

# Emotional content and reality-monitoring ability: fMRI evidence for the influences of encoding processes

Elizabeth A. Kensinger<sup>a,b,\*</sup>, Daniel L. Schacter<sup>a,b</sup>

<sup>a</sup> Department of Psychology, Harvard University, 33 Kirkland Street, Room 884, Cambridge, MA 02138, USA

<sup>b</sup> The Athinoula A. Martinos Center for Biomedical Imaging, Cambridge, MA 02129, USA

Received 27 May 2004; received in revised form 2 January 2005; accepted 6 January 2005

Available online 16 March 2005

## Abstract

Memory for emotional items can be less prone to some types of memory distortion, such as reality-monitoring errors, than memory for neutral items. The present fMRI study examined whether this enhanced reality-monitoring accuracy reflects engagement of distinct processes recruited during encoding of emotional information. Participants only imagined named objects (word-only trials) or imagined named objects and then also viewed photos of them (word-picture trials). Half of the items were emotional (e.g., snake, casket). Later, participants heard object names and indicated whether the corresponding photo had been shown. Reality-monitoring errors occurred when participants attributed an item from a word-only trial to a word-picture trial. Such misattribution errors occurred less frequently for emotional than neutral items. Activity in emotion-processing regions (e.g., orbitofrontal cortex, amygdala) reduced the likelihood of later misattributions, likely due in part to interactions with regions that promote memory accuracy (e.g., the hippocampus). Distinct neural processes also increased the likelihood of reality-monitoring errors, depending on the emotional content of the items. Activity spanning the fusiform and parahippocampal gyri (likely reflecting mental imagery) increased the likelihood of reality-monitoring errors for neutral items, while activity in the anterior cingulate increased the likelihood of reality-monitoring errors for emotional items.

© 2005 Elsevier Ltd. All rights reserved.

**Keywords:** Emotion; Memory; Memory distortion; fMRI; Encoding

## 1. Introduction

A common form of memory distortion arises when individuals are asked to distinguish between previously imagined and perceived events (Johnson & Raye, 1981). Individuals sometimes confuse the two, indicating that imagined events or actions were perceived or performed (*reality-monitoring errors*). In daily life, this confusion can lead to minor annoyances (e.g., Did I remember to take out the trash? Or did I only imagine taking it out?) or may have more serious consequences (e.g., Did I take my medication this morning? Or did I only think about taking it?).

The information encoded during an imagined or perceived event can alter the probability of subsequent reality-monitoring errors. Internally generated and exter-

nally presented items typically differ in the extent to which different types of information are associated with their presentations. Memories for perceived items tend to have greater sensory detail, while memories for imagined items contain more information about the cognitive operations performed during encoding (Johnson & Raye, 1981). Encoding, and subsequent retrieval, of these features often allows correct assignment of an item to perception or imagination. Conversely, memory errors occur when individuals fail to remember the event details that would allow proper assignment to an internal or external source (e.g., failing to remember the cognitive operations associated with forming a mental image, or the sensory details of a perceived event). Reality-monitoring errors also frequently occur when only part of the information in the encoding episode is remembered. For example, remembering sensory details associated with a vivid mental image, but not the cognitive operations performed to form the mental image, may trigger assignment of that visual image

\* Corresponding author. Tel.: +1 617 496 5909; fax: +1 617 496 3122.  
E-mail address: ekensing@wjh.harvard.edu (E.A. Kensinger).

to a presented source (Johnson, Hashtroudi, & Lindsay, 1993).

Consistent with this framework, recent neuroimaging evidence has revealed that encoding-related activity in regions implicated in visual imagery increases the likelihood of reality-monitoring errors. In an event-related potential (ERP) study, Gonsalves and Paller (2000) demonstrated that enhanced positivity at posterior electrodes during encoding of verbal referents of objects (hypothesized to reflect vivid mental imagery of the object) correlated with the probability of later falsely believing that a photo of the object had been shown. An fMRI extension of the study (Gonsalves, Reber, Gitelman, Parrish, Mesulam, & Paller, 2004) confirmed that recruitment of regions important for mental imagery and spatial attention (the precuneus and inferior parietal cortex) increased the likelihood of reality-monitoring errors. Thus, individuals are particularly prone to reality-monitoring errors when a robust or vivid mental image is formed. In these instances, participants presumably remember the sensory details associated with the event, but not the cognitive operations performed, and thus misattribute the imagined item to an external source.

Although prior neuroimaging investigations of memory distortion have examined memory for information devoid of emotional relevance (for review, see Schacter & Slotnick, 2004), recent investigations have begun to address the effect of emotional content on the probability of memory distortion. The results from behavioral studies indicate that, at least in some instances, memory for items with emotional salience is less likely to be distorted than memory for neutral items (Kensinger & Corkin, 2004b; Kensinger & Schacter, unpublished observations; Pesta, Murphy, & Sanders, 2001). Because these studies are behavioral in nature, they cannot readily determine whether processes engaged at encoding (versus retrieval) contribute to the effect. Nor can they elucidate whether the effect is mediated by neural processes that are specifically related to processing of emotional information, or whether the reduction in memory distortion stems from engagement of the same processes that allow accurate remembering of neutral information.

Although parsimony favors the hypothesis that the same processes are recruited to remember emotional and neutral information, data from lesion and neuroimaging studies have provided strong evidence that distinct processes increase the likelihood that emotional information is remembered. These studies have confirmed a role for the amygdala in the memory enhancement effect (reviewed by Hamann, 2001), likely via its connections to the hippocampus (reviewed by McGaugh, 2000). Encoding-related activity in the amygdala and hippocampus correlates with the likelihood of remembering emotional information (Cahill et al., 1996; Canli, Zhao, Brewer, Gabrieli, & Cahill, 2000; Dolcos, Graham, LaBar, & Cabeza, 2003; Hamann, Ely, Grafton, & Kilts, 1999; Kensinger & Corkin, 2004a), and the amount of activity in the amygdala and hippocampus is correlated during successful encoding of emotional information

(Kensinger & Corkin, 2004a; Richardson, Strange, & Dolan, 2004).

While interactions between the amygdala and the hippocampus increase the likelihood of remembering emotional information, it is unclear to what extent activity in those regions also would decrease the frequency of subsequent memory distortions, such as reality-monitoring errors. Given the role of the hippocampus in encoding source information (e.g., Davachi, Mitchell, & Wagner, 2003), and the influence that the amygdala is thought to exert on hippocampal function (reviewed by Phelps, 2004), it is plausible to propose that interactions between these regions would also be critical for reducing reality-monitoring errors for emotional information. This possibility, however, has never been empirically addressed, and there is a conceivable alternate explanation. As noted earlier, previous studies have shown that the likelihood of remembering emotional information tends to be greater than the likelihood of remembering non-emotional information. Other studies have shown that a variety of manipulations that increase the likelihood of remembering previously studied non-emotional information, ranging from increasing numbers of study list presentations to enhancing distinctiveness of studied material, also decrease the incidence of false memories (e.g. Budson, Daffner, Desikan, & Schacter, 2000; Kensinger & Schacter, 1999; Schacter, Israel, & Racine, 1999; Schacter, Verfaellie, Anes, & Racine, 1998). Therefore, increased memory accuracy following encoding of emotional information may simply reflect this general tendency for manipulations that increase the overall level of memory performance to also decrease the incidence of memory errors and distortions (i.e., the “mirror effect”, e.g., Glanzer & Adams, 1990). In other words, perhaps the reductions in false memories for emotional items do not stem from engagement of specific processes recruited for the encoding of emotional information but instead arise from additional recruitment of processes that tend to boost memory accuracy for neutral information. fMRI is well-suited to clarify whether similar or distinct processes support increased accuracy for emotional versus neutral information. Thus, we used fMRI to examine what neural processes, engaged during encoding of emotional and neutral information, were related to the probability of subsequent reality-monitoring errors.

We used a paradigm shown previously to elicit high levels of reality-monitoring errors (see Gonsalves & Paller, 2000). During encoding, individuals indicated whether named objects were bigger or smaller than a shoebox; to make this decision, they were encouraged to use mental imagery. After making their responses, participants were either shown a photo depiction of the object (on half of the trials) or a blank square (Fig. 1). Half of all items were emotional (e.g., *snake*, *casket*) and half were neutral (e.g., *blender*, *canoe*). In some instances, memory was distorted, and participants later indicated that they had studied pictures of items that, in actuality, they only had imagined (i.e., a *reality-monitoring error*). The goals of this study were to elucidate the neural processes influencing the frequency of reality-monitoring

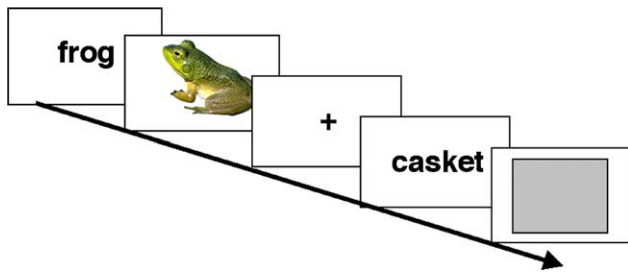


Fig. 1. Methods used to induce high levels of memory distortion. The encoding task, performed in the scanner, required participants to imagine named objects (half emotional, half neutral) and to indicate whether the objects were bigger or smaller than a shoebox. After making their response, participants were either shown the corresponding photo or a blank square. Following a one- to two-day delay, participants performed a surprise recognition test in which they heard words and indicated whether they had viewed the corresponding picture while in the scanner.

errors and to determine whether those processes differed for emotional and neutral items.

