disgust-motivated avoidance might more often be linked to anticipatory sensation or imagery (see [37]).

To conclude, we demonstrate that there might be distinct neural pathways engaged in memory modulation for different emotion categories, over and above of (negative) valence or (high) arousal dimensions. This result provides important insights for theoretical frameworks of both emotion and memory processes, showing that the two approaches (emotion specific and dimensional) should be treated as complementary.

References

1. LaBar, K. S. & Cabeza, R. Cognitive neuroscience of emotional memory. Nat. Rev. Neurosci. 7, 54–64 (2006).
2. Hamann, S. Cognitive and neural mechanisms of emotional memory. 5, 394–400 (2001).
3. Chiu, Y. C., Dolcos, F., Gonsalves, B. D. & Cohen, N. J. On opposing effects of emotion on contextual or relational memory. Front. Psychol. 4, 2–5 (2013).
4. Murray, B. D. & Kensinger, E. A. A review of the neural and behavioral consequences for unitizing emotional and neutral information. Front. Behav. Neurosci. 7, 42 (2013).
5. Yonelinas, A. P. & Ritchey, M. The slow forgetting of emotional episodic memories: an emotional binding account. Trends Cogn. Sci. 19, 259–267 (2015).
6. Murray, B. D. & Kensinger, E. A. The effects of emotion and encoding strategy on associative memory. Mem. Cognit. 40, 1056–69 (2012).
7. Pierce, B. H. & Kensinger, E. A. Effects of emotion on associative recognition valence and retention interval matter. 11, 139–144 (2011).
8. Rimmele, U. et al. Emotion enhances the subjective feeling of remembering, despite lower accuracy for contextual details. Emotion 11, 553–562 (2011).
9. Bergmann, H. C., Rijpkema, M., Fernández, G. & Kessels, R. P. C. The effects of valence and arousal on associative working memory and long-term memory. PLoS One 7, e52616 (2012).
10. Parks, C. M. & Yonelinas, A. P. The importance of unitization for familiarity-based learning. J. Exp. Psychol. Learn. Mem. Cogn. 41, 881–903 (2015).
11. Madan, C. R., Fujiwara, E., Caplan, J. B. & Sommer, T. Emotional arousal impairs association-memory: Roles of amygdala and hippocampus. Neuroimage 156, 14–28 (2017).
12. Murray, B. D. & Kensinger, E. A. The route to an integrative associative memory is influenced by emotion. PLoS One 9, e82372 (2014).
13. Chapman, H. a, Johannes, K., Poppenk, J. L., Moscovitch, M. & Anderson, A. K. Evidence for the differential salience of disgust and fear in episodic memory. J. Exp. Psychol. Gen. 142, 1100–12 (2012).
14. Croucher, C. J., Calder, A. J., Ramponi, C., Barnard, P. J. & Murphy, F. C. Disgust enhances the recollection of negative emotional images. PLoS One 6, (2011).
15. van Hooff, J. C., Devue, C., Vieweg, P. E. & Theeuwes, J. Disgust- and not fear-evoking images hold our attention. Acta Psychol. (Amst). 143, 1–6 (2013).
16. van Hooff, J. C., van Buuringen, M., El M'rabet, I., de Gier, M. & van Zalingen, L. Disgust-specific modulation of early attention processes. Acta Psychol. (Amst). 152, 149–57 (2014).
17. Charash, M. & McKay, D. Attention bias for disgust. J. Anxiety Disord. 16, 529–541 (2002).
18. Susskind, J. M. et al. Expressing fear enhances sensory acquisition. Nat. Neurosci. 11, 843–850 (2008).
19. Kensinger, E. A. & Schacter, D. L. Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. J. Neurosci. 26, 2564–2570 (2006).
20. Cahill, L. & McGaugh, J. L. Mechanisms of emotional arousal and lasting

declarative memory. Trends Neurosci. (1998). doi:10.1016/S0166-2236(97)01214-9
21. Pitkänen, A., Pikkarainen, M., Nurminen, N. & Ylinen, A. Reciprocal connections between the amygdala and the hippocampal formation, perirhinal cortex, and postrhinal cortex in rat. A review. Ann. N. Y. Acad. Sci. 911, 369–91 (2000).
22. Ranganath, C. A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. Hippocampus (2010). doi:10.1002/hipo.20852
23. Haskins, A. L., Yonelinas, A. P., Quamme, J. R. & Ranganath, C. Perirhinal Cortex Supports Encoding and Familiarity-Based Recognition of Novel Associations. Neuron 59, 554–560 (2008).
24. Memel, M. & Ryan, L. Visual integration enhances associative memory equally for young and older adults without reducing hippocampal encoding activation. Neuropsychologia 100, 195–206 (2017).
25. Memel, M. & Ryan, L. Visual integration of objects and scenes increases recollection-based responding despite differential MTL recruitment in young and older adults. Hippocampus (2018). doi:10.1002/hipo.23011
26. Berkers, R. M. W. J., Klumpers, F. & Fernández, G. Medial prefrontal-hippocampal connectivity during emotional memory encoding predicts individual differences in the loss of associative memory specificity. Neurobiol. Learn. Mem. 134, 44–54 (2015).
27. Kriegeskorte, N. Representational similarity analysis – connecting the branches of systems neuroscience. Front. Syst. Neurosci. (2008). doi:10.3389/neuro.06.004.2008
28. Richter, F. R., Cooper, R. A., Bays, P. M. & Simons, J. S. Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. Elife 5, (2016).
29. Bukalo, O. & Holmes, A. Everything in Its Right Place: A Prefrontal-Midbrain Circuit for Contextual Fear Discrimination. Neuron 97, 732–733 (2018).
30. McLaren, D. G., Ries, M. L., Xu, G. & Johnson, S. C. A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. Neuroimage (2012). doi:10.1016/j.neuroimage.2012.03.068
31. Voss, J. L., Bridge, D. J., Cohen, N. J. & Walker, J. A. A Closer Look at the Hippocampus and Memory. Trends Cogn. Sci. 21, 577–588 (2017).
32. Xue, G. The Neural Representations Underlying Human Episodic Memory. Trends in Cognitive Sciences 22, (2018).
33. Kriegeskorte, N., Goebel, R. & Bandettini, P. Information-based functional brain mapping. Proc. Natl. Acad. Sci. U. S. A. 103, 3863–3868 (2006).
34. Staresina, B. P., Henson, R. N. A., Kriegeskorte, N. & Alink, A. Episodic Reinstatement in the Medial Temporal Lobe. J. Neurosci. 32, 18150–18156 (2012).
35. Xiao, X. et al. Transformed Neural Pattern Reinstatement during Episodic Memory Retrieval. J. Neurosci. 37, 2986–2998 (2017).
36. Danker, J. F., Tompary, A. & Davachi, L. Trial-by-Trial Hippocampal Encoding Activation Predicts the Fidelity of Cortical Reinstatement during Subsequent Retrieval. Cereb. Cortex 27, 3515–3524 (2017).

