

Neural processing associated with true and false memory retrieval

YOKO OKADO and CRAIG STARK

Johns Hopkins University, Baltimore, Maryland

We investigated the neural bases for false memory with fMRI by examining neural activity during retrieval processes that yielded true or false memories. We used a reality monitoring paradigm in which participants saw or imagined pictures of concrete objects. (A subsequent misinformation task was also used to increase false memory rates.) At test, fMRI data were collected as the participants determined whether they had seen or had only imagined the object at study. True memories were of seen pictures accurately endorsed as seen, and for false memories were of imagined pictures falsely endorsed as seen. Three distinct patterns of activity were observed: Left frontal and parietal activity was not different for true and for false memories, whereas activity was greater for true than for false memories in occipital visual regions and posterior portions of the parahippocampal gyrus, and activity was greater for false than for true memories in right anterior cingulate gyrus. Possible interpretations are discussed.

Our memories of past experiences are more fragile than many people realize. Not only are they vulnerable to poor encoding and forgetting, but memory retrieval is also highly susceptible to distortion (see, e.g., Loftus & Loftus, 1980; Schacter, 1999). These distorted memories, or *false memories*, are recollections that are inaccurate, misattributed, or “recollections” of entirely novel events (see, e.g., Loftus & Loftus, 1980; Roediger & McDermott, 1995; Schacter, 1999). Some false memories can be so influential and convincing that individuals believe they have experienced a particular event that is entirely novel (Loftus & Pickrell, 1995). False memories can occur because memories are not exact reproductions of past events (see, e.g., Roediger, 1996; Schacter, Norman, & Koutstaal, 1998). Rather, they are complex reconstructions easily interfered with and heavily influenced by previously formed memories and mental representations (see, e.g., Loftus, Miller, & Burns, 1978; Roediger & McDermott, 1995; Schacter et al., 1998).

This reconstructive process makes memories vulnerable to misattribution. Misattribution occurs when a particular memory is attributed to an incorrect time, place, or person (Roediger, 1996; Schacter, 1999). The Deese/Roediger-McDermott (DRM) paradigm, initially developed by Deese (1959) and substantially refined by Roediger and McDermott (1995), is frequently used to examine false recognition since it is a very robust and reliable means of inducing false memories in participants. In this para-

dig, participants study lists of semantically associated words but do not study a critical related lure word. In subsequent recall or recognition tests, the participants will not only frequently endorse as studied these critical lure words, but will also often freely recall these unstudied critical items (Roediger & McDermott, 1995). One explanation for this phenomenon is that participants are actually generating their own nonpresented lure word during the study phase. When asked to recall words, they confuse whether the word had actually been presented or they themselves had conjured it up (Roediger & McDermott, 1995; Schacter et al., 1998).

Marcia Johnson's source monitoring framework (SMF) suggests that during retrieval, memories are reconstructed differently from the way they were initially formed, enabling various errors to be introduced, such as features of similar events or imagined events that activate together to form an inaccurate memory. Thus, evaluation of the memory, through monitoring processes ranging from automatic to deliberate, is critical in determining the veridicality and source accuracy of the recollected memory (see, e.g., Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 2000). Similarly, Schacter et al.'s (1998) constructive memory framework (CMF) explains how memory reconstruction involves focusing on specific aspects of a memory so that irrelevant information is not retrieved, and setting criteria to determine whether a memory has sufficient perceptual vividness, semantic detail, and other phenomenal characteristics to be accurate.

Often, reality monitoring is used to discriminate memories that are generated by reflection or imagination from those based on perception. Externally perceived information and internally generated information may have similar semantic and perceptual content and may be processed similarly, although externally perceived memories typically have more of the information as well as

The research presented in this article was funded by Johns Hopkins University and NSF Grant BCS-0236431. We thank LisaCaitlin Perri, Marcel Poisoit, and the staff of the F. M. Kirby Center for Functional Brain Imaging for their assistance in data collection. Correspondence concerning this article should be addressed to C. Stark, Department of Psychological and Brain Sciences, Johns Hopkins University, Ames Hall, 3400 N. Charles St., Baltimore, MD 21218 (e-mail: cstark@jhu.edu).

contextual detail (Johnson, 1997; Johnson & Raye, 1981). Reality monitoring paradigms are used to investigate how people discriminate externally generated and internally generated memories. Behavioral studies show that failures in reality monitoring can occur when events that were imagined are remembered with considerable sensory detail and mistakenly believed to have actually occurred (e.g., Johnson & Raye, 1981; Johnson, Raye, Wang, & Taylor, 1979; Johnson, Taylor, & Raye, 1977). Because the brain mechanisms involved in mental imagery may be similar to those used in perceiving (see, e.g., Farah, 1989; Kosslyn et al., 1999; Wheeler, Petersen, & Buckner, 2000), it is conceivable that such an overlap can cause imagined events to be misperceived as real.

The propensity of misattribution can be increased by suggestion. Loftus and colleagues discovered that when people are given suggestive and misleading information about a previous event, their recollections of the original event can be altered by the misinformation (e.g., Loftus et al., 1978; Loftus & Palmer, 1974). It has been suggested that the misinformation may be replacing the original memory (Loftus et al., 1978), or it may be providing information that competes with the original memory during retrieval (McCloskey & Zaragoza, 1985). Regardless, it is clear that misinformation decreases the likelihood of accurate memory retrieval and increases the chances that misinformation will be misattributed to the original event.

Although behavioral studies have extensively investigated and provided evidence for the existence and creation of false memories (e.g., Deese, 1959; Johnson et al., 1977; Loftus & Palmer, 1974; Loftus & Pickrell, 1995), there remains much uncertainty about the neural mechanisms underlying the creation of false memories. Recently, neuroimaging studies have been conducted to begin to investigate the neural mechanisms underlying false memories in the production and manifestation at both the encoding and retrieval levels of the memory process (Cabeza, Rao, Wagner, Mayer, & Schacter, 2001; Gonsalves & Paller, 2000; Schacter, Buckner, Koutstaal, Dale, & Rosen, 1997; Schacter et al., 1996). It is not surprising that in many of the neuroimaging studies the DRM paradigm was adopted for the experimental design, since this paradigm robustly elicits false alarms to thematic words. This task is not without its limitations as a neuroimaging paradigm, however. The most pertinent limitation is that it requires many studied words for only a few critical lures. Despite this limitation, modified versions of the DRM paradigm have been used to examine the neural correlates of true and false recognition.