## 2. Methods

### 2.1. Participants

Participants were 19 young adults. The data from three participants were excluded: one due to scanner malfunction, one due to excessive head movement, and one due to failure to comply with task instructions. The remaining 16 young adults (9 women) were ages 18–30. All were right-handed, native English speakers screened to exclude those with contra-indicators for MRI scanning or with a history of depression. No participant was taking centrally active medications. Informed consent was obtained from all participants before both the behavioral and imaging portions of the study. The Harvard University Institutional Review Board approved the behavioral protocol and the Massachusetts General Hospital Institutional Review Board approved the imaging protocol.

### 2.2. Materials and procedure

Materials comprised 450 concrete words and 450 photo objects. Photo objects (Hemera Technologies Inc., 2002, Canada) depicted a single object (e.g., a baseball, a spider) devoid of any background. Words and photo objects were selected as pairs, such that each word named a photo object (e.g., “spider” and a picture of a spider).

Items were divided into three lists of 150 items, and were judged for arousal levels by a separate group of 20 young adults (10 males) and by the participants in this study. Half of the words and objects were high in arousal, with scores greater than 2.5 (mean for words = 2.96, mean for pictures = 3.07) on a scale of –5 to +5. (Negative values indicated that an item was calming or soothing, and positive values indicated that an item caused excitement or agitation.)

The other half of the words and objects were neutral, having received arousal ratings lower than +1 (mean for pictures = 0.67, mean for words = 0.59). The emotional words and pictures were significantly higher in arousal than the neutral words and pictures (all  $t > 15$ ,  $p < 0.0001$ ). Neutral and emotional objects did not differ in the numbers that included animals, people, or objects. Neutral and emotional words did not differ in word length (mean for neutral = 6.3, mean for emotional = 6.2,  $p > 0.5$ ), word frequency (mean for neutral = 18.4, mean for emotional = 24.0,  $p > 0.25$ ; Kucera & Francis, 1967), familiarity (mean for neutral = 503, mean for emotional = 496) or imageability (mean for neutral = 571, mean for emotional = 565, all  $p > 0.25$ ). Familiarity and imageability were estimated from a combination of sources (Gilhooly & Logie, 1980; Paivio, Yuille, & Madigan, 1968; Toglia & Battig, 1978; see Colthart, 1981).

Each participant viewed three study lists (each corresponding to one functional scan). In each of the three study lists, 25 emotional words and 25 neutral words were presented for 2 s, followed by a blank square for 2 s (“word-only” trials). An additional 25 emotional words and 25 neutral words were presented for 2 s, followed by their corresponding photo object for 2 s (“word-picture” trials, Fig. 1). The items shown on word-only versus on word-picture trials were counterbalanced across participants. We did not jitter the spacing between the word trial component and the picture (or square) trial component because we were concerned that participants’ strategies would differ when the trial components were separated by a long delay versus by a short delay.

Participants were instructed to make a button press to indicate whether each word named an object that was bigger or smaller than a shoebox. Approximately half of the items were bigger than a shoebox, and there were no differences in the size judgments made for emotional and neutral items ( $p > 0.25$ ). Participants were told that the task assessed mental imagery performance, and thus were encouraged to use mental imagery to perform the size-judgment task. Participants were instructed to view the square or the photo object that occurred after the word, and they were told that no response was required to those items. Trial types (emotional word-picture, emotional word-only, neutral word-picture, neutral word-only) were pseudorandomly intermixed with one another and with fixation crosses (+) in a way that optimized estimation of the hemodynamic response (Dale, 1999). All four seconds of each trial (e.g., word and picture or word and square) were analyzed as one event in the event-related design.

Following a one- to two-day delay, participants returned to the laboratory to perform a surprise recognition task. Debriefing indicated that no participant realized that his or her memory would be tested for the items studied in the scanner. Participants were presented with a series of 450 spoken words, presented via headphones in a different random order for each participant (taken from the American English Spoken Lexicon database: <http://www ldc.upenn.edu/cgi-bin/aesl/aesl>;

Table 1  
Possible item and response types

	Word-picture item	Word-only item	New item
Picture “viewed”	Word-picture correct attribution	Word-only misattribution	False alarm
Picture “not viewed”	Word-picture misattribution	Word-only correct attribution	Correct rejection

Note that because scanning took place at encoding, responses to “new items” cannot be analysed.

Linguistic Data Consortium, University of Pennsylvania). These methods followed those of Gonsalves and Paller (2000), with auditory presentation of stimuli decreasing the possible influence of perceptual priming on performance. 150 of the words corresponded to items that had been studied in the word-only condition; 150 to items studied in the word-picture condition; and 150 to items that had not been studied. Participants indicated whether each word corresponded to an item for which (a) a picture had been shown at study or (b) no picture had been shown (see Table 1). Thus, the same response was required for novel words that had not been studied and for words that had been presented without their corresponding photo object at study.

### 2.3. Image acquisition and data analysis

Images were acquired on a 3 T Siemens Allegra MRI scanner. Stimuli were back-projected onto a screen in the scanner bore, and participants viewed the words and objects through an angled mirror attached to the head coil. Detailed anatomic data were acquired using a multiplanar rapidly acquired gradient echo (MP-RAGE) sequence. Functional images were acquired using a T2\*-weighted echo planar imaging (EPI) sequence (TR = 2000 ms, TE = 30 ms, FOV = 200 mm; flip angle = 90°). Twenty-one axial-oblique slices (5 mm thickness, 1 mm skip between slices), aligned along the anterior commissure/posterior commissure line, were acquired in an interleaved fashion.

All pre-processing and data analysis were conducted within SPM99 (Wellcome Department of Cognitive Neurology). Standard pre-processing was performed on the functional data, including slice-timing correction, rigid body motion correction, normalization to the Montreal Neurological Institute template (re-sampling at 3 mm cubic voxels), and spatial smoothing (using an 8-mm full-width half maximum isotropic Gaussian kernel).

For each participant, and on a voxel-by-voxel basis, an event-related analysis was first conducted where all instances of a particular event type were modeled through convolution with a canonical hemodynamic response function. For the word-only trials, all participants had at least 12 instances of every event type (e.g., emotional word-only item later misattributed, neutral word-only item later correctly attributed). For the word-picture trials, two participants had too few misattributed items to analyze. The remaining 14 participants had at least nine misattributed word-

picture items. Effects for each event type were estimated using a subject-specific, fixed-effects model. These data were then entered into a second-order, random-effects analysis. To assess the effects of emotion, analyses contrasted activation during encoding of all emotional items to activation during encoding of all neutral items. Conjunction analyses examined regions that were related to subsequent memory performance regardless of emotion type (i.e., both for emotional and neutral items). These analyses were performed by creating a mask from one analysis (e.g., emotional, word-only correct attributions > emotional, word-only misattributions) and overlaying it on the comparable comparison for items of the other emotion type (e.g., neutral, word-only correct attributions > neutral, word-only misattributions). Because the two contrasts were defined independently, the results are independent of any main effect of emotion (i.e., the conjunction analysis is not affected by whether activity is greater for the emotional than the neutral items, or vice versa). While conjunction analyses detected regions that were associated with later memory performance regardless of the emotional content of the items, we also sought to identify regions that were related more strongly to subsequent memory performance for one emotion type than for the other (i.e., regions showing an interaction between emotion type and subsequent memory). To reveal these regions, we constructed interaction contrasts (as in Heckers, Zalesak, Weiss, Ditman, & Titone, 2004; Henson, Shallice, & Rugg, 2000; Lepsien & Pollmann, 2002; Taylor, Welsh, Wager, Phan, Fitzgerald, & Gehring, 2004). For example, to identify regions in which activity was more strongly related to subsequent memory performance for emotional items than for neutral items, the following contrast was performed: [emotional correct attributions – emotional misattributions] – [neutral correct attributions – neutral misattributions].

Unless otherwise noted, activation was considered reliable if the region included at least six resampled voxels, with peak activity at  $p < 0.002$  uncorrected.

All activations are presented in neurological coordinates (i.e., activity in the right hemisphere is presented on the right side of the brain images). Voxel coordinates are reported in Talairach coordinates (Talairach & Tournoux, 1998) and reflect the most significant voxel within the cluster of activation. Event-related time-courses were extracted from active clusters by creating regions-of-interest (ROI) as 8 mm spheres using the ROI toolbox implemented in SPM99. Line graphs plot these time-courses, and bar graphs plot the summation

of the signal change (from baseline) between 2 and 12 s post stimulus onset.

### 3. Results

#### 3.1. Behavioral results

An analysis of variance (ANOVA) was conducted on the recognition scores, with emotion type (emotional, neutral), response type (viewed, not viewed), and item history (word-only, word-picture, new) as within-subject factors. The ANOVA revealed a significant main effect of response type ( $F(1, 15) = 5.26, p < 0.05$ , partial eta-squared = 0.26), as well as interactions between response type and item history ( $F(1, 14) = 208.5, p < 0.0001$ , partial eta-squared = 0.97) and between response type, emotion, and item history ( $F(1, 14) = 9.84, p < 0.01$ , partial eta-squared = 0.58). Post-hoc  $t$ -tests indicated that the three-way interaction resulted because individuals were more likely to correctly say that an item was viewed when it was an emotional word-picture than a neutral word-picture ( $t(15) = 2.13, p < 0.05$ ), marginally more likely to correctly say than an item was not viewed when it was an emotional word-only item versus a neutral word-only item ( $t(15) = 1.87, p < 0.07$ ), and equally likely to correctly say that an item was not viewed if it was an emotional new item or a neutral new item ( $t(15) = 1.02, p > 0.3$ ; Table 2). Thus, emotional salience of the items had the predicted effects: individuals were more likely to correctly attribute emotional compared to neutral items.

An ANOVA conducted on the reaction times to make the size judgment about the named objects revealed no effect of response type, item history, or emotion type nor any significant interactions with emotion type. Thus, the amount of time spent imagining the object and judging its size was similar for the emotional and neutral items, and did not predict

which items would later be assigned a “viewed” versus a “not viewed” response.