37. Woody, S. R. Intersection of Disgust and Fear: Normative and Pathological Views. Clin. Psychol. Sci. Pract. 7, 291–311 (2000).
38. Krusemark, E. a & Li, W. Do all threats work the same way? Divergent effects of fear and disgust on sensory perception and attention. J. Neurosci. 31, 3429–3434 (2011).
39. Critchley, H. D. et al. Activity in the human brain predicting differential heart rate responses to emotional facial expressions. Neuroimage 24, 751–762 (2005).
40. Cisler, J. M., Olatunji, B. O. & Lohr, J. M. Disgust, fear, and the anxiety disorders: a critical review. Clin. Psychol. Rev. 29, 34–46 (2009).
41. Chapman, H. A. Enhanced recall of disgusting relative to frightening photographs is not due to organisation. Cogn. Emot. 32, 1220–1230 (2018).
42. Ferré, P., Haro, J. & Hinojosa, J. A. Be aware of the rifle but do not forget the stench: differential effects of fear and disgust on lexical processing and memory. Cogn. Emot. 32, 796–811 (2018).
43. Critchley, H. D., Mathias, C. J. & Dolan, R. J. Fear conditioning in humans: The influence of awareness and autonomic arousal on functional neuroanatomy. Neuron 33, 653–663 (2002).
44. Phelps, E. A. & LeDoux, J. E. Contributions of the amygdala to emotion processing: From animal models to human behavior. Neuron 48, 175–187 (2005).
45. Ewbank, M. P., Barnard, P. J., Croucher, C. J., Ramponi, C. & Calder, A. J. The amygdala response to images with impact. Soc. Cogn. Affect. Neurosci. 4, 127–133 (2009).
46. Sander, D., Grafman, J. & Zalla, T. Sander et al. (2003). The human amygdala - an evolved system for relevance detection.pdf. Rewievs in the Neurosciences 14, 303–316 (2003).
47. Cunningham, W. A., Van Bavel, J. J. & Johnsen, I. R. Affective flexibility: Evaluative processing goals shape amygdala activity: Research article. Psychol. Sci. 19, 152–160 (2008).
48. Dolcos, F., LaBar, K. S. & Cabeza, R. Remembering one year later: role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. Proc. Natl. Acad. Sci. U. S. A. 102, 2626–2631 (2005).
49. Kensinger, E. A. & Schacter, D. L. Processing emotional pictures and words : Effects of valence and arousal avocado. 6, 110–126 (2006).
50. Inman, C. S. et al. Direct electrical stimulation of the amygdala enhances declarative memory in humans. Proc. Natl. Acad. Sci. 115, 98–103 (2018).
51. de Voogd, L. D., Fern??ndez, G. & Hermans, E. J. Disentangling the roles of arousal and amygdala activation in emotional declarative memory. Soc. Cogn. Affect. Neurosci. 11, 1471–1480 (2016).
52. Ritchey, M., Wing, E. A., Labar, K. S. & Cabeza, R. Neural Similarity Between Encoding and Retrieval is Related to Memory Via Hippocampal Interactions. Cereb. Cortex 23, 2818–2828 (2013).
53. Danker, J. F., Tompary, A. & Davachi, L. Trial-by-Trial Hippocampal Encoding Activation Predicts the Fidelity of Cortical Reinstatement During Subsequent Retrieval. Cereb. Cortex bhw146 (2016). doi:10.1093/cercor/bhw146
54. Ranganath, C. & Ritchey, M. Two cortical systems for memory-guided

behaviour. Nat. Rev. Neurosci. 13, 713–726 (2012).
55. Taylor, K. I., Moss, H. E., Stamatakis, E. A. & Tyler, L. K. Binding crossmodal object features in perirhinal cortex. Proc. Natl. Acad. Sci. 103, 8239–8244 (2006).
56. Qu, L. P., Kahnt, T., Cole, S. M. & Gottfried, J. A. De Novo Emergence of Odor Category Representations in the Human Brain. J. Neurosci. 36, 468–478 (2016).
57. Martin, C. B., Douglas, D., Newsome, R. N., Man, L. L. Y. & Barense, M. D. Integrative and distinctive coding of visual and conceptual object features in the ventral visual stream. Elife 7, 3–5 (2018).
58. Bruffaerts, R. et al. Similarity of fMRI Activity Patterns in Left Perirhinal Cortex Reflects Semantic Similarity between Words. 33, 18597–18607 (2013).
59. Tomás Pereira, I., Agster, K. L. & Burwell, R. D. Subcortical connections of the perirhinal, postrhinal, and entorhinal cortices of the rat. I. afferents. Hippocampus 26, 1189–1212 (2016).
60. Dolcos, F., LaBar, K. S. & Cabeza, R. Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. Neuron 42, 855–863 (2004).
61. Li, M., Lu, S. & Zhong, N. The parahippocampal cortex mediates contextual associative memory: Evidence from an fMRI study. Biomed Res. Int. 2016, (2016).
62. Staresina, B. P., Duncan, K. D. & Davachi, L. Perirhinal and Parahippocampal Cortices Differentially Contribute to Later Recollection of Object- and Scene-Related Event Details. J. Neurosci. 31, 8739–8747 (2011).
63. Aminoff, E. M., Kveraga, K. & Bar, M. The role of the parahippocampal cortex in cognition. Trends Cogn. Sci. 17, 379–390 (2013).
64. Dandolo, L. C. & Schwabe, L. Time-dependent memory transformation along the hippocampal anterior-posterior axis. Nat. Commun. 9, 1–11 (2018).
65. Bisby, J. A., Horner, A. J., Horlyck, L. D. & Burgess, N. Opposing effects of negative emotion on amygdalar and hippocampal memory for items and associations. Soc. Cogn. Affect. Neurosci. 11, 981–990 (2016).
66. Okada, G. et al. The effect of negative and positive emotionality on associative memory: An fMRI study. PLoS One 6, (2011).
67. Ritchey, M., Wang, S.-F., Yonelinas, A. P. & Ranganath, C. Dissociable medial temporal pathways for encoding emotional item and context information. bioRxiv (2018). doi:10.1101/248294
68. Saarimaki, H. et al. Discrete Neural Signatures of Basic Emotions. Cereb. Cortex 1–11 (2015). doi:10.1093/cercor/bhv086
69. Saarimäki, H. et al. Distributed affective space represents multiple emotion categories across the human brain. Soc. Cogn. Affect. Neurosci. 13, 471–482 (2018).
70. Clark-polner, E., Johnson, T. D. & Barrett, L. F. Multivoxel Pattern Analysis Does Not Provide Evidence to Support the Existence of Basic Emotions. 1–5 (2016). doi:10.1093/cercor/bhw028
71. Hermans, E. J. et al. Persistence of Amygdala-Hippocampal Connectivity and Multi-Voxel Correlation Structures During Awake Rest After Fear Learning Predicts Long-Term Expression of Fear. Cereb. Cortex 27, 3028–3041 (2017).
72. Méndez-Bértolo, C. et al. A fast pathway for fear in human amygdala. Nat.

Neurosci. 19, 1041–1049 (2016).
73. Vuilleumier, P., Armony, J. L., Driver, J. & Dolan, R. J. Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. Neuron 30, 829–841 (2001).
74. Olatunji, B. O., Berg, H. E. & Zhao, Z. Emotion regulation of fear and disgust: differential effects of reappraisal and suppression. Cogn. Emot. 31, 403–410 (2017).
75. Engelhard, I. M., Leer, A., Lange, E. & Olatunji, B. O. Shaking that icky feeling: Effects of extinction and counterconditioning on disgust-related evaluative learning. Behav. Ther. 45, 708–719 (2014).
76. Sharvit, G., Vuilleumier, P., Delplanque, S. & Corradi-Dell'Acqua, C. Cross-modal and modality-specific expectancy effects between pain and disgust. Sci. Rep. (2015). doi:10.1038/srep17487
77. Corradi-Dell'Acqua, C., Tusche, A., Vuilleumier, P. & Singer, T. Cross-modal representations of first-hand and vicarious pain, disgust and fairness in insular and cingulate cortex. Nat. Commun. 7, 10904 (2016).
78. Qi, S. et al. How cognitive and reactive fear circuits optimize escape decisions in humans. Proc. Natl. Acad. Sci. 201712314 (2018). doi:10.1073/pnas.1712314115
79. Kaneda, T., Shigemune, Y. & Tsukiura, T. Lateral and medial prefrontal contributions to emotion generation by semantic elaboration during episodic encoding. Cogn. Affect. Behav. Neurosci. 17, 143–157 (2017).
80. Wright, P., Randall, B., Clarke, A. & Tyler, L. K. The perirhinal cortex and conceptual processing: Effects of feature-based statistics following damage to the anterior temporal lobes. Neuropsychologia 76, 192–207 (2015).
81. Ritchey, M., LaBar, K. S. & Cabeza, R. Level of processing modulates the neural correlates of emotional memory formation. J. Cogn. Neurosci. 23, 757–771 (2011).
82. Domínguez-borràs, J. et al. Fear spreading across senses: visual emotional events alter cortical responses to touch, audition, and vision. Cereb. Cortex 68–82 (2017). doi:10.1093/tropej/fmw080
83. Marchewka, A. et al. Arousal rather than basic emotions influence long-term recognition memory in humans. Front. Behav. Neurosci. 10, 1–10 (2016).
84. Zimmer, U., Höfler, M., Koschutnig, K. & Ischebeck, A. Neuronal interactions in areas of spatial attention reflect avoidance of disgust, but orienting to danger. Neuroimage 134, 94–104 (2016).
85. Skipper, L. M., Ross, L. A. & Olson, I. R. Sensory and semantic category subdivisions within the anterior temporal lobes. Neuropsychologia 49, 3419–3429 (2011).
86. Tibon, R., Greve, A. & Henson, R. The missing link? Testing a schema account of unitization. Mem. Cogn. 1–18 (2018). doi:10.3758/s13421-018-0819-3
87. D'Angelo, M. C., Noly-Gandon, A., Kacollja, A., Barense, M. D. & Ryan, J. D. Breaking down unitization: Is the whole greater than the sum of its parts? Mem. Cogn. 45, 1306–1318 (2017).
88. Ritchey, M., Montchal, M. E., Yonelinas, A. P. & Ranganath, C. Delay-dependent contributions of medial temporal lobe regions to episodic memory retrieval. Elife 4, 1–19 (2015).