Using both positron emission tomography (PET; Schacter et al., 1996) and functional magnetic resonance imaging (fMRI; Cabeza et al., 2001; Schacter et al., 1997), researchers have observed that true and false recognition in the DRM task were associated with similar patterns of activity throughout much of the brain, suggesting that, by and large, true and false recognition are associated with similar neural mechanisms. Frontal

lobe activity was observed during both true and false recognition (which differed in both cases from activity for novel foil items), which suggests that this region may be involved in strategic monitoring processes that are activated when participants are attempting to determine whether a lure word was actually previously presented. Regions such as the anterior cingulate (Cabeza et al., 2001; Schacter et al., 1997), anterior prefrontal cortex (Schacter et al., 1997; Schacter et al., 1996), left ventrolateral prefrontal cortex (Cabeza et al., 2001), and bilateral dorsolateral prefrontal cortex (Cabeza et al., 2001; Schacter et al., 1997) also elicited similar activation patterns for true and false recognition.

In contrast with these similarities between true and false recognition, several differences have also been observed. Cabeza et al. (2001) reported that ventromedial prefrontal cortex was associated with more activity for false recognition in comparison with true recognition. This is consistent with numerous neuroimaging studies that have suggested that various prefrontal cortical regions are differentially involved with various types of memory retrieval, such as retrieval associated with effort or success, or performance of postretrieval verification or monitoring (e.g., Buckner, Koutstaal, Schacter, Dale, et al., 1998; Buckner, Koutstaal, Schacter, Wagner, & Rosen, 1998; Cabeza & Nyberg, 2000; Rugg, Fletcher, Frith, Frackowiak, & Dolan, 1996; Schacter et al., 1996; Wagner, Desmond, Glover, & Gabrieli, 1998) that might differ for true and for false recognition.

Also, the left parietal cortex appears to respond in a graded fashion depending on the depth of encoding during study and the amount of information that is recovered from memory during test (Nessler, Mecklinger, & Penney, 2001). Cabeza et al. (2001) found that activity in this region was greatest for retrieval that yielded true memories, then for false memories, and then for the correct rejection of new items. Likewise, Henson, Rugg, Shallice, Josephs, & Dolan (1999) found that this region showed a graded response to remember, know, and new judgments (see Tulving, 1985). Interestingly, Wheeler and Buckner (2003) found that the left parietal cortex responds to perceived oldness independent of the number of times the stimulus was studied or the modality of information retrieved. Even when a stimulus was new, if participants perceived it as old, the left parietal cortex was active. These findings suggest that the left parietal cortex serves as an index of how much contextual information is retrieved and how much the information is perceived as a previously encountered item or event.

The medial temporal lobe (MTL) plays a vital role in memory for facts and events (see Milner, Squire, & Kandel, 1998, for a review). As such, it is a clear target for neuroimaging research of false memory. The only study that has specifically examined the neural differences in the MTL between the retrieval of true and false memories (Cabeza et al., 2001) again used a version of the DRM paradigm. Cabeza et al. found an interesting dissociation between the parahippocampal gyrus and the

hippocampus. The parahippocampal gyrus was associated with greater activity for targets than for either foils or the critical lures (where activity was similar between the two). In contrast, the hippocampus was associated with greater activity for targets and critical lures than for foils. Cabeza et al. proposed that the parahippocampal gyrus (and not the hippocampus) is sensitive to the sensory properties of recovered information, and therefore could potentially differentiate between items seen before and items not seen before (even if both are endorsed during recognition).

Similarly, event-related potential (ERP) studies using the DRM paradigm showed that unilateral presentation of words during study elicited ipsilateral brain activity during test for centrally presented studied items but not for the highly related nonstudied items (Fabiani, Stadler, & Wessels, 2000). This suggests that true memories of studied items leave sensory signals of study experiences that make the memory traces distinctive. Nonstudied items lack these sensory signals. The P300 component of the ERP has also been identified with shorter latencies for false recognition of nonstudied lure items in comparison with true recognition of studied items (Miller, Baratta, Wynveen, & Rosenfeld, 2001).

Using the DRM paradigm, Nessler et al. (2001) reported that the nature of the encoding task determines whether there are neural differences between true and false memories. When the encoding task focused on the conceptual similarity of the studied items, brain activity for true and false recognition was similar, whereas they differed when the encoding task focused on specific item features. For example, during the latter task, there was a lack of frontomedial cortical activity during false recognition, an area frequently active during feelings of familiarity, when there was no such absence of activity during the first task.

Rather than examining the neural mechanisms of false memories elicited by the DRM paradigm, Gonsalves and Paller (2000) investigated false memories using an explicit reality monitoring paradigm. Using ERP, Gonsalves and Paller showed that during encoding, midline occipital and parietal activity during the imagination of a picture was more positive for those that were later mistakenly believed to have been actually perceived than for those that were later correctly identified as having been only imagined. In consistency with the SMF (see, e.g., Johnson et al., 1993), Gonsalves and Paller suggested that the more vivid the visual imagery during encoding, the more likely the imagined pictures are to become false memories of perceived pictures. During retrieval, there was greater midline occipital and parietal activity for accurate perceived picture memories than for false picture memories. This suggested that the imagery for false memories at retrieval was not as strong as the imagery for true memories. The pattern of results observed in the parietal cortex has also been noted by others examining true and false memory retrieval and imagery (Allan, Wilding, & Rugg, 1998; Cabeza et al., 2001; Henson et al., 1999; Nessler et al., 2001).

Thus, true and false memories appear to share much of the same neural activation patterns in similar brain regions, but evidence also suggests that clear differences exist. Since the nature of the encoding task appears to have an effect on whether neural differences are elicited between true and false memories, it is critical to use paradigms other than the frequently employed DRM paradigm to further examine these distinguishing markers.

One key difference between the DRM and reality monitoring paradigms is that in the DRM paradigm the falsely recognized items are not presented at study, whereas in the Gonsalves and Paller (2000) study the falsely recognized items are presented, at least in the form of a cue. This may affect activity during retrieval, depending on whether items presented at test correspond in any way to items presented at study.

We used the reality monitoring paradigm employed in Gonsalves and Paller's (2000) ERP study to investigate the neural differences between retrieval processes that yield true and false memories with fMRI to help provide further insight into the brain regions and neural circuitry associated with the creation of false memories.