#### 3.2. Neuroimaging results: effects of emotion

We first conducted a voxel-based whole-brain contrast analysis to reveal the regions that showed greater activation during viewing of emotional items as compared to neutral items. The results (Table 3) were consistent with prior studies of emotional processing: regions including the medial prefrontal cortex (PFC), orbital PFC, amygdala,<sup>1</sup> and anterior temporal lobe have frequently been found to play a role in emotional processing (reviewed by Barbas, 2000; Bechara, Damasio, & Damasio, 2000; Phan, Wager, Taylor, & Liberzon, 2002; Zald, 2003). A region of the fusiform gyrus also was disproportionately engaged during processing of emotional items, consistent with proposals that emotional processing may trigger modulation of visual-processing regions (see LeDoux, 1995; Tabert et al., 2001 for further discussion). The fusiform gyrus also has been shown to be activated in response to emotional stimuli, including facial expressions (Lewis et al., 2003; Winston, O’Doherty, & Dolan, 2003), fearful body expressions (Hadjikhani & de Gelder, 2003), and emotional photographs (Lang, Bradley, & Cuthbert, 1998; Mouras et al., 2003) and could explain the emotion related activation seen here.<sup>2</sup>

#### 3.3. Neuroimaging results: processes leading to reality-monitoring errors

The critical questions of this study centered on the neural processes that were related to the frequency of reality-monitoring errors. We first examined the neural processes that were associated with memory misattributions for the neutral items. We conducted a whole-brain, voxel-based contrast analysis, to identify regions in which activity was more strongly related to subsequent memory misattributions for the neutral items than for the emotional items (i.e., [neutral misattributions – neutral correct attributions] – [emotional misattributions – emotional correct attributions]). The only region revealed by this interaction contrast spanned the parahippocampal and fusiform gyri (BA 35 and 37; Talairach coordinates  $-20, -42, -6$ ; Fig. 2). This result is consistent

Table 2  
Effect of emotional content on reality-monitoring ability

	Picture “viewed”
Word-picture	
Emo	0.69 (0.04)
Neu	0.65 (0.04)
Word-only	
Emo	0.37 (0.05)
Neu	0.42 (0.05)
New (unstudied)	
Emo	0.20 (0.04)
Neu	0.19 (0.03)

Viewed (word-picture) items assigned a “viewed” response indicate correct attributions (mean, S.E.). Imagined (word-only) items assigned a “viewed” response reflect misattributions (mean, S.E.). Unstudied items assigned a “viewed” response signify baseline false alarms (mean, S.E.). Note that because participants assigned each item either a “viewed” or “not viewed” response, the proportions corresponding to “not viewed” are (1 – proportion “viewed”).

<sup>1</sup> Prior studies have indicated an effect of gender on the laterality of amygdalar activation, with men showing right-lateralized activity and women showing left-lateralized activity (reviewed by Cahill, 2003). Thus, we conducted post-hoc analyses to see whether there was an interaction between gender and laterality of activation in the amygdala. There was a suggestion of a gender effect: more men than women showed only right-lateralized amygdalar activation, and more women showed only left-lateralized amygdalar activation. However, effects of gender were not significant in either the left or right amygdala. Gender effects were not present in any other regions that were related to emotional processing.

<sup>2</sup> Negative and neutral stimuli were matched for the number containing people or animals, so it could not be due to global features of the stimuli such as presence of a face.

Table 3  
Regions more active during encoding of emotional items than neutral items

Region	Hemisphere	Brodmann area	Talairach coordinates (x, y, z)	Related to word-only CAs
Inferior PFC	L	47	−50, 21, −3	Emo
		45	−53, 24, 4	
		10, 11, 47	−43, 38, −8	
Medial PFC	R	9, 32	2, 42, 23	Both
	R, L	9, 10	2, 56, 8; −2, 51, 20	
Anterior temporal lobe	L	38	−28, 13, −20	
Inferior temporal lobe	R, L	21	56, −1, −24; −53, −3, −20	
Parahippocampal gyrus	R	35	21, −41, −3	
Fusiform gyrus and parahippocampal gyrus	R	36, 20	27, −26, −19	Emo
posterior hippocampus	L		−20, −28, −9	Both
Amygdala	R, L		21, −3, −21; −20, −3, −23	Emo <sup>a</sup>
Striatum	L		−12, 2, 11	Both
Cerebellum	L		−21, −41, −8	

Far right column indicates regions in which activation also was related to an increased likelihood of correct attributions (CA) for all the word-only trials (both), or specifically for emotional (emo) or neutral (neu) word-only trials.

<sup>a</sup> Activity in a nearby region of the right extended amygdala (21, −6, −9) was related to memory accuracy for the emotional items only.

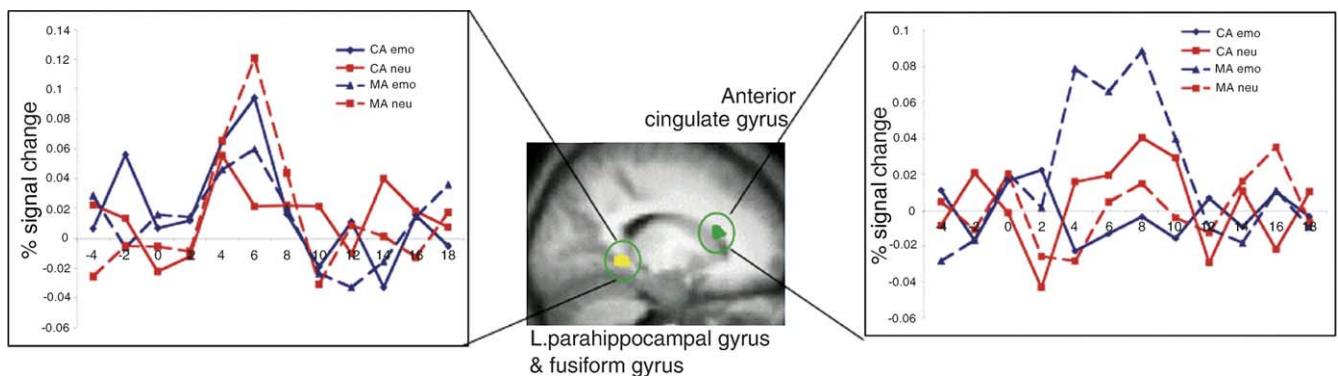


Fig. 2. The regions in which activity was related to word-only misattributions (reality-monitoring errors) differed depending on the emotional salience of the information. Activation in a region spanning the L. fusiform and parahippocampal gyri (shown in yellow; left panel) was associated with subsequent reality-monitoring errors for neutral, but not emotional items. Activation in the anterior cingulate (shown in green; right panel) was related to subsequent reality-monitoring errors for emotional, but not neutral items. CA, subsequent correct attribution; MA, subsequent misattribution.

with that of Gonsalves and colleagues (Gonsalves & Paller, 2000; Gonsalves et al., 2004), who found that activity in regions important for visual imagery increased the likelihood of subsequent reality-monitoring errors.<sup>3</sup> Activation in the fusiform gyrus has been associated with the formation of mental images in a number of investigations (e.g., Ishai, Haxby, & Ungerleider, 2002; Roland & Gulyas, 1995; Wise et al., 2000), and it is likely that, for the neutral items, individuals are later confusing these vividly imagined items for presented ones (discussed by Johnson et al., 1993).

To examine which regions were related more strongly to subsequent memory performance for the emotional items

than for the neutral items, we performed another interaction contrast analysis: [emotional misattributions – emotional correct attributions] – [neutral misattributions – neutral correct attributions]. This analysis revealed that activity in the anterior cingulate gyrus (BA 24/32; Talairach coordinates −16, 31, 12) corresponded with false memory formation for emotional words, but not for neutral words (Fig. 2). A conjunction analysis revealed that no region was related to false memory formation for both the emotional and the neutral words. Thus, distinct processes appear to support subsequent reality-monitoring errors, depending on the emotional salience of the to-be-remembered information.

### 3.4. Neuroimaging results: processes reducing the likelihood of reality-monitoring errors

We next examined whether dissociable processes also reduced the likelihood of reality-monitoring errors for neutral and emotional information. To address this question, we conducted interaction analyses to identify regions that

<sup>3</sup> The imagery-related regions that corresponded with subsequent memory misattributions in the study by Gonsalves et al. (2004) were the precuneus and inferior parietal cortex. It is not obvious why distinct regions implicated in mental imagery were revealed by the present study. Perhaps the divergent findings arose due to differences in the delay after which memory was tested (2 days in the present study versus 20 min in Gonsalves et al., 2004), the lower word frequency of the items used in the present study, or the presence of emotional items within the study list.