89. Diano, M. et al. Dynamic Changes in Amygdala Psychophysiological Connectivity Reveal Distinct Neural Networks for Facial Expressions of Basic Emotions. Sci. Rep. 7, 45260 (2017).
90. Dolcos, F. et al. Emerging directions in emotional episodic memory. Front. Psychol. 8, 1–25 (2017).
91. Lambert, H. K. et al. Hippocampal Contribution to Context Encoding across Development Is Disrupted following Early-Life Adversity. J. Neurosci. 37, 1925–1934 (2017).
92. Knowles, K. A., Cox, R. C., Armstrong, T. & Olatunji, B. O. Cognitive mechanisms of disgust in the development and maintenance of psychopathology: A qualitative review and synthesis. Clin. Psychol. Rev. 0–1 (2018). doi:10.1016/j.cpr.2018.06.002
93. Susskind, J. M. et al. Expressing fear enhances sensory acquisition. Nat. Neurosci. (2008). doi:10.1038/nn.2138
94. Riegel, M. et al. Nencki Affective Word List (NAWL): the cultural adaptation of the Berlin Affective Word List – Reloaded (BAWL-R) for Polish. (2015). doi:10.3758/s13428-014-0552-1
95. Wierzba, M. et al. Basic Emotions in the Nencki Affective Word List (NAWL BE): New Method of Classifying Emotional Stimuli. PLoS One 10, e0132305 (2015).
96. Murray, B. D. & Kensinger, E. A. The route to an integrative associative memory is influenced by emotion. PLoS One 9, 1–8 (2014).
97. Slotnick, S. D. & Schacter, D. L. A sensory signature that distinguishes true from false memories. Nat. Neurosci. 7, 664–672 (2004).
98. Wickens, T. D. Elementary signal detection theory. New York 1, (2002).
99. Ekman, P. An argument for basic emotions. Cognition & Emotion 6, 169–200 (1992).
100. Briesemeister, B. B., Kuchinke, L. & Jacobs, A. M. Emotion word recognition: discrete information effects first, continuous later? Brain Res. 1564, 62–71 (2014).
101. Theios, J. Reaction Time Measurements in the Study of Memory Processes: Theory and Data. Psychol. Learn. Motiv. 7, 43–85 (1973).
102. Weidemann, C. T. & Kahana, M. J. Assessing recognition memory using confidence ratings and response times. R. Soc. Open Sci. 3, 150670 (2016).
103. Whitfield-Gabrieli, S. & Nieto-Castanon, A. Conn : A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. Brain Connect. 2, 125–141 (2012).
104. Poldrack, R. A. & Farah, M. J. Progress and challenges in probing the human brain. Nature 526, 371–379 (2015).
105. Dimsdale-Zucker, H. R. & Ranganath, C. Chapter 27 – Representational Similarity Analyses: A Practical Guide for Functional MRI Applications. (2019). doi:10.1016/B978-0-12-812028-6.00027-6
106. Ashburner, J. et al. SPM12 Manual. Functional Imaging Laboratory Wellcome Trust Centre for Neuroimaging Institute of Neurology, UCL 12 Queen Square, London WC1N 3BG, UK (2014). doi:10.1111/j.1365-294X.2006.02813.x
107. Hofstetter, C., Achaibou, A. & Vuilleumier, P. NeuroImage Reactivation of visual cortex during memory retrieval : Content speci fi city and emotional modulation. Neuroimage 60, 1734–1745 (2012).

108. Marchewka, A. et al. What Is the Effect of Basic Emotions on Directed Forgetting? Investigating the Role of Basic Emotions in Memory. Front. Hum. Neurosci. 10, Article 378 (2016).
109. Wierzba, M. et al. Cognitive control over memory – individual differences in memory performance for emotional and neutral material. Sci. Rep. 8, 3808 (2018).
110. Mumford, J. A., Poline, J. B. & Poldrack, R. A. Orthogonalization of regressors in fMRI models. PLoS One 10, 1–11 (2015).
111. Flandin, G. & Friston, K. J. Analysis of family-wise error rates in statistical parametric mapping using random field theory. Proc. Natl. Acad. Sci. 113, 201602413 (2016).
112. Worsley, K. J. et al. A unified statistical approach for determining significant signals in images of cerebral activation. Hum. Brain Mapp. 4, 58–73 (1996).
113. Richter, F. R., Cooper, R. A., Bays, P. M. & Simons, J. S. Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. Elife 5, 1–18 (2016).
114. Rolls, E. T., Joliot, M. & Tzourio-Mazoyer, N. Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. Neuroimage 122, 1–5 (2015).
115. Brett, M., Anton, J.-L. L., Valabregue, R. & Poline, J.-B. Region of interest analysis using an SPM toolbox - Abstract Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2-6, 2002, Sendai, Japan. Neuroimage 16, Abstract 497 (2002).
116. Rolls, E. T., Joliot, M. & Tzourio-Mazoyer, N. Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. Neuroimage 122, 1–5 (2015).
117. Maldjian, J. A., Laurienti, P. J., Kraft, R. A. & Burdette, J. H. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19, 1233–1239 (2003).
118. Whitfield-Gabrieli, S. & Nieto-Castanon, A. CONN Functional Connectivity SPM Toolbox 2017. (2017).
119. Benjamini, Y. & Hochberg, Y. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B 57, 289–300 (1995).
120. Ritchey, M., Wing, E. A., LaBar, K. S. & Cabeza, R. Neural Similarity Between Encoding and Retrieval is Related to Memory Via Hippocampal Interactions. Cereb. Cortex 23, 2818–2828 (2013).
121. Kragel, P. a. & LaBar, K. S. Advancing Emotion Theory with Multivariate Pattern Classification. Emot. Rev. 6, 160–174 (2014).
122. Mumford, J. A., Turner, B. O., Ashby, F. G. & Poldrack, R. A. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. Neuroimage 59, 2636–2643 (2012).
123. Oosterhof, N. N., Connolly, A. C. & Haxby, J. V. CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. Front. Neuroinform. 10, 1–27 (2016).
124. Tulving, E. & Thomson, D. M. Encoding specificity and retrieval processes in episodic memory. Psychol. Rev. 80, 352–373 (1973).

Methods

Participants

Fifty-nine native Polish speakers (29 F and 30 M; aged 20-33; M = 24,81, SD = 3.14) without history of any neurological disorders or treatment with psychoactive drugs, right-handed, with normal or corrected-to-normal vision, gave written informed consent and participated in the study. They were mostly students and working people living in Warsaw, with a minimum of secondary education. Six subjects were excluded from the experimental group due to technical problems during experimental procedure and one additional subject did not take part in the second experimental session. Thus, behavioural and neuroimaging data collected from fifty-two subjects was further analyzed (24 F and 28 M; aged 20-33; M = 24,83, SD = 3.21). The local Research Ethics Committee at Faculty of Psychology, University of Warsaw approved the experimental protocol of the study.