THE PRESENT EXPERIMENT

Using a variation of Gonsalves and Paller's (2000) experimental paradigm, we examined and compared neural activity during true and false recognition. We report that activity in frontal and parietal regions, though memory related, did not differ between true and false recognition. Activity in several occipital regions and in the anterior cingulate, however, differed for true relative to false recollection. Occipital activity and anterior cingulate activity differed in numerous respects, however, which is indicative of their different roles in memory retrieval. Furthermore, we note that although, overall, there was little differential activity in the MTL, the activity that was observed mirrored the activation pattern observed in the occipital regions.

Method

The experiment was divided into three phases: *study*, *lie test*, and *test*. In the study phase, words naming concrete objects were presented auditorily and the participants were asked to vividly imagine the object (see Figure 1). An actual picture of the object followed the presentation of half of the objects.

A misinformation task (i.e., the lie test) was inserted between this study phase and the test phase, in an effort to increase the occurrence of false memories (see, e.g., Loftus et al., 1978; Loftus & Palmer, 1974). During the lie test, the participants were given a version of the test phase in which they were encouraged to lie about seeing a picture that they had only imagined during the previous study phase. Here, the participants played a game in which the scoring was rigged in such a way as to encourage them to claim they had previously seen a picture of an object that they had only imagined.

Later, during the test phase (in which absolute accuracy was stressed), having lied and claimed to have seen the picture earlier in the lie test induced a misattribution error, which served as a self-generated source of misinformation. As in previous studies of external misinformation (e.g., Loftus et al., 1978; Loftus & Palmer,

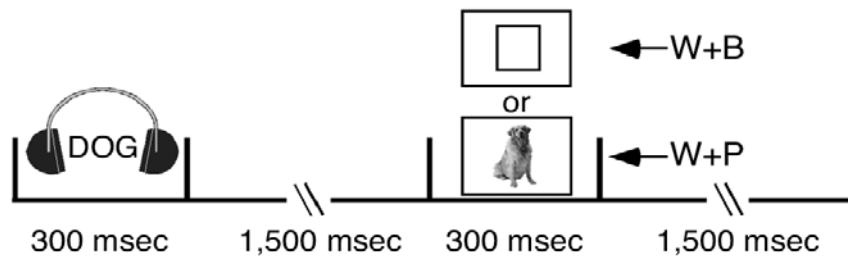


Figure 1. Schematic diagram of the study phase. Words (in this example, DOG) were presented auditorily, followed by a 1,500-msec delay. Either a picture of the object or a blank rectangle was presented for 300 msec, followed by a 1,500-msec delay. The participants imagined a picture of the object when the word was presented, and a response was required during the first 1,500-msec delay.

1974), the participants' memories appear to have been altered to the point that they falsely believed they had actually seen the picture, since the false alarm rate was raised significantly.

The test phase was the only in-scanner portion of the experiment. In the test phase, the participants heard names of objects they had seen accompanied by a picture, objects that were presented without a picture, and new objects. The participants were asked to determine if they had actually seen a picture of the object during the study phase. We were specifically interested in trials in which the participants accurately endorsed perceived pictures as perceived (true memory) and trials in which they inaccurately endorsed imagined pictures as perceived (false memory).

The sole purpose of including the lie test was to increase the number of false memories available for fMRI analysis, which it appears to reliably do. In a control experiment conducted outside the scanner, half of the studied stimuli were presented in the lie test and half were not presented to each participant. Overall, the false memory rate (i.e., the rate of inaccurate endorsement of words presented alone as having been presented along with a picture) during the test phase was significantly higher [$t(13) = 2.32, p < .05$] for those items presented in the lie test (17%) than for those items that were not presented in the lie test (13%). Although this increases the number of trials available for fMRI data analysis, it is nearly impossible to predict what cognitive or neural processes are engaged during the lie test, since true errors made by the participants cannot be distinguished from deliberate errors. Likewise, it would be exceptionally difficult, if not impossible, to determine in the present paradigm whether any particular false memory had its source in the original study phase, in the lie test, or in some combination of both. Since we were interested specifically in examining retrieval processes yielding false memories and the reconstructive nature of the retrieval process, we collected fMRI data only from the test phase.

Participants

Fourteen native English speakers (8 male, 6 female) were recruited from the Johns Hopkins University community. The participants were between the ages of 19 and 31 years, were right-handed, and had no history of neurological or psychiatric illness. All the participants were naive to the experimental materials and hypotheses and gave informed written consent to participate.

Materials

The stimuli consisted of 450 color photographs of objects on a white background and auditorily presented names of each object (recorded as sound files by a single female native English speaker). Outside the scanner, pictures were presented on a computer screen and the object names were played over headphones. Inside the scanner, object names were played over pneumatic headphones and responses were collected using a fiber optic button box.

Procedure

The study phase consisted of a total of 300 trials: 150 spoken words followed by pictures (W+P trials) and 150 spoken words followed by a blank rectangle (W+B trials). The duration of the spoken words was approximately 300 msec, followed by a delay (~1,500 msec), for a total of exactly 1,800 msec (see Figure 1). Corresponding pictures of the objects or blank rectangles were presented for 300 msec, followed by a 1,500-msec delay. The participants were instructed to visualize an image of the object and decide if the object was larger or smaller than a shoebox, indicating their response on the computer keyboard after the word was presented. The participants were informed that a corresponding picture or filler rectangle would appear on the screen soon after the word was presented. The words presented with pictures and those presented with blank rectangles were counterbalanced. In addition to trials with stimuli, there were 50 null trials in the form of auditory noise (3,600 msec) during which the participant made no response. All of the stimuli were randomly intermixed. The study phase was divided into two runs.