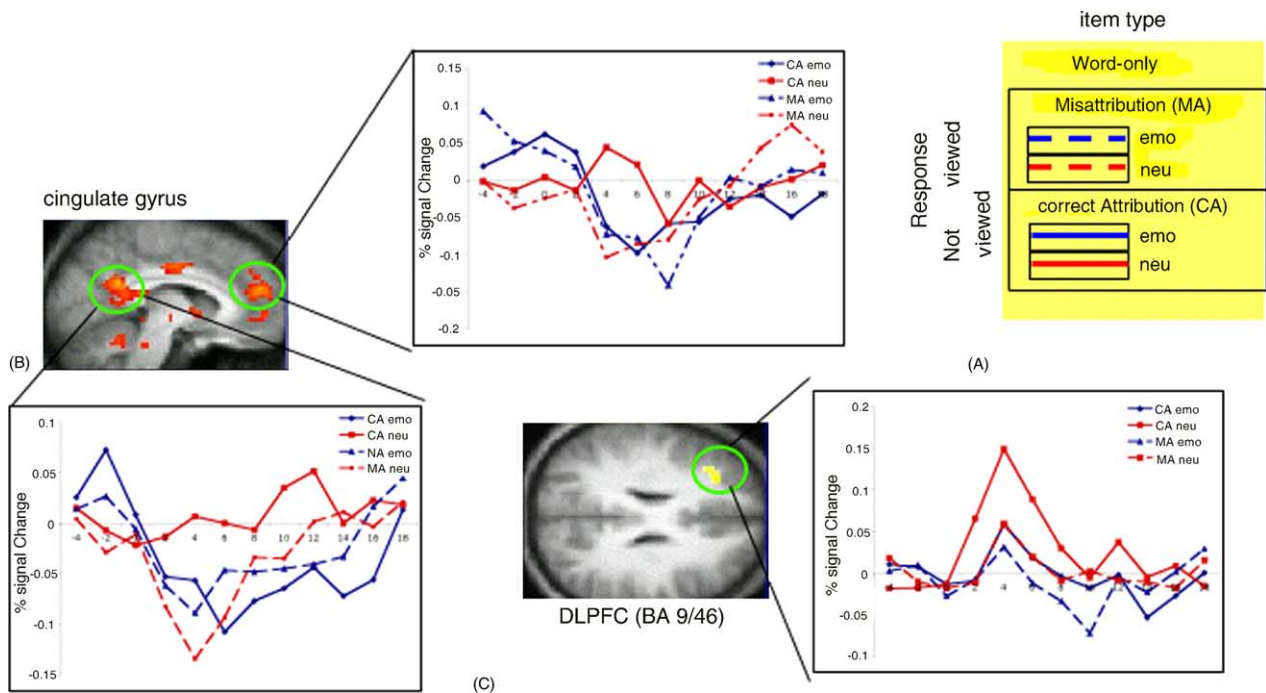


Fig. 3. Regions in which activity was associated with word-only correct attributions (CA) > word-only misattributions (MA) for neutral, but not emotional, items. Analyses were restricted to word-only trials (panel A), and data were sorted based on subsequent memory attributions. Activation along the cingulate gyrus (panel B) and in dorsolateral PFC (BA 9 and 46; panel C) was greater for later correct attributions than for later misattributions, but this pattern held only for the neutral items and not for the emotional items.

showed an interaction between emotion type and response type for the word-only trials. First, we examined whether there were regions in which the effect of response type (word-only correct attributions > word-only misattributions) was greater for the neutral items than for the emotional items (i.e., [emotional correct attributions – emotional misattributions] – [neutral correct attributions – neutral misattributions]). Activation in the left dorsolateral PFC (BA 9/46; Talairach coordinates  $-33, 36, 27$ ) and along the cingulate gyrus (anterior cingulate, BA 24/32, Talairach coordinates  $-3, 30, 15$ ; posterior cingulate, BA 23/31, Talairach coordinates  $15, -51, 24$ ) showed this pattern of results (Fig. 3). These regions have been implicated in a range of cognitive processes (reviewed by Duncan & Owen, 2000): dorsolateral PFC most commonly in working memory and executive control (reviewed by Kane & Engle, 2002; Levy & Goldman-Rakic, 2000), and the cingulate gyrus in attention modulation (Small, Gitelman, Gregory, Nobre, Parrish, & Mesulam, 2003), visuospatial processing (Vogt, Finch, & Olson, 1992) and spatial imagery (Knauff, Mulack, Kassubek, Salih, & Greenlee, 2002). Thus, additional cognitive processes carried out as individuals mentally imagined neutral objects appeared to contribute to subsequent correct attributions of the word-only items (consistent with studies suggesting that memory for cognitive operations performed can be used to assign a memory to an internal source; Johnson et al., 1993).

We then examined whether there were regions in which the effect of response type (word-only correct attribu-

tions > word-only misattributions) was larger for the emotional items than for the neutral items (i.e., [emotional correct attributions – emotional misattributions] – [neutral correct attributions – neutral misattributions]). A large subset of the regions identified as important for emotional processing were also identified in this interaction contrast analysis (far right column of Table 3). Time-course analyses confirmed that these regions were related to memory accuracy (word-only correct attributions > word-only misattributions) for the emotional items, but not for the neutral items (Fig. 4). In fact, for the neutral items, the pattern of activation sometimes was in the opposite direction, with activation for subsequent misattributions nonsignificantly greater than activation for subsequent correct attributions (e.g., Fig. 4, panel B). These results provide strong evidence for encoding processes that reduce memory misattributions specifically for emotional items, and not for neutral items. Further, the extensive overlap of the regions relating to memory accuracy for emotional items and those implicated more broadly in emotional processing (Table 3; see also Bechara et al., 2000; Phan et al., 2002; Zald, 2003) suggests that the way in which emotional information is processed tends to confer a reality-monitoring advantage. Individuals may orient attention toward aspects of emotional events that will later allow correct assignment to internal or external sources.

Although these results provide evidence that separable neural processes are related to memory accuracy for emotional and neutral items, we also were interested in examining

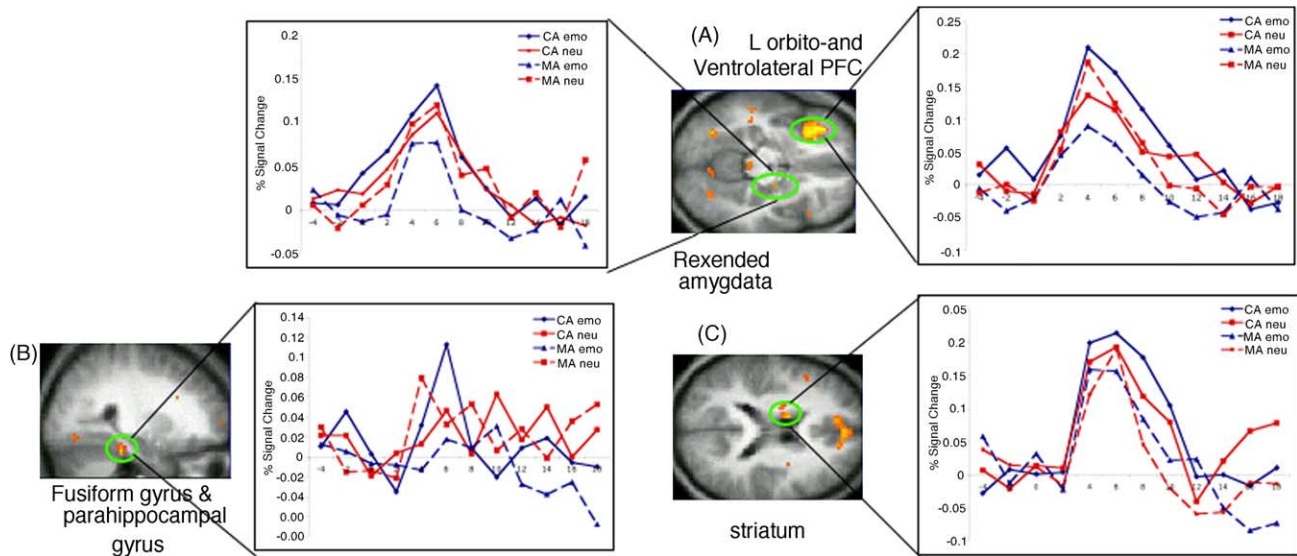


Fig. 4. Regions in which activity was related to word-only correct attributions (CA) > word-only misattributions (MA) for emotional, but not neutral, items. Encoding-related activity in the extended amygdala and orbitofrontal cortex (panel A), fusiform/parahippocampal gyrus (panel B), and striatum (panel C) was linked to subsequent correct attributions only for the emotional items. These regions also were found to be more active during processing of emotional as compared to neutral items (see Table 3).

what processes might be associated with accurate memory assignment regardless of emotional content. Thus, we performed a conjunction analysis to identify regions in which activity was greater for items later correctly attributed than for those later misattributed, regardless of the items' emotional content (i.e., for both emotional and neutral word-only trials). We found that activation in the right inferior PFC and the left posterior hippocampus was related to subsequent correct attributions for both item types (Fig. 5). These regions closely parallel the network previously implicated in subsequent memory analyses (Reber et al., 2002; Weis, Klaver, Reul, Elger, & Fernandez, 2004; reviewed by Paller & Wagner, 2002). In addition, activity in the medial PFC also was associated with correct attributions for emotional and neutral items. This region has been implicated in a variety of processes, including tactile imagery (Yoo, Freeman, McCarthy, & Jolesz, 2003) and self-referential processing (Gusnard, Akbudak, Shulman, & Raichle, 2001). These processes were likely engaged for all items as participants imagined manipulating the objects to determine if they would fit into a shoebox. The time-course of activity in this medial PFC region—showing decreased activity compared to baseline—is consistent with that of prior studies, and likely reflects the region's role in self-referential processing, which is likely to occur even during the baseline events (see Gusnard et al., 2001, for further discussion).

### 3.5. Characterizing regional responses: memory accuracy or memory assignment?

An interesting question regarding the regions identified in the above comparisons from word-only trials was whether

they were related specifically to subsequent memory accuracy (correct attributions > misattributions), or to subsequent memory assignment (e.g., later calling an item "not viewed" versus "viewed," regardless of its actual encoding condition). To address this question, we examined how these regions responded to the viewed items that were later correctly attributed and to those that were later misattributed. Because participants tended to have low rates of misattribution for the viewed items, these analyses were restricted to the 14 participants (eight women) who had at least nine misattributed word-picture trials of each emotion type.