Stimuli

360 words were selected from the Nencki Affective Word List (NAWL) [94]. According to the collected affective ratings and a novel method of classification based on Euclidean distances [95], a total of 120 words eliciting disgust, 120 words eliciting fear, and 120 neutral words was selected. As presented in Fig. 1, the stimuli were counterbalanced on all the other affective scales (valence, arousal, and intensities of other basic emotions). Subsequently, in each emotional category, 60 word pairs (6-25 characters) were formed in order to limit possible associations and suggest a specific meaning. The word pairs consisted of emotionally and semantically congruent words (M = 3.61, SD = .65) as rated by 5 independent raters (6-point scale, with -1 = words are opposites, 0 = words are not semantically congruent and 5 = words are highly semantically congruent). Thirty-five out of 60 word pairs in each emotion category (disgust, fear, neutral) were used in both encoding and retrieval sessions as targets, whereas 25 were additionally added as lures to the recognition session. The final set of word pairs can be illustrated with the following examples (translated into English) for disgust: swab - stinky, drainpipe - rat, shit – pigeon; for fear: throat - knife, bomber - drop, truck – collision; and for neutral state: circle - axis, word - Latin, facade - stone. The full list of Polish word pairs in the order of presentation during encoding and recognition is included in the Supplementary materials, Table S1 and S2.

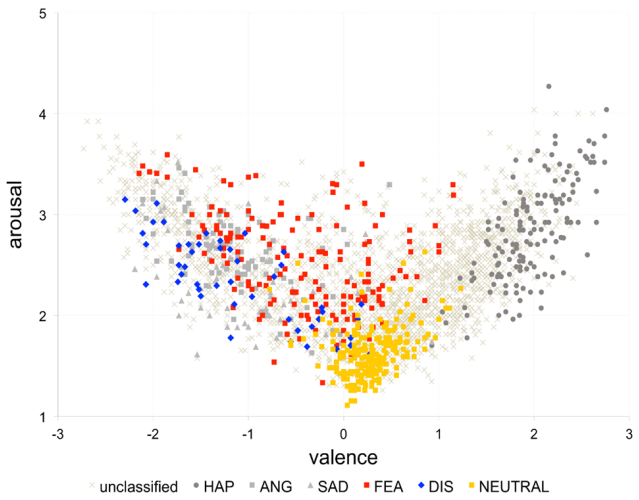

Fig. 1 Distribution of the NAWL words assigned to experimental conditions of basic emotions (disgust, fear and neutral) in the affective space of valence and arousal. Each dot represent a single word, blue – disgust, red – fear, yellow – neutral, word labels given to exemplary dots marked in circles; HAP – happiness, ANG – anger, SAD – sadness, FEA – fear, DIS - disgust (interactive browser: http://exp.lobi.nencki.gov.pl/nawl-analysis)

Study design and experimental paradigm

The study consisted of three experimental sessions, two of which (encoding and recognition) were conducted with the use of MRI scanner, with a delay period of 15-19 days (M = 16.86, SD = 1.22) (Fig. 2). After the second session and a short break, the third experimental session started during which subjective affective ratings were collected through an original platform available on a local server (http://exp.lobi.nencki.gov.pl/). All the experimental sessions took place in the Laboratory of Brain Imaging, Neurobiology Center, Nencki Institute of Experimental Biology in Warsaw, Poland.

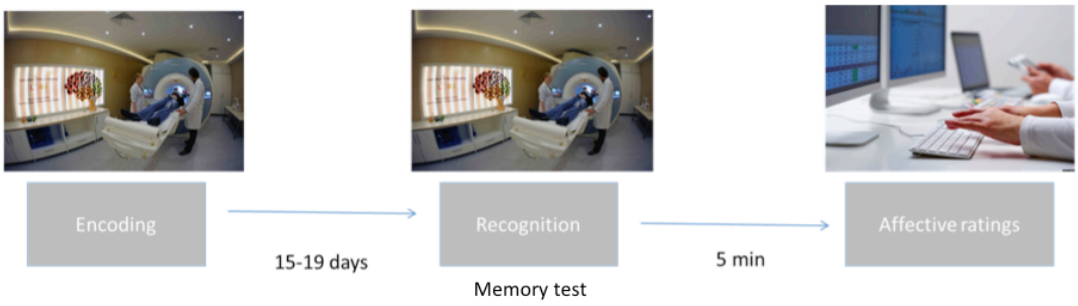

Fig. 2 Overview of experimental procedure.

Before entering the scanner, subjects were given the details of scanning procedure and performed a brief practice session of memory task (10 word pairs) in a mock MRI scanner. The experimental procedure was programmed using the Presentation software (Neurobehavioral Systems, Inc., Albany, CA, USA) and displayed on an MR-

compatible high-resolution LCD monitor positioned at the back of the scanner. Subjects observed the stimuli through the mirror system placed on the MR coil. The order of experimental trials in both sessions was pseudorandom under the following constraints: no more than 3 consecutive trials of the same emotion category (disgust, fear, neutral), no more than 3 of the same part of speech (noun, verb, adjective), no more than 4 old (targets) or new (lures) in the recognition session, and a maximized difference in semantic congruency (as rated by independent raters) of each two consecutive word pairs. In order to avoid the serial-position and recency effects, word pairs selected for both encoding and recognition sessions were divided into three parts (A, B, C), and presented in three variants of their order, counterbalanced across all the participants (encoding: ABC n = 21, BCA n = 20, CAB n = 18; recognition: ABC n = 20, BCA n = 20, CAB n = 19).

The first experimental session was encoding, as depicted in Fig. 3. The participants were presented with 105 word pairs (35 disgusting, 35 fearful and 35 neutral), instructed to imagine a single mental representation for each word pair and memorize as many of them as possible (4s + jittered fixation cross for 3-7s). The participants were presented with all the word pairs twice in order to strengthen the memory trace. The second presentation took place during a separate scanning session, right after the first one, in an identical order. During the second presentation, each word pair (4s + jittered fixation cross of 3-7s) was followed by a rating of integrative imageability success [96], i.e. how successful a participant was when trying to imagine a single mental representation for a particular word pair (3s; 4-point scale, with 1 = not successful and 4 = very successful). The instruction was formulated as follows: "Try to memorize as many pairs of words as possible. In order to do so, imagine each two words as one, in a common mental representation. You will be presented with the pairs twice, in the same order. After the second presentation, you will be asked to rate to what extent you can integrate these two words into one mental image (1-4 scale)." If not stated differently, fMRI data from both presentations are included in a single model in the presented analyses.

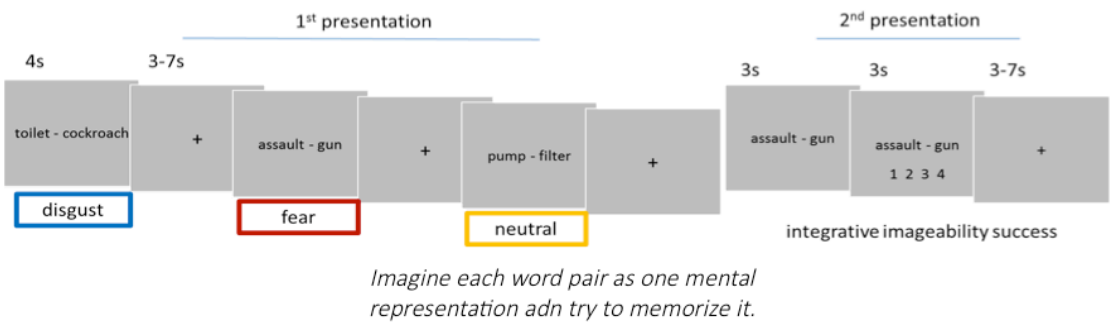

Fig. 3 Example trial structure for all the emotion conditions during encoding.

After 15-19 days, the subjects took part in the recognition session (Fig. 4). During an item recognition task, participants were presented with 105 word pairs from the encoding list and 75 other word pairs as lures, which gives 180 word pairs in total (2s + jittered fixation cross for 3-7s), and asked to determine (3s) whether a word pair was old (studied earlier) or new (not studied before). Additionally, they were asked to

indicate if they were sure or unsure of their responses (2s). The instruction was formulated as follows: "You will be presented with word pairs expressing different emotions or neutral. Your task will be to indicate if you can remember having seen them [O – old] or you cannot remember it [new – N], and if you are sure and remember your specific mental representation [97] [S – sure] or you are unsure and do not remember your specific mental representation [U – unsure]. The recognition session was divided into two runs with a short break in between, for the comfort of participants. The fMRI data from both runs are included in a single model in the presented analyses.