The lie test consisted of 225 trials (75 of the W+P trials and all 150 of the W+B trials from the study phase) of auditorily presented words. For each word, the participants were asked to decide whether or not they had seen an actual picture of the object in the study phase. On each trial, the participants could earn or lose points with each response, with the goal of accumulating as many points as possible (indicated on the screen). Truthful responses were awarded 2 ± 1 points. When the participants lied and endorsed a word studied without a picture, they would gain 8 ± 1 points on 70% of the trials and lose 4 ± 1 points on 30% of the trials. On the latter set of trials (those checked for accuracy), an animated figure would appear along with a sound effect as the participant's score was reduced. Finally, on trials in which the participants lied (or were inaccurate) and did not endorse a trial that had been previously presented with a picture, they lost 2 ± 1 points if the trial was checked for accuracy but gained 4 ± 1 points if it was not checked. Therefore, the participants were strongly encouraged to lie and endorse words that had not been studied with a picture, and they were only moderately encouraged to lie and not endorse words that had been studied with a picture. The participants were informed that they would earn or lose a certain number of points with every response they made. They were given the goal of collecting as many points as possible and were told that they would receive feedback on every trial in the form of a bar graph and the actual number of points collected. The participants were informed that although it benefited them to respond truthfully as to whether or not they had seen a picture, it might benefit them even more to lie if the particular trial was not checked. They were told that only a small number of the trials would be checked. The participants indicated their response on the computer keyboard. The trials were self-paced.

The participants were then placed in the fMRI scanner for the test phase. The test phase consisted of 450 trials: 150 W+P trials, 150 W+B trials, and 150 trials of spoken words that had not been presented at study (foils). Words were presented auditorily every 2,500 msec. The participants were instructed to decide whether or not they had seen an actual picture of the object in the study phase. They were instructed that this portion of the experiment pertained only to the study phase and to tell the truth to the best of their ability. They indicated their “yes” and “no” responses using a button box. In addition, there were 75 null trials in the form of auditory noise (2,500 msec), during which the participants made no response. These trials served as a baseline condition for the fMRI data analysis. All of the stimuli were randomly intermixed. The test phase was divided into five runs.

fMRI Methods

Imaging was performed on a Philips Gyroscan 1.5T MRI scanner equipped with a whole-brain SENSE coil. Thirty-five T2*-weighted triple-oblique functional images were collected per 3-D volume using a single-shot echoplanar pulse sequence (64×64 matrix, TE = 40 msec, flip angle = 90° , in-plane resolution = 4×4 mm, thickness = 4 mm, TR = 2.5 sec). Slices were aligned with the principal axis of the left and right hippocampus (as determined by a series of sagittal localizer MRI scans for each participant). This was done to optimize the signal from the medial temporal lobes and to minimize partial-volume effects, so that voxels could be clearly constrained to lie within subregions of the MTL. A total of 525 volumes were collected. The task began in synchrony with the acquisition of the fifth volume to allow for T1 stabilization. After the functional scans, a high-resolution structural MRI was acquired (MP-RAGE pulse sequence, 1 mm³ resolution, 150 triple-oblique axial slices in the same orientation as the functional images) for anatomical localization.

fMRI Data Analysis

Image analysis was performed using Analysis of Functional Neuroimages (Cox, 1996). Functional MRI data were first resampled in time using a Fourier algorithm to align all slices to a common time base. Functional images were then resampled in space to coregister the images and reduce the effects of head motion in three dimensions. During this process, six vectors were created that code for all possible translations and rotations of the brain. fMRI data from all test runs were concatenated.

Following this processing, the behavioral data were coded into eight trial types of interest, and a general linear model (GLM) of the fMRI time series data was constructed using these vectors. The stimulus vectors coding for the eight trial types were named according to study phase condition and participant response as follows: (1) S–S, seen pictures endorsed as seen (“yes” response to W+P items); (2) S–S (no MI), “yes” response to W+P items that were not presented during the lie test and, therefore, received no misinformation; (3) S–N, seen pictures rejected as not seen (“no” response to W+P items); (4) S–N (no MI), “no” response to W+P items that were not presented during the lie test and, therefore, received no misinformation; (5) I–S, imagined pictures endorsed as seen (“yes” response to W+B items); (6) I–N, imagined pictures rejected as not seen (“no” response to W+B items); (7) U–S, unstudied words endorsed as seen (“yes” response to foils); and (8) U–N, unstudied words rejected as not seen (“no” response to foils). Unfortunately, with only 10 trials per participant on average (range = 2–22), reliable estimates of the hemodynamic response in the U–S condition could not be obtained and, therefore, this trial type was not included in the group analysis. In addition, the GLM included nuisance vectors coding for first- and second-order drift in the MR signal and for 3-D head motion.

The GLM was constructed using a deconvolution technique (Ward, 2000) that estimates the impulse response function within

each voxel and performs a multiple linear regression. The sum of the beta coefficients for the time points corresponding to the expected peak in the hemodynamic response (-2.5 – 10 sec after stimulus onset) was taken as the model’s estimate of the response to each trial type. We should note that this response magnitude is relative to the activity present during the null task (auditory noise). An activity level of zero for a particular condition (and a flat hemodynamic response) does not imply a lack of activity in the region during this condition. Rather, it implies merely a similar level of activity in the condition of interest and in the null task, which may very well be active (Stark & Squire, 2001b).

Initial spatial normalization was done using each participant’s structural MRI to transform data according to the common atlas of Talairach and Tournoux (1988). This transformation was applied to the statistical maps of the beta coefficients and in the process the data were resampled to 2.5 mm³. The spatially normalized statistical maps of the beta coefficients were blurred using a Gaussian filter with a full-width half maximum of 4 mm to help account for variations in the functional anatomy across participants.

The main analyses involved a voxel-wise two-factor analysis of variance (ANOVA) on the beta coefficients with a fixed factor (condition: all seven trial types) and a random factor (subject) to identify any activation differences between the trial types. The ANOVA was used to define functional regions of interest (ROIs) that demonstrated significant activation differences across any of the seven trial types that passed a voxel-wise threshold of $p < .01$ and a spatial extent threshold of 328 mm³ ($p < .05$, correcting for multiple comparisons). Post hoc t tests were performed on the average beta coefficients within the ROIs functionally defined by the ANOVA. These t tests were pairwise comparisons of the seven conditions. Hemodynamic responses were extracted for three trial types of interest (S–S, I–S, and I–N) for the identified regions that showed significant differences among these three conditions.