### 3.6. Regions related to word-only misattributions

#### 3.6.1. Neutral items

We started by examining activity in the left fusiform region (BA 35 and 37; Talairach coordinates  $-20, -42, -6$ ) that had been revealed in the interaction contrast of [neutral word-only misattributions – neutral word-only correct attributions] – [emotional word-only misattributions – emotional word-only correct attributions]. Consistent with Gonsalves and colleagues' (2000, 2004) hypotheses, this region showed greater activity for the word-picture trials that were later correctly attributed to visual presentation versus those that were later misattributed. This pattern held for both emotional and neutral items (Fig. 6).

In addition, a conjunction analysis conducted on the viewed items (correct attributions > misattributions for both emotional and neutral items) revealed a nearby region of the left fusiform/parahippocampal gyrus (BA 35 and 36; Talairach coordinates  $-27, -31, 18$ ) as the only region that was associated with accurate word-picture memory attribution for



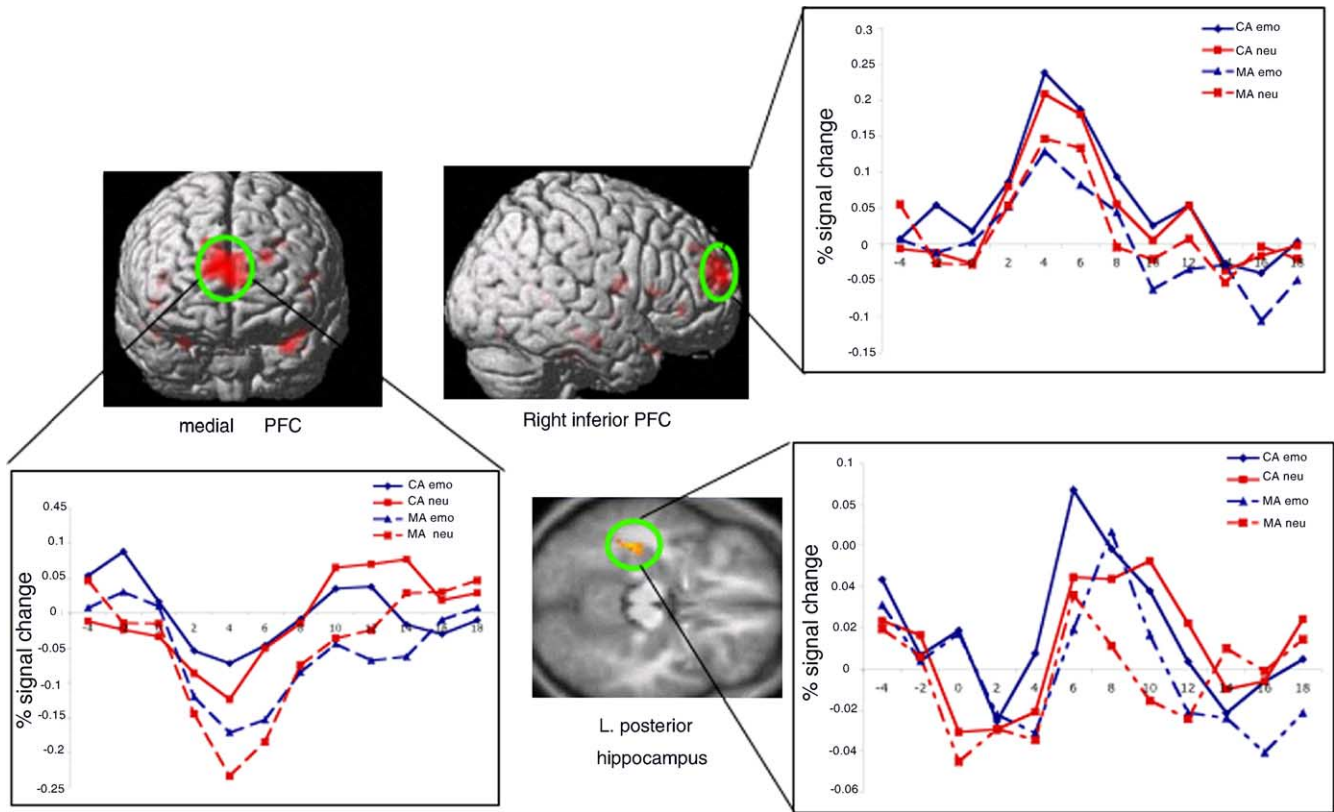


Fig. 5. Regions in which activity was greater for word-only correct attributions (CA) than word-only misattributions (MA) for emotional and neutral items. Conjunction analyses revealed that activations in the medial PFC, right inferior PFC (BA 45), and left posterior hippocampus were related to memory accuracy for both emotional and neutral items.

both emotional and neutral items. These results support the hypothesis that greater activity in visual regions during encoding of word-picture trials leads to assignment of a memory to an externally presented source. Interestingly, in both of these fusiform regions, greater activity during word-only trials also led to later misattribution to picture presentation for the neutral items. However, no such relation existed for the emotional items.

### 3.7. Emotional items

We then examined the region of the anterior cingulate gyrus (BA 24/32; Talairach coordinates  $-16, 31, 12$ ) that had been identified in the interaction contrast comparing [emotional word-only misattributions – emotional word-only correct attributions] – [neutral word-only misattributions – neutral word-only correct attributions]. This region

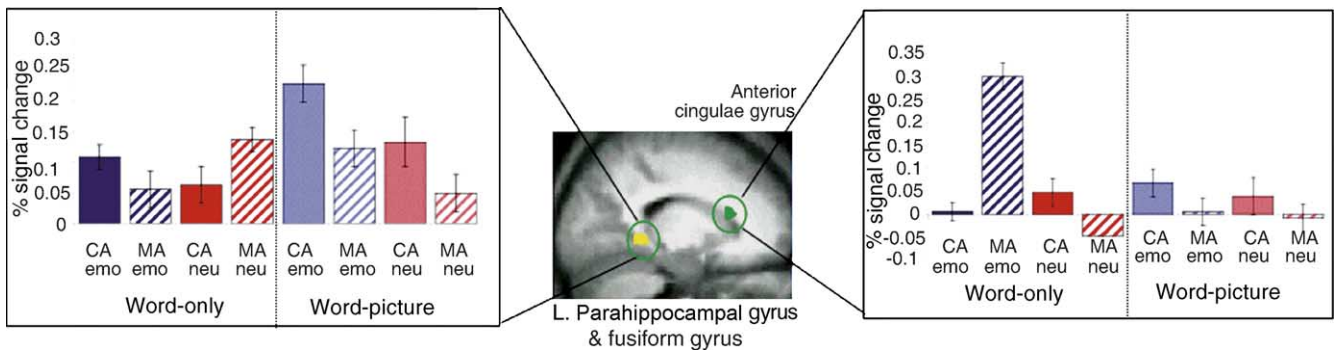


Fig. 6. Regions in which activity was related to word-only misattributions (MA): memory accuracy vs. memory assignment. For neutral items, activity in the fusiform gyrus was associated with later assignment of an item as “viewed,” regardless of its actual study history. For emotional items, fusiform activity led to correct attributions (CA) of word-picture trials as being “viewed,” but did not lead to misattributions for the word-only items. Activity in the anterior cingulate gyrus showed no relation to memory attribution for the word-picture items.

showed no relation to subsequent memory for the viewed items (Fig. 6). Although activity in the anterior cingulate gyrus occurs across a range of tasks requiring cognitive control processes (e.g., Ochsner et al., 2001; reviewed by Carter, Botvinick, & Cohen, 1999), a number of studies have implicated the region in motor and visual imagery (e.g., Decety, 1996; Gulyas, 2001). Thus, it is plausible that activity in this region is related to vivid mental imagery generation. In contrast to the pattern of activity in the fusiform region, however, activity in the anterior cingulate gyrus was not strongly associated with perception of the visual objects. This pattern is consistent with the finding that increased activity in the anterior cingulate gyrus was associated with an increased likelihood of falsely remembering that a photo had been presented when it had not been studied, but was not associated with an increased ability to remember a photo that had been viewed.

### 3.8. Regions related to word-only correct attributions

#### 3.8.1. Neutral items

Of the regions revealed in the interaction contrast analysis to be more closely related to accurate memory assignment for the neutral word-only items than for the emotional word-only items (i.e., left dorsolateral PFC and cingulate gyrus), only dorsolateral PFC activity was related to later memory for the word-picture trials. In particular, this region (BA 9/46; Talairach coordinates  $-33, 36, 27$ ) showed greater activity for word-picture trials later misattributed than for those later correctly attributed (Fig. 7). Thus, this region was associated with later memory assignment: greater activity led a participant to later respond that a picture had been “not viewed,” regardless of the item’s actual study condition. This pattern is consistent with the hypothesis that this dorsolateral PFC activity reflected the cognitive operations performed during mental imagery: participants may have later retrieved the details of these cognitive operations, and thus attributed any visual information they remembered to mental imagery rather than to visual presentation (see Johnson et al., 1993, for further discussion). Interestingly, a conjunction analysis of [neutral word-picture misattributions > neutral word-picture correct attributions] and [emotional word-picture misattributions > emotional word-picture correct attributions] revealed that activity in a dorsolateral PFC region, contralateral to that discussed above (BA 9/46; Talairach coordinates  $33, 27, 24$ ), led to a later response of “not viewed” for both emotional and neutral items.<sup>4</sup>

<sup>4</sup> It also is plausible that this pattern of activity (subsequent “not viewed” response at retrieval) reflects subsequent forgetting. That is, a “not viewed” response was given both to novel items that were never studied and to items that were imagined, but for which a picture was never studied. Given that the regions shown to relate to subsequent “not viewed” responses do not parallel those observed in other investigations of subsequent forgetting (reviewed in Wagner & Davachi, 2001), we are hesitant to believe that this explanation could account for the data. Nevertheless, future investigations will be required to distinguish these possibilities.