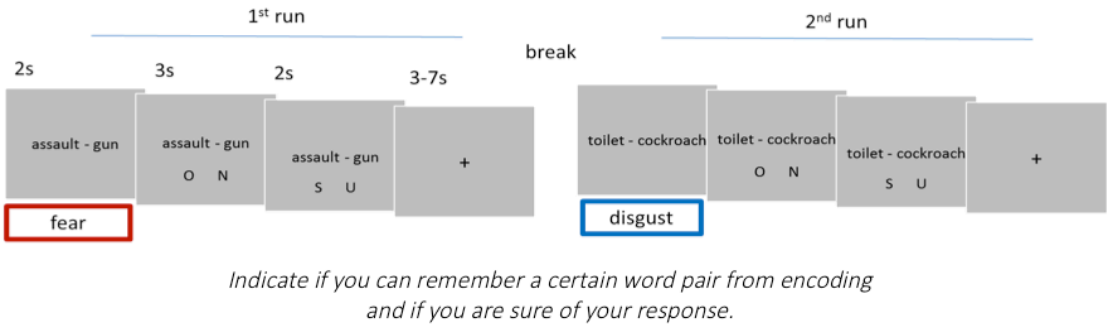

Fig. 4 Example trial structure for all the emotion conditions during recognition.

In order to further control the individual differences in affective processing, after fMRI scanning the subjects took part in the third experimental session of affective assessment (Fig. 5). During this session, each participant was presented with the instructions, and subsequently with all the 180 word pairs used during encoding and recognition sessions. The task of participants was to assess emotional properties of each word pair on all the given affective scales. Each word pair was presented in a full-screen view for 1 s. Then, the first scale of impact [45] appeared (1 for very low and 9 for very high). The participants were asked to consider each word pairs as a whole and assess whether they felt that its meaning created an instant and personal sense of impact on them. Following, the assessed word pair appeared in smaller font in the upper part of a new screen, next to all the other rating scales: an intensity of evoked disgust and fear (1 for not at all and 7 for very much), valence (-3 for negative, 0 for neutral and 3 for positive), and arousal (1 for unaroused and 5 for aroused) on the self-assessment manikin (SAM) (Lang et al., 1980). A specific word pair and rating scales remained visible to the participants until they completed all the ratings and pressed the "Next" button. Completing the whole affective assessment session took approximately 30 min.

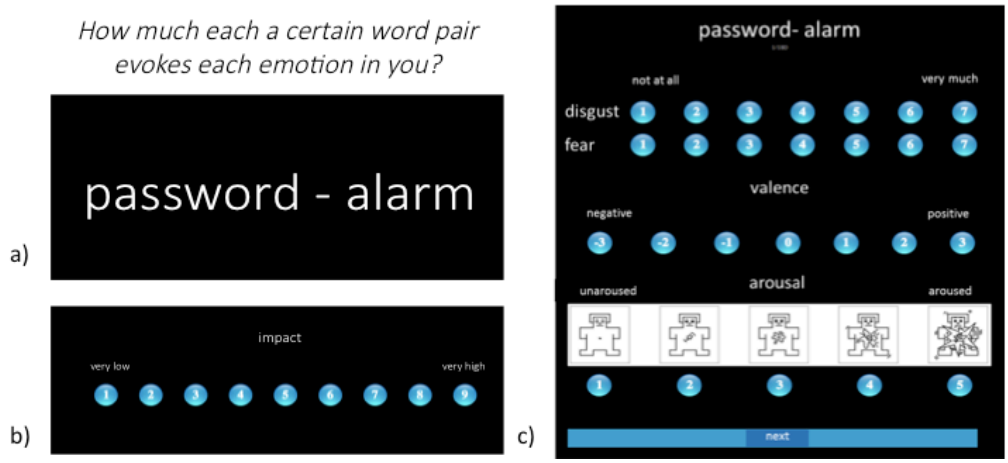

Fig. 5 Example trial structure with a) first presentation of a word pair, b) scale of impact, c) all the other affective scales.

Behavioral data analysis

First, memory for emotional and neutral stimuli was assessed as the proportion of correct responses in the recognition test. To analyze memory as a function of emotion category (3 levels: disgust, fear, neutral) and novelty (2 levels: old, new) as within-subject factors, repeated-measures (rm) analysis of variance (ANOVA) was performed on the recognition rate (proportion of correct responses) data. All the reported ANOVAs were performed with typically used Greenhouse-Geisser adjustments to the degrees of freedom when the assumption of sphericity was violated. Post hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons.

Since the subjects had to make memory decisions under conditions of uncertainty and could potentially use certain strategies of responding, signal detection theory (SDT) [98] was also applied to the collected data. Apart from the proportion of correct responses, that is old recognized as old (hits), the proportion of new incorrectly recognized as old (false alarms), the proportion of new correctly recognized as new (correct rejections) and the proportion of old incorrectly recognized as new (misses) were calculated. Based on these proportions, a sensitivity index (d') was calculated to provide a separation between the means of the signal and noise distributions, compared against the standard deviation of this distribution: Z (hit rate) – Z (false alarm rate). It was calculated separately for each emotion category (disgust, fear, neutral). Finally, rm ANOVAs were performed on the obtained d' values to test the effect of emotion (3 levels: disgust, fear, neutral) as a within-subject factor, on memory performance.

Next, in order to eliminate the possible confounds and assess the effect of emotion on memory performance more accurately, additional factor was included in the analysis. It has been long debated whether theoretical approaches of affective dimensions (Osgood & Succi, 1957) and basic emotions [99] are mutually exclusive or complementary [100] in explaining emotional experience. Therefore, we aimed at explaining whether any potential effect of basic emotions on memory could be

explained by an associated level of arousal. Repeated-measures analysis of covariance (ANCOVA) was performed on d' with emotion (3 levels: disgust, fear, neutral) as within-subject factor, for old and new word pairs separately, with a covariate of the difference in mean level of arousal between disgusting and fearful stimuli (for the details of this difference, see further analyses of affective ratings).

Reaction times (RT) were also analyzed to control for any differences in task difficulty as reflected by the speed of memory task performance [101] due to emotion. Recently, it has been also indicated that reaction times provide similar information to explicit ratings of memory confidence in their ability to quantify recognition memory decisions [102]. We performed rm ANOVA with emotion (3 levels: disgust, fear, neutral), novelty (2 levels: old, new) and correctness (2 levels: correct, incorrect) as within-subject factors. In order to exclude effects driven by the comparison of emotional with neutral word pairs, another rm ANOVA was performed on emotional word pairs only. Its goal was to differentiate between the effects of different emotions (2 levels: disgust, fear), novelty (2 levels: old, new) and correctness (2 levels: correct, incorrect) as within-subject factors.

The next research question was whether emotion effects are related to the strength of integration of encoded pieces of information or to the subjective strength of memory trace during recognition [12,28]. To answer these questions, we analyzed additional responses given during encoding and recognition sessions concerning the integrative imageability success and the sureness of recognition responses, respectively. First, we performed rm ANOVA concerning the level of integrative imageability success, with emotion (3 levels: disgust, fear, neutral) and correctness (2 levels: correct, incorrect) as within-subject factors (only old word pairs were analyzed, as the integrative imageability success was rated during encoding session). Second, we performed rm ANOVA of the percentage of sure responses, with emotion (3 levels: disgust, fear, neutral), and novelty (2 levels: old, new) as within-subject factors. Including correctness as an additional factor was impossible due to a low number of participants having all the possible combinations of factors and therefore, in this analysis only correct responses were included.

Individual affective ratings were used to analyze possible differences between emotion categories (disgust, fear, neutral) assigned to word pairs, as an experimental manipulation check. Mean valence, arousal, sadness, and fear ratings were calculated within at the subject-level, and rm ANOVA was performed with emotion (3 levels: disgust, fear, neutral) as within-subject factor for all the affective parameters. Paired t-tests for dependent samples were also performed to directly compare the mean affective ratings between disgusting and fearful stimuli.

fMRI data acquisition and preprocessing

Magnetic resonance data was acquired on a 3T MAGNETOM Trio TIM system (Siemens Medical Solutions) equipped with a whole-head 32-channel coil. The following images were acquired within a single subject scanning session: structural localizer image, structural T1-weighted (T1w) image (TR: 2530 ms, TE: 3.32 ms, flip

angle: 7°, PAT factor = 2, voxel size 1 x 1 x 1 mm, field of view 256 mm, volumes: 1), field map (TR: 488 ms, TE1: 5 ms, TE2: 7.46 ms, flip angle: 60°, voxel size 3 x 3 x 3 mm, field of view 216 mm, volumes: 1), first series of functional EPI images (TR: 2500 ms, TE: 30 ms, flip angle: 90°, PAT factor = 2, voxel size 3 x 3 x 2.5 mm, field of view 216 mm, volumes: 386 in encoding run 1, and 441 in recognition run 1), and second series of functional EPI images (same parameters, volumes: 470 in encoding run 2, 441 in recognition run 2). The whole scanning session during encoding took approximately 37 minutes, and during recognition – approximately 40 minutes, independent of participants' speed of responding.