The MTL analysis was performed using the ROI–AL (regions of interest–alignment) technique described by Stark and Okado (2003). The technique begins with a definition of the structures in the MTL (hippocampal region, temporopolar, perirhinal, entorhinal, and parahippocampal cortices) bilaterally according to the techniques described by Insausti et al. (1998). The parahippocampal cortex was further defined bilaterally as the portion of the parahippocampal gyrus caudal to the perirhinal cortex and rostral to the splenium of the corpus callosum (by this definition, the rostral extent includes only slightly more tissue than as defined by Pruessner et al., 2002). The hippocampal region (the CA fields of the hippocampus, dentate gyrus, and subiculum) was also defined bilaterally. ROI–AL uses these anatomically defined ROIs to calculate an additional transformation matrix to fine-tune the cross-participant alignment in the MTL. ROI–AL significantly improves the overlap across participants, increasing statistical power and precision in localization of cross-participant tests. In the ROI–AL analysis, a voxel-wise threshold of $p < .01$ and a spatial extent threshold of 125 mm³ was used ($p < .05$, correcting for multiple comparisons with the MTL alone).

RESULTS

Behavioral Results

The overall response rates for each category are shown in Figure 2a. To assess accuracy and false memory rates, overall discriminability (d') scores were calculated from the S–S rate for each of the two types of false alarm rates. The d' for S–S and I–S was 1.38, and the d' for S–S and U–S items was 1.98. These d' scores significantly differed [$t(13) = 5.34, p < .001$], as did the overall rates for the two false alarm conditions [$t(13) = 3.43$,

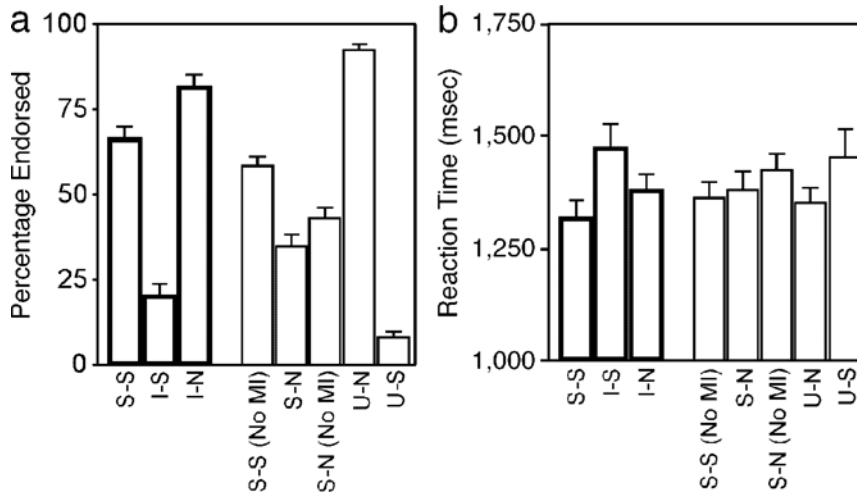


Figure 2. Behavioral data from the test phase. Percentage of trials endorsed as having been presented with a picture at time of study (a) and reaction time for the decision (b) are plotted for each trial type as means across all 14 participants. Error bars indicate the standard errors of the means.

$p < .01$], suggesting that the combination of the reality monitoring paradigm and the lie test significantly increased the false alarm rate to above chance levels and that the participants were actually experiencing false memories. During the lie test, the participants claimed 47.6% of W+B items were studied as pictures. Fifty-three percent of the I-S items from the test phase were items that were falsely endorsed during the lie test.

Of the seven possible trial types, three critical trial types were selected a priori for further analyses in an effort to analyze differences between true and false memories: S-S, I-S, and I-N (the left three bars in Figures 2 and 3c-e). Differences between S-S and I-S activations are suggestive of differences between true and false memory retrieval. Both trial types were presented in the study phase (S-S as W+P and I-S as W+B) and in the lie test, and both were endorsed in the test phase (S-S correctly endorsed and I-S incorrectly endorsed). Similarly, a contrast between I-S and I-N can highlight differences between true and false memory retrieval by indicating differences between the false endorsement of an imagined picture as being real and the correct rejection of an imagined picture as having been only imagined (in both of these conditions, the stimuli have been treated identically prior to the test phase). Finally, by contrasting true retrieval of pictures and correct rejection of only imagined pictures, differences between S-S and I-N activations are suggestive of accurate retrieval. However, it should be noted that although this contrast is often used to assess memory retrieval success, it can be contaminated and weakened by activity associated with incidental encoding during the retrieval task (Stark & Okado, 2003).

Figure 2b shows the mean reaction times (RT) for all trial types. In the three conditions of interest, the mean

RTs were fastest for S-S items (1,313 msec) and slowest for I-S items (1,469 msec). A repeated measures ANOVA showed that there was a significant effect of these three trial types on RT [$F(1,13) = 22.5, p < .001$]. Pairwise t tests showed significant RT differences between S-S and I-S [156 msec; $t(13) = -4.74, p < .01$], between I-S and I-N [95 msec; $t(13) = 2.71, p < .05$], and between S-S and I-N [61 msec; $t(13) = -3.06, p < .01$].

fMRI Results

A voxel-wise two-factor ANOVA was first conducted on the whole brain data to identify regions that showed any activation differences among the seven trial types. Results of this analysis identified 12 areas in which activity significantly differed across trial types (Figure 3a). These areas were treated as functionally defined ROIs, and the average hemodynamic response functions for each trial type were calculated for each region.

The 12 regions identified demonstrated three distinct patterns of activity in the functional ROI analysis with respect to our three conditions of interest. The first notable pattern observed in the identified regions was similar increased activity for S-S and I-S in comparison with I-N (Figure 3c). This pattern of activity was observed in the left parietal cortex and left frontal regions. Regions in which S-S and I-N differed significantly and I-S and I-N differed significantly included the left parietal cortex [BA 7/39/40; $t(13) = 3.29, p < .01$ and $t(13) = 2.57, p < .05$, respectively], left precentral gyrus [BA 9; $t(13) = 2.45, p < .05$ and $t(13) = 3.19, p < .01$, respectively], and left inferior frontal gyrus [BA 10/46; $t(13) = 2.61, p < .05$ and $t(13) = 2.98, p < .05$, respectively]. The left caudate showed this similar pattern of activation, with S-S and I-S demonstrating increased activity in comparison with I-N, but showed no signifi-