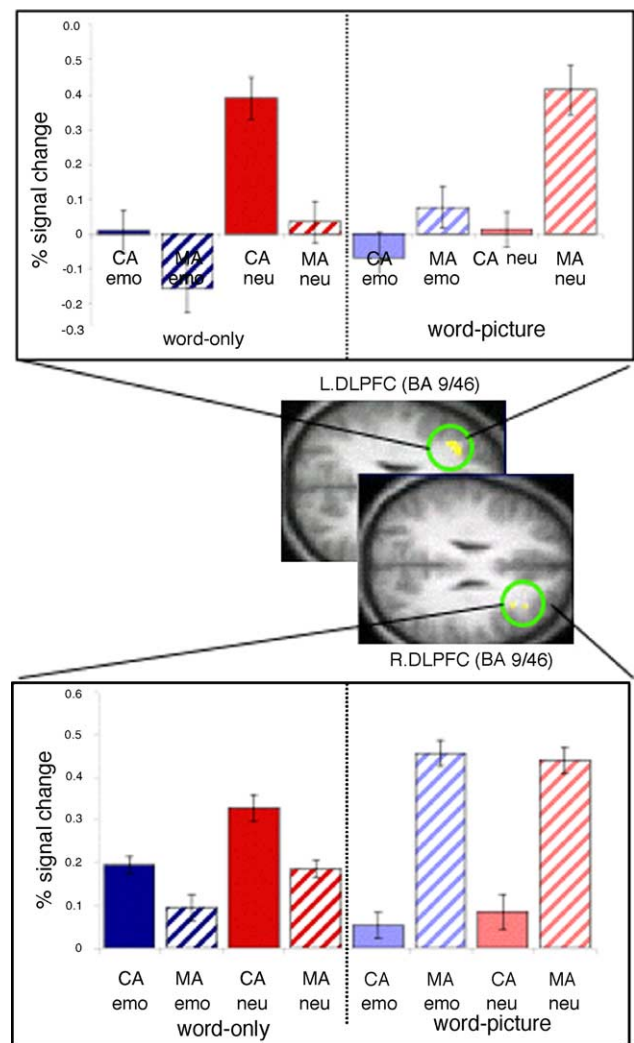


Fig. 7. Activity in the DLPFC predicted later assignment of an item as “not viewed.” Activity in the left dorsolateral PFC (DLPFC) was related to a later assignment of “not viewed” for all neutral items, regardless of their study history. This region showed no relation to later memory for the emotional items. In contrast, activity in the right DLPFC was related to a later assignment of “not viewed” for both emotional and neutral items.

#### 3.8.2. Emotional items

Of the regions revealed by the interaction contrast analysis to be more closely related to accurate memory assignment for the emotional items than for the neutral items, activity in the amygdala and the left ventrolateral/orbital PFC was associated with accurate memory attribution for the emotional word-picture items (but not the neutral word-picture items; Fig. 8). Activity was greater during encoding of emotional word-picture trials that would later be given a correct (“viewed”) response versus an incorrect (“not viewed”) response.

The region spanning the right fusiform and parahippocampal gyrus showed no relation to memory assignment for the emotional word-picture trials, but it was associated with assignment to visual presentation for the neutral word-picture items.

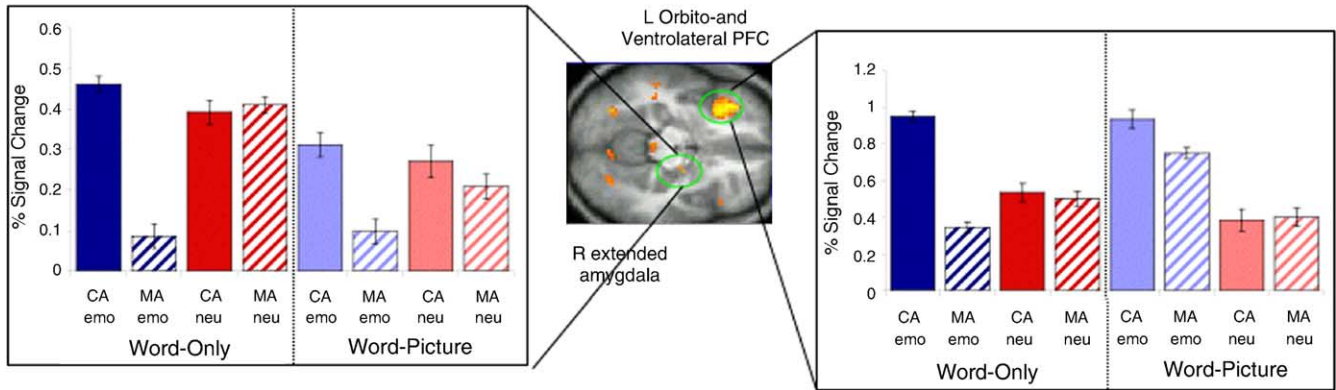


Fig. 8. Activity in the amygdala and ventrolateral/orbitofrontal PFC was associated with memory accuracy for the emotional items. Activity in these regions reduced the likelihood of later misattributions, both for word-only and for word-picture emotional trials. The regions showed no relation to memory accuracy for the emotional items.

3.8.3. Both neutral and emotional items

With regard to the regions identified in the conjunction analysis as corresponding to accurate memory attribution for both item types, only the hippocampus also was related to accurate memory attribution for the word-picture items. Activity in this region was greater during the encoding of word-picture trials that would later be correctly attributed to external presentation versus during the encoding of word-picture trials for which individuals would later indicate that no picture had been viewed (Fig. 9). This relation held for both the emotional and the neutral items.

Activity in the medial PFC, in contrast, showed a relation to memory assignment for the emotional, but not the neutral, items: increased activity was related to the tendency to assign an emotional item from a word-picture trial a “not viewed” response at retrieval. It may be that individuals are more likely to carry out self-referential imagery for the emotional than the neutral items (as evidenced by the overall increased medial PFC activity for emotional than neutral items), and then later use retrieval of this information to assign a memory to a nonpresented source.

Activity in the right inferior PFC did not show a significant relation to later memory assignment for the neutral or emotional items.

3.8.4. Amygdala–hippocampal interactions

Prior studies have indicated that interactions between the amygdala and the hippocampus are critical for emotional memory enhancement effects (e.g., Dolcos et al., 2003; Kensinger & Corkin, 2004a; Richardson et al., 2004). To examine whether these interactions also were important in leading to accurate memory attributions, we correlated activity in the functional regions of the amygdala (Talairach coordinates: 21, -1, -21; -20, -3, -23) and the posterior hippocampus (Talairach coordinates: -20, -28, -9), identified in the earlier contrast analyses.

The critical finding from these correlations was that activity in the amygdala and the hippocampus was correlated when individuals were encoding emotional items that were later correctly attributed ( $r = 0.89, p < 0.01$  for word-only trials in which emotional items were later correctly attributed;  $r = 0.81, p < 0.01$  for word-picture trials in which emotional

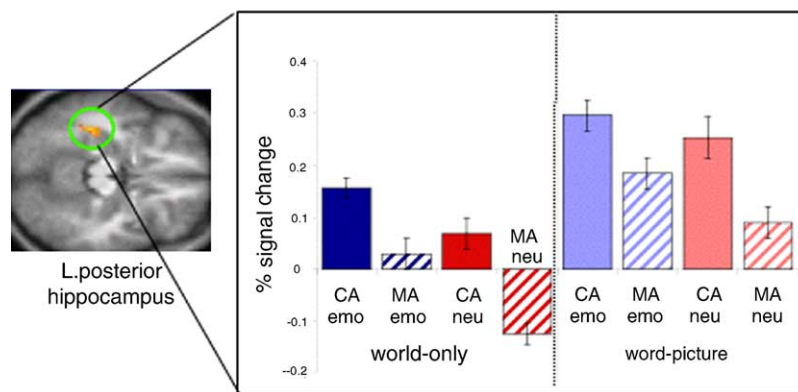


Fig. 9. Activity in the hippocampus was related to memory accuracy for both emotional and neutral items. Increased activity in this region was associated with later correct attributions for all items (emotional and neutral, word-only and word-picture).

items were later correctly attributed), but not when they were encoding any other item types. Fisher's  $z$ -test confirmed that the correlation for the emotional items later correctly attributed was significantly greater than the correlation for emotional items later incorrectly attributed, or for neutral items (all  $z > 1.97$ ,  $p < 0.05$ ).

#### 4. Discussion

By showing participants only the verbal referents of some objects, and the verbal label and corresponding image of other objects, we induced reality-monitoring errors in all of our participants. Below, we outline the ways in which the results of this study extend our knowledge about the processes contributing to reality-monitoring errors for neutral items, and the ways in which an item's emotional salience affects those processes.

##### 4.1. Reality-monitoring errors for neutral items

We found that encoding-related activation spanning the fusiform and parahippocampal gyri (likely a marker of robust or vivid mental imagery; Ishai et al., 2002; Kosslyn, Thompson, Costantini-Ferrando, Alpert, & Spiegel, 2000; Roland & Gulyas, 1995; Wise et al., 2000) enhanced the likelihood of reality-monitoring errors for neutral items. This finding converges with that of Gonsalves et al. (2004), who found that activity in regions implicated in mental imagery (precuneus and inferior parietal cortex) increased the probability of later making reality-monitoring errors for neutral items. Presumably, individuals later retrieved the visual information generated during mental imagery, but believed that it had come from visual presentation of the item (as discussed by Johnson et al., 1993; Johnson & Raye, 1981).