At the initial step of fMRI data preprocessing, the DICOM series were converted to NIfTI using the MRIConvert (ver. 2.0.7; Lewis Center for Neuroimaging, University of Oregon). Then, the brain imaging data was preprocessed using Statistical Parametric Mapping (SPM12; Wellcome Department of Cognitive Neuroscience, University College London, London, UK) running under Matlab 2013b (Mathworks, Inc., Natnick, MA, USA). The preprocessing of data started with the correction of functional images for distortions related to magnetic field inhomogeneity and correction for motion by realignment to the first acquired image. Then, structural (T1w) images from single subjects were co-registered to the mean functional image and segmented into separated tissue classes (grey matter, white matter, cerebrospinal fluid) using the default tissue probability maps (TPM). Structural and functional images were normalized to the MNI space and resliced to preserve the original resolution. Finally, functional images were smoothed with the 6 mm FWHM Gaussian kernel. Unsmoothed data was used for the ROI-to-ROI functional connectivity analyses and pattern similarity analyses. For the purpose of all functional connectivity analyses, additional preprocessing steps were applied. Outlier time points were identified in functional images using the Artifact Detection Toolbox (ART) and a CompCor strategy was used for control of physiological and movement artifacts, both implemented in the CONN software, ver. 17.f. [103].

fMRI data analysis

fMRI data was analysed in line with three different approaches: univariate activation analysis, functional connectivity analysis and multivoxel pattern analysis (representational similarity analysis). The assumptions and rules of these approaches differ substantially. However, these methods answer different questions and provide complementary information [104,105]. A standard univariate fMRI analysis approach would examine the difference at each voxel between the averages across experimental conditions (for instance: emotion categories). Functional connectivity analysis investigates if there is a temporal synchronization in the activation of different brain regions related to the experimental conditions. Representational similarity analysis (RSA) examines the correlations between multivoxel activity patterns for each item (single trials) across encoding and retrieval.

Univariate activation analysis

fMRI data was initially analyzed with SPM12 based on a mass-univariate approach and general linear models (GLMs). At the subject level, each single event was modeled

with onsets corresponding to the presentation of a word pair (together with a response given afterwards, as it was not jittered in time from a word pair presentation period) and with corresponding durations. To account for movement-related variance, six nuisance regressors were included as representing the differential of the movement parameters from the realignment. Data was high-pass filtered (1/128 Hz), corrected for intrinsic autocorrelations, and convolved with a standard canonical hemodynamic response function (HRF) to approximate the expected blood-oxygen-level dependent (BOLD) signal. Because of high variability in the inclusion of the PRC region in the individual brain masks, we constructed an alternative version of individual brain masks always including this region, using the Imcalc tool (http://tools.robjellis.net/) for the SPM toolbox. In order to incorporate PRC masks into standard individual masks, we used the logical "OR" expression and the masking threshold of –Inf. The resulting individual masks including PRC were included in an alternative version of the first analysis, as described below.

In the first analysis, the interaction between emotion and memory performance was examined for encoding and recognition sessions separately. Functional volumes from encoding session were split into conditions along to the factors of emotion (DIS, FEA, NEU) and subsequent memory performance (corr, incorr). Thus, the following experimental conditions were specified: DIS corr, DIS incorr, FEA corr, FEA incorr, NEU corr, NEU incorr and miss. The number of trials falling into each condition was dependent on individual subsequent memory performance and varied between subjects (ranging from 10 to 35). Given that numerous subjects had less than 10 trials in the incorrect conditions, only correct conditions were further included in the group-level analysis. At the group level, flexible factorial design [106] was chosen because of the abovementioned variability in the number of trials among subjects and possible subject effects. Event-related stick-function regressors were used to perform ANOVA with emotion (three levels: DIS, FEA, NEU) as a within-subject factor, as well as a subject factor. Whole-brain random-effects contrasts were evaluated to obtain estimates of activity in response to each trial type relative to implicit baseline. The maps were created for the following directional T-contrasts: EMO – NEU (positive effect of emotion), DIS – FEA and FEA – DIS (difference between emotions).

Functional volumes from recognition session were split at the subject level into conditions according to emotion (DIS, FEA, NEU), memory performance (corr, incorr) and novelty (old, new). As a result, the following conditions were specified: old DIS corr, old DIS incorr, old FEA corr, old FEA incorr, old NEU corr, old NEU incorr, analogical conditions for new word pairs, and miss. Once again, the number of each condition trials was dependent on individual memory performance (ranging from 10 to 35 for old and from 10 to 25 for new conditions). Given that numerous subjects had less than 10 trials in the incorrect conditions, again only correct conditions were further included in the group-level analysis. Similar to data analysis from encoding, at the group level, ANOVA was performed using flexible factorial design [106], with the following factors: emotion (three levels: DIS, FEA, NEU) and novelty (two levels: old, new) as within-subject factors, as well as a subject factor. Maps were created for the following directional T-contrasts: EMO – NEU (positive effect of emotion), DIS – FEA and FEA – DIS (difference between emotions). In order to test the positive effect of

emotion on correct memory of old word pairs while controlling for more general emotional effects of new word pairs [107], maps were created for the following contrasts: old EMO – NEU masked exclusively ($p < .05$) by new EMO – NEU (positive effect of emotional memory), old DIS – FEA masked exclusively ($p < .05$) by new DIS – FEA (difference between memorized emotions) and old FEA – DIS masked exclusively ($p < .05$) by new FEA – DIS (difference between memorized emotions).

The second analysis was similar to the first one and also examined the interaction between emotion and memory performance during correct encoding and recognition. However, in order to isolate the effects of basic emotions (disgust, fear), modulation by the level of arousal was regressed out [83,108]. In order to perform this kind of operation, at the subject-level, individual arousal ratings collected during the third experimental session were included in the model. They were manually convolved with HRF and added as an additional regressor of no interest. Event-related stick-function regressors isolated from the effects of arousal were subsequently used to perform ANOVAs at the group level in a manner described above.

The third analysis was performed to further examine how the general effects of memory were modulated by the individual ratings of each basic emotion (fear, disgust), affective dimensions (valence, arousal), and impact. At the subject-level, individual ratings were applied to respective experimental conditions (corr and incorr at encoding; old corr, old incorr, new corr, new incorr at recognition) as linear (first-order) modulators [83,109]. Conditions with zero variance in a parameter vector were discarded from further analysis. A single parameter was added to the model at a time to avoid orthogonalization effects between parameters [110]. However, each parametric value was mean-centered to orthogonalize this variable with respect to the corresponding condition variable. Altogether, five additional models were created for each subject. The effects (stick contrasts) of parametrically modulated conditions were then included in the group-level analyses. At the group level, ANOVAs were performed for each parameter separately, by defining a within-subject factor of memory performance (corr and incorr at encoding; old corr, old incorr, new corr, new incorr at recognition). The positive correlation was tested for each emotion parameter and each condition (the higher value of fear, disgust, arousal and impact parameter, the more emotional a given word pair). Only in the case of valence, negative correlation was analyzed (the lower value of valence parameter, the more emotional a given word pair).