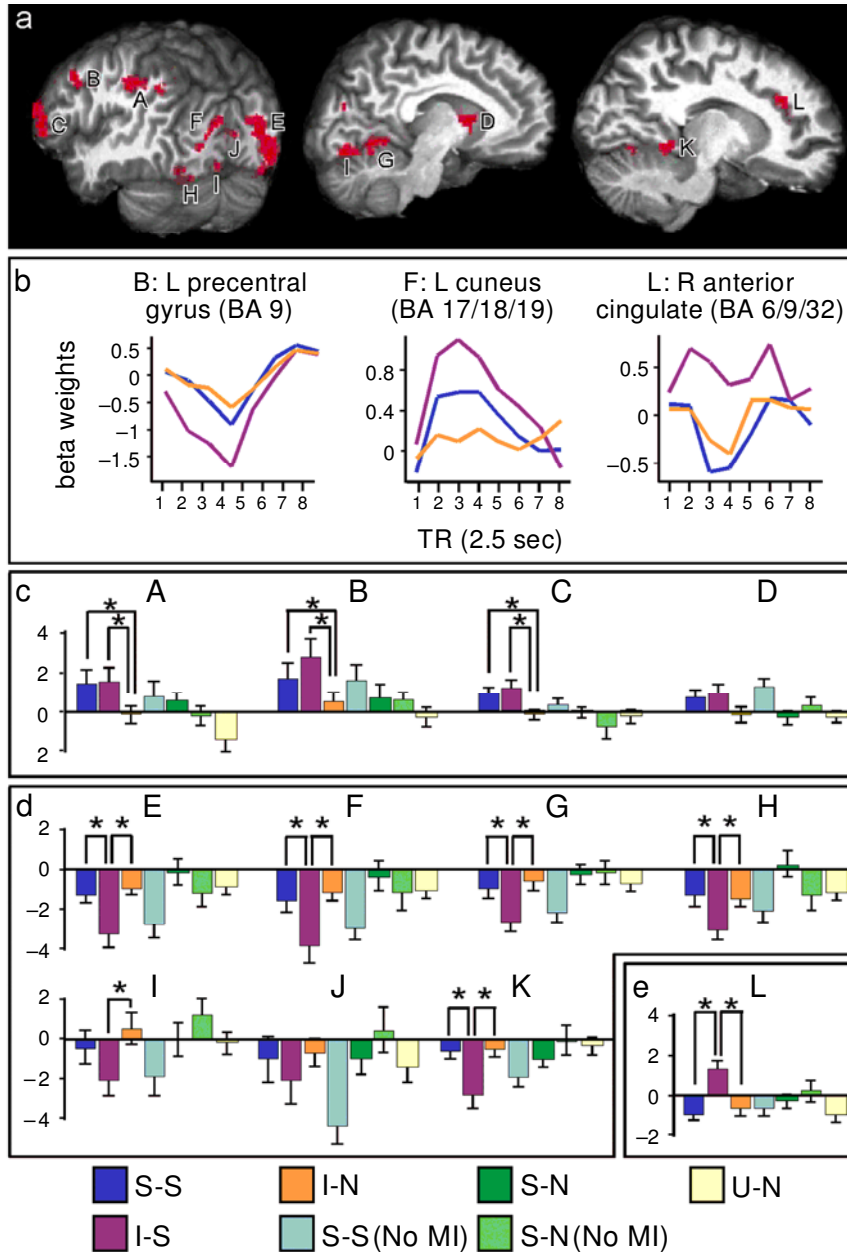


Figure 3. (a and b) Locations and sample hemodynamic responses for regions where activity at test varied among the seven trial types analyzed (there were too few trials in the U-S condition to analyze). (a) The 12 regions where activity varied as a function of trial type are shown as colored overlays on 3-D renderings of a brain. A: left parietal cortex (BA 7/39/40); B: left precentral gyrus (BA 9); C: left inferior frontal gyrus (BA 10/46); D: left caudate; E: right middle occipital gyrus (BA 18/19); F: left cuneus (BA 17/18/19); G: left lingual gyrus (BA 19); H: left fusiform gyrus (18/19); I: bilateral lingual gyrus (BA 18); J: bilateral cuneus (BA 17/18/27/30); K: right posterior parahippocampal gyrus (BA 30/35/36/37); L: right anterior cingulate gyrus (BA 6/9/32). (b) Average hemodynamic response functions (sum of beta coefficients vs. image acquisition [TR]) for the three trial types of interest (S-S, I-S, and I-N) are shown for three sample regions (left precentral gyrus, left cuneus, and right anterior cingulate) that demonstrate the patterns of activity observed. (c-e) Activity for all seven trial types analyzed in all 12 functionally defined ROIs. Bars show the mean fMRI response (sum of beta coefficients) across participants, and error bars show the standard errors of the means. Asterisks indicate significant differences in activity between conditions of interest (S-S, I-S, and I-N). (c) Regions demonstrating the first pattern of activity included A, left parietal cortex; B, left precentral gyrus; C, left inferior frontal gyrus; and D, left caudate. (d) Regions demonstrating the second pattern of activity included E, right middle occipital gyrus; F, left cuneus; G, left lingual gyrus; H, left fusiform gyrus; I, bilateral lingual gyrus; J, bilateral cuneus; and K, right posterior parahippocampal gyrus. (e) The third pattern of activity was observed in L, right anterior cingulate gyrus.

cant differences between S–S and I–S, between I–S and I–N, or between S–S and I–N.

The second pattern of activity (Figure 3d) was associated with similar levels of activity for S–S items and I–N items that was substantially greater than the activity associated with I–S items. This pattern of activity was observed primarily in the occipital region as well as in a portion of the parahippocampal gyrus just posterior to the parahippocampal cortex. Regions in which S–S and I–S differed significantly and I–S and I–N differed significantly included right middle occipital gyrus [BA 18/19; $t(13) = 3.33, p < .01$ and $t(13) = -3.52, p < .01$, respectively], left cuneus [BA17/18/19; $t(13) = 3.53, p < .01$ and $t(13) = -3.47, p < .01$, respectively], left lingual gyrus [BA 19; $t(13) = 4.48, p < .01$ and $t(13) = -4.99, p < .001$, respectively], left fusiform gyrus [BA 18/19; $t(13) = 2.57, p < .05$ and $t(13) = -2.91, p < .05$], and right posterior parahippocampal gyrus [BA 30/35/36/37; $t(13) = 4.44, p < .01$ and $t(13) = -4.57, p < .01$, respectively]. Bilateral lingual gyrus (BA 18) showed significant activation differences for only the I–S and I–N [$t(13) = -3.18, p < .01$] contrast. Bilateral cuneus (BA 17/18/23/30) elicited the same activation pattern, but showed no significant differences between S–S and I–S, between I–S and I–N, or between S–S and I–N.