By demonstrating that encoding-related activity can be related to a decreased (as well as an increased) likelihood of reality-monitoring errors, our results also extend the results of Gonsalves and colleagues. Activity in the dorsolateral PFC and the left posterior hippocampus showed this relation: greater activity during encoding of word-only items which were later correctly attributed than during encoding of word-only items which were later misattributed. Importantly, however, the reason for that relation appears to have differed for the two regions. Dorsolateral PFC activity led an individual to indicate that an item's photo was "not viewed," regardless of whether or not the photo actually had been shown. It is plausible that dorsolateral PFC activity at encoding was related to cognitive operations performed during mental imagery. It is logical that at retrieval, participants later remembered those cognitive operations, and thus attributed any memory for visual information to imagination rather than to external presentation (as proposed by Johnson et al., 1993). This pattern of activity, taken together with the pattern in the fusiform gyrus, would support the hypothesis that individuals are likely to make reality-monitoring errors when the characteristics of

memory from an internal source are most similar to those from an external source (e.g., contain vivid perceptual detail and lack information about cognitive operations performed; Johnson et al., 1993; Johnson & Raye, 1981).

In contrast to the dorsolateral PFC activity, hippocampal activity at encoding was related to accurate memory attribution, not simply to assignment of an item as being "not viewed." Thus, increased hippocampal activity decreased the probability of reality-monitoring errors (i.e., word-only misattributions), but also reduced the likelihood of word-picture misattributions (i.e., indicating that an object from a word-picture trial had been "not viewed"). This finding is consistent with the hippocampus' role in the encoding of source information (e.g., Davachi et al., 2003), and is also consistent with Gonsalves et al. (2004) finding that activity in the hippocampus reduced the likelihood of word-picture misattributions.

These results demonstrate that not only are distinct neural signatures associated with encoding processes that lead to remembering versus forgetting (reviewed by Paller & Wagner, 2002), but also there are neural processes engaged at encoding that predict whether information will be remembered in an accurate or distorted fashion (see also Gonsalves & Paller, 2000; Gonsalves et al., 2004).

##### 4.2. Effects of emotion on reality-monitoring errors

While this study provides novel evidence regarding the encoding processes that lead to reality-monitoring errors for neutral items, the main goal of this study was to examine whether those processes differed depending on the emotional nature of the items presented at study. As expected, memory misattributions were less frequent for items with emotional salience than for the neutral items (see also Kensinger & Corkin, 2004b; Kensinger & Schacter, unpublished observations; Pesta et al., 2001, for evidence of reduced memory distortion for emotional information). The critical question, therefore, was whether this enhanced accuracy for emotional items resulted from differences in the processes engaged during the encoding of the emotional items.

The processes that were associated with an increased likelihood of reality-monitoring errors were distinct for the neutral and the emotional items, and there were no regions that showed this relation for both item types. Fusiform activity was not related to an increased likelihood of reality-monitoring errors for emotional items, as it was for neutral items. In fact, if anything, the pattern of activity was in the opposite direction (i.e., activity decreasing the likelihood of false memory formation for emotional items). Instead, activity in the anterior cingulate gyrus was associated with subsequent reality-monitoring errors for the emotional items.<sup>5</sup> Further investigations will be required to elucidate why the emotional salience of items modulated

<sup>5</sup> It is interesting that Gonsalves et al. (2004) found that activity in the anterior cingulate increased the probability of later reality-monitoring errors for neutral items.

the processes associated with an increased probability of reality-monitoring errors: it may be that the characteristics of the constructed mental images differed for the two item types; or it may be that interactions between limbic regions and visual-processing regions (e.g., fusiform gyrus) altered the relation of visual activity to memory misattributions<sup>6</sup> (see LeDoux, 1995; Tabert et al., 2001).

In terms of decreasing the likelihood of reality-monitoring errors, the hippocampus and right dorsolateral PFC showed identical patterns of activity for the emotional word-only trials as they did for the neutral word-only trials. Both regions showed greater activity during the encoding of word-only trials later correctly attributed than during the encoding of word-only trials later misattributed. The activity in the hippocampus appeared to correspond with subsequent memory accuracy, because activity also was greater for emotional word-picture trials which were later correctly attributed than for those which were later misattributed. In contrast, activity in the right dorsolateral PFC led to a later judgment of an item as “not viewed,” regardless of its study history. These results are important in indicating that there is overlap between the encoding processes that predict later memory judgments for both emotional and neutral items.

While there were these regions of overlap, there also were important points of divergence, both in the regions linked to later memory assignment (“not viewed” versus “viewed”) and in the regions that were related to accurate memory attribution. In terms of later memory assignment for emotional and neutral items, activity in the left dorsolateral PFC was associated with a subsequent judgment of “not viewed” for the neutral items, but did not show this relation for the emotional items. In contrast, activity in the medial PFC led to a later judgment of “not viewed” for the emotional items, but not for the neutral items. Further investigation will be necessary to elucidate the reasons for this dissociation. It is plausible that differences in the cognitive operations recruited to construct the mental images can explain the separable patterns.

The results also highlighted a dissociation in the processes that were related to accurate memory attribution for emotional and neutral items. Emotional information was less likely to be misattributed when regions implicated in emotional processing (in Table 3, and also by Barbas, 2000; Bechara et al., 2000; Phan et al., 2002; Zald, 2003) were engaged during encoding: activity in the left ventrolateral/orbital PFC and in the amygdala was greater for emotional word-only trials that were later correctly attributed than it was for those later misattributed. Because both of these regions showed greater activity during correctly attributed

word-picture trials than during misattributed word-picture trials, encoding-related activity in these regions appears to correspond with accurate memory attribution for emotional items. Activation in these emotion-processing regions, in contrast, was unrelated to memory accuracy for the neutral items.

It has been debated to what extent emotional information is better-remembered due to the same types of processes (e.g., increased elaboration or rehearsal) that benefit memory for neutral information (Christianson & Engelberg, 1999; Guy & Cahill, 1999; Hamann, 2001). The data reported here suggest that, while there are some encoding processes that boost the likelihood of accurate memory attribution for all items, there are also encoding processes that are related to accurate memory attribution only for the emotional items.

The correlation between the amygdala and hippocampus suggests that at least part of the memory advantage for emotional items is due to interactions between regions important specifically for emotion processing and those that more generally support encoding (of nonemotional as well as emotional information). Thus, amygdalar–hippocampal interactions may be important not only for increasing the likelihood that emotional information is remembered (e.g., Hamann et al., 1999; Kensinger & Corkin, 2004a; Richardson et al., 2004) but also for decreasing the likelihood that memories for emotional information will be misattributed (e.g., to external presentation versus internal generation).<sup>7</sup>

Given that reality-monitoring performance for emotional items exceeded that for nonemotional items, it may also be of interest to examine brain activity during encoding under conditions in which reality-monitoring performance for nonemotional items is matched to that of the emotional items (perhaps by using distinctive but nonemotional words). If, as we contend, the effects observed here are attributable specifically to the emotional nature of the material, then matching reality-monitoring performance for emotional and nonemotional words should produce the same patterns of results. If the effects are attributable to differences in levels of performance, then under conditions of matched performance, brain activation for emotional and nonemotional items should be indistinguishable.

From the perspective of the emotion and memory literature, the results of the present study nicely extend previous research indicating that regions specifically involved in emotional processing are associated with subsequent remembering of emotional information (e.g., Cahill et al., 1996). From the perspective of the memory distortion literature, our data provide some of the strongest evidence to-date that reduction in susceptibility to false memories need not be the result of

<sup>6</sup> The data do not seem likely to be explained by a bias for participants to not image the emotional objects as vividly or for as long as the neutral objects: on written debriefing forms, no participant reported difficulty imagining the emotional objects nor a failure to comply with the encoding task instructions. Further, the overall activity in the fusiform gyrus was similar for the emotional and the neutral word-only items (and, in fact, was greater for emotional than neutral items in the right fusiform; Table 1).

<sup>7</sup> The present study design does not allow examination of the overlap between processes contributing to item memory (i.e., memory that an item was presented, but no memory for its encoding condition) and contextual memory (i.e., memory for the item’s encoding condition). Therefore, future studies will be required to examine whether these processes mediate enhanced item memory for emotional items as well as enhanced contextual memory.

a general or global improvement in memory performance. Rather, the finding that activation in emotion-processing regions was related to memory accuracy specifically for emotional but not for neutral items indicates that reduction in memory misattributions may reflect the operation of domain-specific processes that have specific effects on a particular type of processing. Future studies would do well to explore the distinction between domain-general and domain-specific effects in false memory reduction suggested by the findings reported here.

## 5. Conclusions

In summary, the results of the present study indicate that, for both emotional and neutral items, encoding processes correlate with whether memories later will be accurately or inaccurately attributed. Participants made fewer misattribution errors for emotional than for neutral items, and the neuroimaging data indicated that while hippocampal activity reduced the likelihood of memory misattributions for all items, activity in the amygdala and the orbitofrontal cortex reduced the likelihood of memory misattributions specifically for emotional, but not for neutral, items. At least part of the role of these emotion-specific regions likely resulted from their interactions with regions that promote accurate encoding of both emotional and nonemotional items (as suggested by the significant amygdala–hippocampal correlation for emotional items later correctly attributed). Thus, just as distinct processes contribute to the likelihood of remembering emotional and neutral items, so do distinct processes modulate reality-monitoring accuracy for emotional and neutral information.

## Acknowledgements

We thank Alex Salskov for help with participant recruitment and testing and Rachel Garoff and Scott Slotnick for helpful discussion. This research was supported by NIH grants MH60941 (to D.L.S.) and MH070199 (to E.A.K.), and a Massachusetts Biomedical Research Corporation Tosteson Postdoctoral Fellowship (to E.A.K.).