If not stated otherwise, in all the analyses, a voxel-wise height threshold of $p < .001$ (uncorrected) combined with a cluster-level extent threshold of $p < .05$, corrected for multiple comparisons using the family-wise error (FWE) rate [111] was applied in the whole brain analyses. The small volume correction (svc.) [112] was applied using regions-of-interest (ROI) anatomical masks defined based on a priori hypotheses about the engagement of MTL regions in emotion and memory interactions. Within each of these ROIs, we considered reliable activations whose effects survived the small volume FWE correction at the voxel level. This type of correction for multiple comparisons was used in the previous papers related to the topic of emotion and memory due to a small volume of subcortical structures of interest [82,113]. The

coordinates of significant effects are reported in the Montreal Neurological Institute (MNI) space and are labeled according to Automated Anatomical Labeling (AAL2) [114] atlas with the use of bspmview (http://www.bobspunt.com/bspmview). Results were visualized with the use of MRIcroGL (http://www.mccauslandcenter.sc.edu/mricrogl/home).

Region-of-Interest localization and analyses

To further test specific hypotheses concerning differences in emotional modulation of successful encoding (followed by correct recognition) in the regions suggested by available literature, an additional ROI analysis was performed using a MarsBaR toolbox [115]. Contrast estimate values were extracted from the first univariate ANOVA analysis mentioned above in the following regions, bilaterally: amygdala (AMY), perirhinal cortex (PRC), hippocampus (HC) and parahippocampal gyrus (PHG). Anatomical masks for the ROI analysis were specified according to the AAL2 [116] atlas implemented in the WFU PickAtlas [117] toolbox, version 3.0.5. The PRC masks used in these analyses were manually segmented as described in [88].

As for emotion effects in these ROIs in encoding data, first we performed a rm ANOVA on the stick contrast estimates extracted from selected regions, with the ROI (8 levels: AMY, PRC, HC, PHG, each left and right) and emotion (3 levels: disgust, far, neutral) as within-subject factors. To unpack these effects, we run a post hoc analysis of simple effects restricted to the two emotion categories (DIS, FEA). Specifically, we performed a paired t-test for dependent samples to directly compare the mean contrast estimates between disgust- and fear- related stimuli in each ROI.

As for emotion effects in these ROIs in retrieval data, we performed an rm ANOVA on the stick contrast estimates for old word pairs from selected regions, with the ROI (8 levels: AMY, PRC, HC, PHG, left and right) and emotion (3 levels: disgust, far, neutral) as within-subject factors. Second, we performed a post hoc analysis of simple effects, namely a paired t-test for dependent samples to directly compare the mean contrast estimates between correctly recognized old disgusting and fearful stimuli in each ROI. All the reported ANOVAs were performed with typically used Greenhouse-Geisser adjustments to the degrees of freedom when the assumption of sphericity was violated. Post hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons.

fMRI functional connectivity analysis

Functional connectivity analysis aimed at investigating the temporal synchronization between the activation of brain regions involved in memory and emotion interactions was performed using the CONN software, ver. 17.f. [103]. Task-related connectivity changes were analyzed based on the subject-level SPM models from the first of univariate analyses described above. Experimental conditions with onsets and durations were imported from those models. Structural and functional images had been already preprocessed in SPM 12, including a normalization to MNI-space. An additional step of ART-based scrubbing was run for the outlier scans identification and a first-level covariate containing the offending scans was created for each subject and session. For the voxel-level analyses, BOLD signal estimates were extracted from the

smoothed data, whereas for the ROI-level analyses, BOLD signal estimates were extracted from unsmoothed data in order to avoid spillage from nearby regions [118]. The default ROI masks were imported for cortical and subcortical areas (AAL atlas), and a few commonly used networks and areas as described above.

In a denoising step, linear regression and band-pass filtering were applied in order to remove unwanted motion, physiological, and other artifactual effects from the BOLD signal before computing connectivity measures. Three possible sources of confounds were defined: BOLD signal from the white matter and CSF masks, within-subject covariates (movement and scrubbing parameters) and the main effects of task conditions (direct BOLD signal changes associated with the presence/absence of a task) and regressed out. The band-pass filter of [0.008 Hz – Inf] was applied to keep higher-frequency information related to the task.

At the subject-level, generalized PsychoPhyisiological Interaction (gPPI) [30] was used to compute the interaction between the seed/source BOLD timeseries and a chosen condition-specific factor when predicting each voxel or target ROI BOLD timeseries. Bivariate regression was computed for each pair of source and target ROIs (ROI-to-ROI analysis). The obtained regression coefficients were used for the group-level analysis of increases and decreases in functional connectivity. Similar to the univariate activation analyses, the following contrasts were tested for encoding session: EMO – NEU, DIS – FEA, and FEA – DIS. The results were reported at a two-sided threshold of $p < 0.05$, with false discovery rate (FDR) [119] and correction for multiple comparisons at the seed-level.

fMRI multivariate pattern analysis

Previous analyses of brain activity were performed on the encoding and recognition data separately. However, another very interesting aspect is how emotion modulates the relationship between brain activation during encoding and retrieval [120]. In order to examine emotion modulation of the similarity of transient, stimulus-evoked BOLD activity between encoding and recognition, we performed a split-half correlation based on the representational similarity analysis (RSA) [27]. In this multivariate approach, data from individual voxels within a region are jointly analyzed, which is closer to a contemporary view on the brain representation of different mental states [121].

The first step of this analysis was creating the separate subject-level GLM models for each of 105 encoding trials and each of 180 recognition trials to estimate the single-trial response. A Least-Square Single (LSS) method was used to model a certain trial as a regressor of interest and combine all other trials as a single nuisance regressor [122]. Each trial's onsets and durations were extracted from the existing subject-level models as described in the first univariate analysis above. The obtained single-trial beta images were then renamed and moved to a common folder, sorted identically for encoding and recognition, and concatenated into 4D NIfTI format for further analyses.

Then, the split-half correlation RSA was performed with the CoSMO MVPA [123] toolbox and in-house MATLAB scripts. Both searchlight [33] and ROI-based methods were used. First, we applied a searchlight method, where a sphere-shaped mask 'travels' through the brain, and at each location a specific measure (here: a correlation coefficient) is assigned to the center voxel of the sphere, resulting in an informative map (here: of correlation differences). In order to generate searchlight maps for each participant, a sphere of 3-voxels radius (on average, 110 voxels per sphere) was defined as a mask. In a given searchlight sphere, the multivoxel response patterns (betas) for each encoding and recognition trial (item) were extracted. In order to estimate response pattern similarities between each encoding and recognition item pair, pair-wise Pearson correlations were calculated. The resulting Fisher-transformed r-to-z score was assigned to the center voxel of the sphere for each specific encoding-recognition pairings. Thus, each voxel had over 11000 values associated with it, 105 of which represented the voxelwise correlation between matching encoding - recognition item pairs, whereas the remaining values represented the voxelwise correlation between the non-matching encoding – recognition item pairs.

Second, using the ROI-based approach to the split-half correlation RSA, we defined the ROIs for the analysis. We expected that the similarity of brain activation patterns between encoding and recognition could arise not only from the reactivation of perceptual processes, but also from higher-order cognitive and affective processes in MTL regions. Therefore, all the ROIs selected for the ROI univariate analysis were also selected for the split-half correlation RSA. Since the searchlight method is actually a repeated equivalent of an ROI-based method, also in each ROI the multivoxel response patterns (betas) for each encoding and recognition trial (item) were extracted. Pair-wise Pearson correlations were calculated to estimate the response pattern similarities between each encoding and recognition item pair. The resulting Fisher-transformed r-to-z score was assigned to each ROI and again, each ROI obtained multiple values of correlation between matching and non-matching encoding – recognition pairs.

Finally, the index of encoding-retrieval similarity (ERS) [52,124] was calculated for each sphere and ROI per participant, as a mean correlation for the matching pairs minus a mean correlation for the non-matching pairs. To this end, positive values were put on the diagonal of the contrast matrix, and small negative values were put on the off-diagonal area of the contrast matrix, which is depicted in Fig. XX. The resulting single value was assigned to the central voxel of each sphere and each ROI. However, these values of the ERS index would not be still informative regarding the experimental conditions. Therefore, a contrast matrix was created for the resulting values of ERS index to directly compare different levels of correctness (2 levels: correct, incorrect) and emotion (3 levels: disgust, fear, neutral) factors. The ERS index was also calculated separately for each emotion condition for further statistical comparisons. In the case of searchlight method, the t values resulting from the contrasts were again assigned to each central voxel of the sphere, and the obtained maps of t-values were tested against a null hypothesis of zero using one-sample t-test across subjects. One sample t-test across subjects was also performed on the single t-values assigned to each ROI. Finally, rm ANOVAs were performed for each ROI on the t-values calculated

for each emotion condition separately, with emotion (3 levels: disgust, fear, neutral) as within-subject factors.