The third pattern of activity noted in the identified regions was increased activity for I–S items in comparison with S–S and I–N items (Figure 3e). This pattern of activity was observed in the right anterior cingulate gyrus (BA 6/9/32). This region showed significant activation differences between S–S and I–S [$t(13) = -5.48, p < .001$] and between I–S and I–N [$t(13) = 3.25, p < .01$].

A separate analysis was done on the data from the MTL, since the MTL plays a vital role in declarative memory tasks such as those used in the current experiment (see Milner et al., 1998, for a review). Using the ROI–AL technique (Stark & Okado, 2003) to optimize the alignment and statistical power within the MTL, a separate voxel-wise two-factor ANOVA was performed on the aligned data with a clustering threshold adjusted to reflect the number of voxel-wise comparisons within the MTL. Four regions in the MTL showed activity differences across the trial types. Two regions showed activity similar to that of the second activation pattern discussed. That is, S–S and I–N items were similarly more active in comparison with I–S items. A region within the left temporopolar portion of perirhinal cortex showed significant activation differences between S–S (beta weight = 0.74, $SEM = 0.23$) and I–S [beta weight = $-1.7472, SEM = 0.41$; $t(13) = 2.72, p < .05$] and between I–S and I–N [beta weight = $-0.29, SEM = 0.31$; $t(13) = -3.3, p < .01$]. Right parahippocampal gyrus (also observed in the whole-brain ANOVA results) showed significant activation differences between S–S (beta weight = $-0.27, SEM = 0.47$) and I–S [beta weight = $-2.48, SEM = 0.592$; $t(13) = 5.59, p < .001$] and between I–S and I–N [beta weight = $-0.64, SEM = 0.47$;

$t(13) = -6.18, p < .001$]. This region included a portion of the most posterior extent of the parahippocampal cortex and extended significantly posterior to our definition of the parahippocampal cortex, with the majority of the region lying outside of the parahippocampal cortex. The two other regions identified in the MTL, an area within the right temporopolar portion of perirhinal cortex and an area within left entorhinal cortex, did not show any significant activation differences across any of the trial types of interest. For both of these regions, however, activity for S–N differed significantly from that of the rest of the trial types [specifically, S–N (no MI) for right temporopolar cortex was least active, and S–N for left entorhinal cortex was most active].

DISCUSSION

In this study, we compared the neural bases of memory retrieval processes that yielded true and false memories. Three key contrasts were examined: (1) neural differences between true recognition of pictures that had been seen and false recognition of pictures that had been only imagined (S–S vs. I–S, respectively); (2) false recognition of imagined pictures and correct rejection of imagined pictures (I–S vs. I–N, respectively); and (3) true recognition of seen pictures and correct rejection of imagined pictures (S–S vs. I–N, respectively). The first two contrasts help identify differences between true and false memory retrieval, whereas the third helps identify activity associated with accurate memory retrieval.

We observed that certain regions of the brain showed strikingly distinct patterns of activation for retrieval processes that yield true and false memories. In left parietal and left frontal cortices, S–S and I–S trials were similarly more active than I–N trials. Upon further investigation, left parietal and left frontal activity differed outside of our conditions of interest, in that the left parietal cortex showed low levels of activity associated with novel U–N items. Bilateral occipital regions and right posterior parahippocampal gyrus (posterior to parahippocampal cortex) showed significantly increased activity for both S–S and I–N trials in comparison with I–S trials. In right anterior cingulate gyrus, there was more activity for I–S trials than for S–S and I–N trials. These distinct neural response patterns are suggestive of differential processing by the various brain regions for the same information.

Before we turn to a discussion of the possible sources of activity for S–S, I–S, and I–N trials, it will be useful to understand the potential effects of the misinformation task used in the lie test. This misinformation phase was deliberately employed to convert I–N trials into I–S trials (and, potentially, S–N trials into S–S trials), and it was successful in increasing the overall false memory rate. It is clear that from the data at hand, it would be difficult if not impossible to understand what cognitive or neural processes were engaged during this task and how these processes have affected the results. However, since

the primary interest of this study was to examine the false memory reconstructive retrieval process and not how encoding affects false memories, the use of the lie test should not affect what one can conclude about activity for I-S items during the test phase. We would caution careful interpretation wherever a difference between S-S and S-S (no MI) activity was observed. Although both represent accurate memory retrieval, S-S (no MI) trials were not presented during the lie test, whereas S-S trials were. If activity differs for S-S and S-S (no MI) trials, we must clearly consider the influence of the misinformation task on brain activity.

In addition to any effects of the misinformation task, it is worthwhile to note that activity from our trials of interest could be due to several possible sources: S-S activity (e.g., episodic memory of picture during study phase or vivid memory of picture without episodic component); I-S activity (e.g., false episodic memory of picture during study phase or vivid memory of picture without episodic component); and I-N activity (e.g., episodic memory of no picture during study phase or poor image of picture retrieved).

With these characterizations in mind, we can attempt to understand the computational implications of the three observed patterns of activity. The first activation pattern observed for the three conditions of interest (Figure 3c) is consistent with the hypothesis that activity in left parietal and left frontal regions is affected by the amount of information retrieved (episodic/source and item-based components of the memory). According to this hypothesis, one would predict that not only should S-S and I-S activity be high relative to I-N, but U-N activity should also be lowest. The left parietal cortex demonstrated this pattern of a graded response of how much contextual information was recovered or believed to be recovered. When the participants believed they had studied a picture, regardless of whether they in fact had, there was a trend in this region for activity to be greatest. Items that were studied but forgotten and those that were correctly rejected showed a trend of less activity, suggesting that the participants remembered studying the words but did not remember them in enough detail to believe they had seen pictures. Novel items that were correctly rejected showed the least amount of activity, suggesting that the participants could recall little or no contextual information associated with the novel words. Similar results have been obtained by others investigating true and false memories and imagery, but those authors reported that retrieval activity that resulted in true memories was greater than activity that resulted in false memories (Allan et al., 1998; Cabeza et al., 2001; Gonsalves & Paller, 2000; Henson et al., 1999; Nessler et al., 2001). It may be that the participants in the present study recalled as much perceptual, sensory, and contextual detail about the imagined objects as they did about the perceived objects. Wheeler and Buckner (2003) found that as long as participants perceived an item as old, left parietal cortex was active regardless of accuracy, modality of item

presentation, and number of times the item was studied. Although Johnson and Raye (1981) suggest that perceived and imagined events differ in a number of ways (e.g., quality and quantity of detail recovered, cognitive operations used to form memory), it appears that activity of the left parietal cortex was correlated with the amount of information recovered or believed to be recovered from memory.