## References

- Barbas, H. (2000). Connections underlying the synthesis of cognition, memory, and emotion in primate prefrontal cortices. *Brain Research Bulletin*, *52*, 319–330.
- Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*, 295–307.
- Budson, A. E., Daffner, K. R., Desikan, R., & Schacter, D. L. (2000). When false recognition is unopposed by true recognition: Gist-based memory distortion in Alzheimer's disease. *Neuropsychology*, *14*, 277–287.
- Cahill, L. (2003). Sex- and hemisphere-related influences on the neurobiology of emotionally influenced memory. *Progress in Neuropharmacology and Biological Psychiatry*, *27*, 1235–1241.
- Cahill, L., Haier, R. J., Fallon, J., Alkire, M., Tang, C., Keator, D., et al. (1996). Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proceedings of the National Academy of Sciences of the United States of America*, *93*, 8016–8021.
- Canli, T., Zhao, Z., Brewer, J., Gabrieli, J. D., & Cahill, L. (2000). Event-related activation in the human amygdala associates with later memory for individual emotional experience. *Journal of Neuroscience*, *20*, RC99.
- Carter, C. S., Botvinick, M. M., & Cohen, J. D. (1999). The contribution of the anterior cingulate cortex to executive processes in cognition. *Reviews in Neuroscience*, *10*, 49–57.
- Christianson, S. A., & Engelberg, E. (1999). Organization of emotional memories. In T. Dalgleish & M. Power (Eds.), *Handbook of cognition and emotion*. New York: John Wiley & Sons.
- Colthart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, *33A*, 497–505.
- Dale, A. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, *8*, 109–114.
- Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 2157–2162.
- Decety, J. (1996). Do imagined and executed actions share the same neural substrate? *Brain Research: Cognitive Brain Research*, *3*, 87–93.
- Dolcos, F., Graham, R., LaBar, K., & Cabeza, R. (2003). Coactivation of the amygdala and hippocampus predicts better recall for emotional than for neutral pictures. *Brain and Cognition*, *51*, 221–223.
- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neuroscience*, *23*, 475–483.
- Gilhooly, K. J., & Logie, R. H. (1980). Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1944 words. *Behavior Research Methods and Instrumentation*, *12*, 395–427.
- Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 5–16.
- Gonsalves, B., & Paller, K. A. (2000). Neural events that underlie remembering something that never happened. *Nature Neuroscience*, *3*, 1316–1321.
- Gonsalves, B., Reber, P. J., Gitelman, D. R., Parrish, T. B., Mesulam, M.-M., & Paller, K. A. (2004). Neural evidence that vivid imagining can lead to false remembering. *Psychological Science*, *15*, 655–660.
- Gulyas, B. (2001). Neural networks for internal reading and visual imagery of reading: A PET study. *Brain Research Bulletin*, *54*, 319–328.
- Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 4259–4264.
- Guy, S. C., & Cahill, L. (1999). The role of overt rehearsal in enhanced conscious memory for emotional events. *Consciousness and Cognition*, *8*, 114–122.
- Hadjikhani, N., & de Gelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and amygdala. *Current Biology*, *13*, 2201–2205.
- Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Science*, *5*, 394–400.
- Hamann, S. B., Ely, T. D., Grafton, S. T., & Kilts, C. D. (1999). Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience*, *2*, 289–293.
- Heckers, S., Zalesak, M., Weiss, A. P., Ditman, T., & Titone, D. (2004). Hippocampal activation during transitive inference in humans. *Hippocampus*, *14*, 153–162.
- Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science*, *287*, 1269–1272.
- Ishai, A., Haxby, J. V., & Ungerleider, L. G. (2002). Visual imagery of famous faces: Effects of memory and attention revealed by fMRI. *Neuroimage*, *17*, 1729–1741.

- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*, 67–85.
- Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin and Review*, *9*, 637–671.
- Kensinger, E. A., & Corkin, S. (2004a). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 3310–3315.
- Kensinger, E. A., & Corkin, S. (2004b). The effects of emotional content and aging on false memories. *Cognitive, Affective, and Behavioral Neuroscience*, *4*, 1–9.
- Kensinger, E. A., & Schacter, D. L. (1999). When true memories suppress false memories: Effects of aging. *Cognitive Neuropsychology*, *16*, 399–415.
- Kensinger, E. A. & Schacter, D. L. (unpublished observations). Reality monitoring and memory distortion: Effects of emotion.
- Knauff, M., Mulack, T., Kassubek, J., Salih, H. R., & Greenlee, M. W. (2002). Spatial imagery in deductive reasoning: A functional MRI study. *Brain Research: Cognitive Brain Research*, *13*, 203–212.
- Kosslyn, S. M., Thompson, W. L., Costantini-Ferrando, M. F., Alpert, N. M., & Spiegel, D. (2000). Hypnotic visual illusion alters color processing in the brain. *American Journal of Psychiatry*, *157*, 1279–1284.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry*, *44*, 1248–1263.
- LeDoux, J. E. (1995). Emotion: Clues from the brain. *Annual Review of Psychology*, *46*, 209–235.
- Lepsien, J., & Pollmann, S. (2002). Covert reorienting and inhibition of return: An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*, 127–144.
- Levy, R., & Goldman-Rakic, P. S. (2000). Segregation of working memory functions within the dorsolateral prefrontal cortex. *Experimental Brain Research*, *133*, 23–32.
- Lewis, S. R. J., Lanoue, M. D., Miller, G. A., Heller, W., Edgar, C., Huang, M., et al. (2003). Visual processing of facial affect. *Neuroreport*, *14*, 1841–1845.
- McGaugh, J. L. (2000). Memory—a century of consolidation. *Science*, *287*, 248–251.
- Mouras, H., Stoleru, S., Bittoun, J., Glutron, D., Pelegrini-Issac, M., Paradis, A. L., et al. (2003). Brain processing of visual sexual stimuli in healthy men: A functional magnetic resonance imaging study. *Neuroimage*, *20*, 855–869.
- Ochsner, K. N., Kosslyn, S. M., Cosgrove, G. R., Cassem, E. H., Price, B. H., Nierenberg, A. A., et al. (2001). Deficits in visual cognition and attention following bilateral anterior cingulotomy. *Neuropsychologia*, *39*, 219–230.
- Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery and meaningfulness values for 925 words. *Journal of Experimental Psychology Monograph Supplement*, *76*, 1–20.
- Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Science*, *6*, 93–102.
- Pesta, B. J., Murphy, M. D., & Sanders, R. E. (2001). Are emotionally charged lures immune to false memory? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 328–338.
- Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*, *16*, 331–348.
- Phelps, E. A. (2004). Human emotion and memory: Interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology*, *14*, 198–202.
- Reber, P. J., Siwec, R. M., Gitelman, D. R., Parrish, T. B., Mesulam, M. M., Paller, K. A., et al. (2002). Neural correlates of successful encoding identified using functional magnetic resonance imaging. *Journal of Neuroscience*, *22*, 9541–9548.
- Richardson, M. P., Strange, B. A., & Dolan, R. J. (2004). Encoding of emotional memories depends on amygdala and hippocampus and their interactions. *Nature Neuroscience*, *7*, 278–285.
- Roland, P. E., & Gulyas, B. (1995). Visual memory, visual imagery, and visual recognition of large field patterns by the human brain: Functional anatomy by positron emission tomography. *Cerebral Cortex*, *5*, 79–93.
- Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory and Language*, *40*, 1–24.
- Schacter, D. L., & Slotnick, S. D. (2004). The cognitive neuroscience of memory distortion. *Neuron*, *44*, 149–160.
- Schacter, D. L., Verfaellie, M., Anes, M. D., & Racine, C. (1998). When true recognition suppresses false recognition: Evidence from amnesic patients. *Journal of Cognitive Neuroscience*, *10*, 668–679.
- Small, D. M., Gitelman, D. R., Gregory, M. D., Nobre, A. C., Parrish, T. B., & Mesulam, M. M. (2003). The posterior cingulate and medial prefrontal cortex mediate the anticipatory allocation of spatial attention. *Neuroimage*, *18*, 633–641.
- Tabert, M. H., Borod, J. C., Tang, C. Y., Lange, G., Wei, T. C., Johnson, R., et al. (2001). Differential amygdala activation during emotional decision and recognition memory tasks using unpleasant words: An fMRI study. *Neuropsychologia*, *39*, 556–573.
- Talairach, J., & Tournoux, P. (1998). *Co-planar stereotaxic axis of the human brain*. New York: Thieme.
- Taylor, S. F., Welsh, R. C., Wager, T. D., Phan, K. L., Fitzgerald, K. D., & Gehring, W. J. (2004). A functional neuroimaging study of motivation and executive function. *Neuroimage*, *21*, 1045–1054.
- Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. New York: Erlbaum.
- Vogt, B. A., Finch, D. M., & Olson, C. R. (1992). Functional heterogeneity in cingulate cortex: The anterior executive and posterior evaluative regions. *Cerebral Cortex*, *2*, 435–443.
- Wagner, A. D., & Davachi, L. (2001). Cognitive neuroscience: Forgetting of things past. *Current Biology*, *11*, R964–R967.
- Weis, S., Klaver, P., Reul, J., Elger, C. E., & Fernandez, G. (2004). Temporal and cerebellar brain regions that support both declarative memory formation and retrieval. *Cerebral Cortex*, *14*, 256–267.
- Winston, J. S., O’Doherty, J., & Dolan, R. J. (2003). Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *Neuroimage*, *20*, 84–97.
- Wise, R. J., Howard, D., Mummery, C. J., Fletcher, P., Leff, A., Buchel, C., et al. (2000). Noun imageability and the temporal lobes. *Neuropsychologia*, *38*, 985–994.
- Yoo, S. S., Freeman, D. K., McCarthy, J. J., 3rd, & Jolesz, F. A. (2003). Neural substrates of tactile imagery: A functional MRI study. *Neuroreport*, *14*, 581–585.
- Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research: Brain Research Review*, *41*, 88–123.