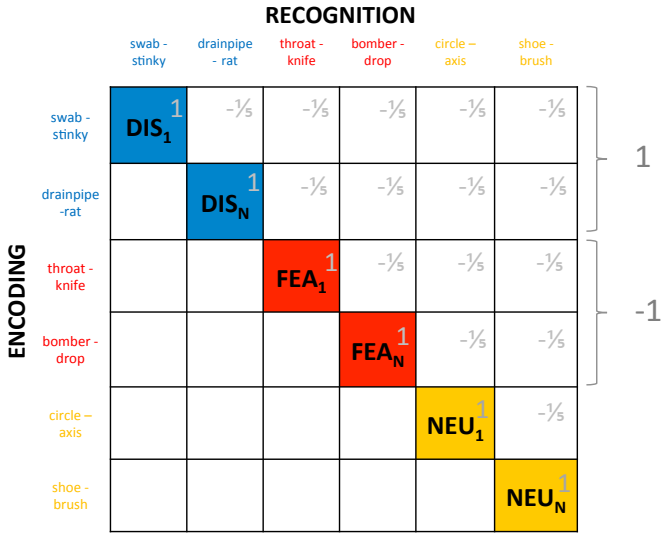

Fig. 6 Contrast matrix of the ERS calculated for the matching (diagonal) and non-matching (off-diagonal) pairs of encoding and retrieval trials. DIS – disgust, FEA – fear, NEU – neutral. It was calculated for each emotion separately for further statistical testing. The contrast matrix was also modified to directly compare the ERS values between emotion and correctness conditions.

Encoding – reinstatement analysis

Due to variability in encoding-retrieval similarity across trials, the correspondence between encoding and retrieval might not be fully illustrated by the results of multivariate ERS analysis averaged across conditions. Thus, the goal of the final analyses was twofold (based on [36]). First, it was aimed to investigate the univariate mechanism of reinstatement across encoding and retrieval in the univariate activations. Specifically, we queried the retrieval data for emotion-specific regions from encoding whose trial-by-trial activity from recognition correlated with trial-by-trial estimates of encoding activity in specific ROIs, namely bilateral AMY, HC, PHG and PRC. Second, we wanted to ask how univariate activation during encoding was related to the reinstatement as measured by the ERS. It was aimed to query the encoding data for regions whose trial-by-trial activity correlated with trial-by-trial ERS calculated for emotion-specific masks. The emotion-specific masks limited the scope of both analyses to the regions that were preferentially activated during encoding of specific emotions, and whose activity during recognition could be therefore considered to be 'reactivation'. The masks were created from the following contrasts: DIS – FEA, FEA – DIS and NEU – (DIS + FEA) all the masks including voxels at $p < .001$ uncorrected.

In the first reinstatement analysis, we initially used in-house scripts based on Marsbar to extract contrast estimates in the abovementioned ROIs for each encoding trial. As presented in Fig. XX, these trial-by-trial contrast estimates were further used as parametric regressors applied to the recognition data in a model that accounted for parametric effects within each emotion and correctness condition. In other words,

each trial at recognition obtained a parametric value corresponding to the AMY, PRC, HC and PHG activation for that trial at encoding. Each individual subject's model included the following regressors for old conditions: DIS corr, DIS corr parametric, DIS incorr, DIS incorr parametric, FEA corr, FEA corr parametric, FEA incorr, FEA incorr parametric, NEU corr, NEU corr parametric, NEU incorr, NEU incorr parametric; and new DIS corr, new DIS incorr, new FEA corr, new FEA incorr, new NEU corr, new NEU incorr and miss. The results were limited to emotion-specific regions from encoding. This approach was aimed to identify regions that exhibit trial-by-trial retrieval activity correlated with trial-by-trial AMY, HC, PHG and PRC encoding activity across experimental conditions.

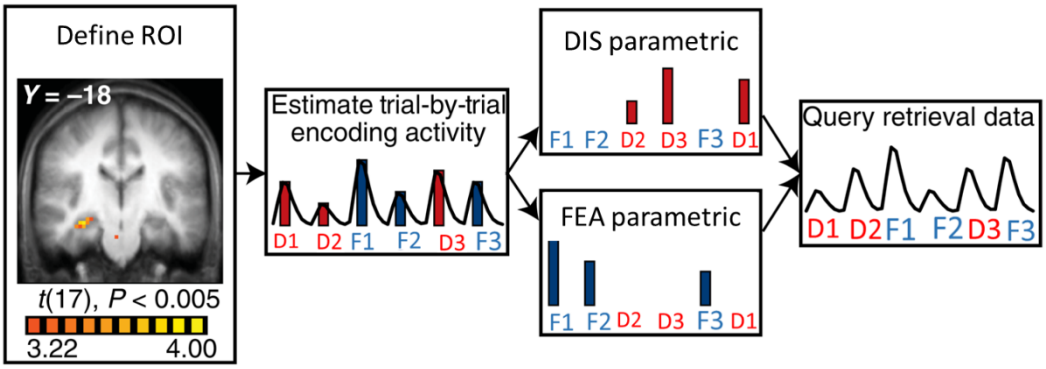

Fig. 7 Schematic of the encoding – reinstatement analysis with parametric modulation of recognition data by specific ROI activation during encoding. D1 – 1st disgust trial, D2 – 2nd disgust trial, D3 – 3rd disgust trial, F1 – 1st fear trial, F2 – 2nd fear trial, F3 – 3rd fear trial (modified after [36]).

In the second reinstatement analysis, emotion-specific masks from encoding were used as ROIs for the ERS analysis as described above (DIS – FEA used for disgusting trials, FEA – DIS used for fearful trials and NEU – (DIS + FEA) used for neutral trials). These trial-by-trial ERS estimates were used as parametric regressors applied to encoding data in two models, one of which accounted for parametric effects of ERS within correctness factor only. Thus, as presented in Fig. XX, each trial at encoding obtained a parametric value corresponding to the ERS value for that trial. In this model, ERS values for disgusting, fearful and neutral trials were collapsed into a single parametric regressor and this analysis was meant to find encoding regions that predicted trial-by-trial ERS across emotions. Since ERS values for each emotion were calculated based on different trials and different set of voxels, they were separately mean-centered. Each individual subject's model included the following regressors: corr, corr parametric, incorr, incorr parametric and miss. Next, in order to investigate the various effects for emotions, another parametric model was constructed which included separate regressors for each level of emotion and correctness factors. Here, each individual subject's model included the following regressors: DIS corr, DIS incorr, FEA corr, FEA incorr, NEU corr, NEU incorr and miss. In general, using this approach, we could identify regions exhibiting activation during encoding that correlated with broad similarity in patterns of activity between encoding and their corresponding retrieval trials, across experimental conditions.

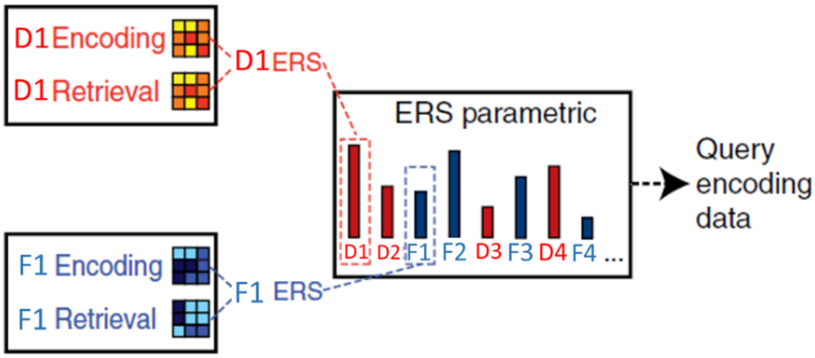

Fig. 8 Schematic of the encoding – reinstatement analysis with parametric modulation of recognition data by specific ROI activation during encoding. D1 – 1st disgust trial, D2 – 2nd disgust trial, D3 – 3rd disgust trial, D4 – 4th disgust trial, F1 – 1st fear trial, F2 – 2nd fear trial, F3 – 3rd fear trial, F4 – 4th fear trial (modified after [36]).