In the left frontal regions, however, S-S, I-S, and I-N activity showed the same pattern, whereas U-N activity was not significantly lower than I-N activity. These regions may instead be associated with response monitoring of information believed to have been studied or successful source retrieval. If so, S-S and I-S activity should be similar and higher than all other conditions, since in each of these trial types participants are retrieving pictures and images believed to have been seen. This pattern was indeed found in the left frontal regions and is consistent with Cabeza et al.'s (2001) observation that true and false recognition evoked activity in bilateral dorsolateral prefrontal cortex. Similarly, other studies have reported activity in left anterior prefrontal cortex that is associated with the retrieval of perceptual information (Ranganath, Johnson, & D'Esposito, 2000). The prefrontal cortex is also frequently associated with successful memory retrieval, including accurate episodic remembering (Buckner, Koutstaal, Schacter, Dale, et al., 1998; Buckner, Koutstaal, Schacter, Wagner, & Rosen, 1998; Cabeza & Nyberg, 2000; Ranganath et al., 2000). Thus, the activity observed in the frontal regions in the present study is consistent with several findings in the literature, suggesting that these regions are associated with response monitoring or successful source retrieval. As is suggested by Johnson's (Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 2000) SMF and Schacter's (Schacter et al., 1998) CMF, recovered information must undergo an evaluation and criterion-setting process to determine the accuracy of the memory.

The second activation pattern, observed primarily in visual areas (Figure 3d), although quite robust, is difficult to interpret from the present data. The pattern of the three conditions of interest suggests that numerous occipital regions and the posterior right parahippocampal gyrus contain information about the difference between true and false responses, since I-S trials showed significantly less activity than S-S or I-N trials (except in bilateral lingual gyrus and bilateral cuneus, where the same trend was observed but the differences were not significant). By facilitating an examination of the predictions a hypothesis such as this would make for the other four stimulus conditions, the present design allows us to test such a hypothesis to a greater extent than has previously been possible in the neuroimaging of false memory. For example, if activity were simply a function of the truthfulness of the response, one might predict that all accurate responses should be equally active and greater in activity than all inaccurate responses, which also should be equally active. It is also possible that re-

retrieval of seen pictures elicited more activity than retrieval of imagined pictures believed to have been seen. Perceived objects are rich in sensory detail, whereas imagined objects tend to lack such perceptual detail (Johnson & Raye, 1981); therefore, it is not surprising that true memories evoked more activity in the occipital region than did false memories of imagined pictures. In such reality monitoring studies, it is often found that retrieval of real pictures contains more details and information than retrieval of imagined pictures (Johnson & Raye, 1981). Furthermore, although the retrieval of imagined pictures has elicited activity in both primary and higher order visual cortex (Kosslyn et al., 1999; Wheeler et al., 2000), some reports on whether imagined pictures elicit primary visual cortical activity have been inconsistent with each other (see Cabeza & Nyberg, 2000, for a review), raising the possibility that imagined pictures do not possess all the elements of perceived pictures necessary for the brain to respond as if a real picture were being remembered. It is also possible that the amount of activity is determined by the total amount of study exposure. It is certainly possible that one effect of retrieving and reencoding these stimuli in the lie test was to enhance the strength and vividness of the memory for these pictures. Similarly, the increase in the false memory rate resulting from the lie test is consistent with the idea that the misinformation phase enhanced the vividness of the participants' memory for items initially only imagined. The difference between S-S and S-S (no MI) activity in all the areas in the occipital lobes (except left fusiform gyrus) and right posterior parahippocampal gyrus is indicative of a substantial role of the lie test, be it in the form of a second exposure or as a more complex effect. It is difficult to interpret any influence on neural activity that the misinformation effect may have on these brain regions, but this activation pattern was robust and prevalent, suggesting that these regions are processing the retrieval of real and imagined pictures in a systematic fashion. The presence of the additional five-stimulus conditions in our experimental design (four of which can be analyzed) has allowed us to reject all three of these hypotheses when they are considered in isolation. Thus, although this pattern was clearly robust, it appears to be the result of an interaction of the factors discussed above (or other, unknown factors) that is beyond explanation from the present data. Further experimentation designed to more directly address this issue is therefore warranted.

The third observed pattern of activity of the three conditions of interest (Figure 3e) in the right anterior cingulate gyrus suggests that this region is heavily associated with effort. The neural activation and behavioral RT pattern demonstrate a strong correlation in that I-S activity and RT are highest and S-S activity and RT are lowest, and the other trial types fall in between (Figure 2b and Figure 3e). Interestingly, this region has been identified to be active during erroneous responses as well as during correct responses when there is a high level of response competition or conflict (Botvinick, Nystrom, Fissell,

Carter, & Cohen, 1999; Carter et al., 1998). For example, Carter et al. demonstrated that the anterior cingulate cortex detects situations in which errors are likely to occur. A prediction of this hypothesis is that S-S and U-N trials should elicit the smallest amount of activation, followed by S-S (no MI) and I-N trials, followed by activity associated with S-N trials, and, finally, by I-S trials (which were associated with the longest RTs, are erroneous responses, and have clear potential for competition between conflicting responses). Exactly this pattern was observed in the data. This interpretation of effort and high conflict driving the right anterior cingulate is clearly consistent with the data, since this region was the only one to demonstrate increased activity for false memories in comparison with the other trial types. It is reasonable to assume that high levels of conflict and effort are involved when imagined pictures are believed to be real.

It is worth noting that although the MTL is heavily associated with memory encoding and retrieval processes (see, e.g., Milner et al., 1998), we did not observe prominent activity throughout the MTL as one might expect in a declarative memory retrieval task such as the one used here. It is important to note that the literature on memory retrieval is marked by numerous such failures to observe MTL activity, since the contrasts used to index retrieval success are often confounded with incidental encoding processes (Stark & Okado, 2003; Stark & Squire, 2000, 2001a). Here, the lack of hippocampal and other MTL activity may ultimately be due to confounding effects such as incidental encoding or the presence of episodic or source memory components in all the trial types analyzed. Such a commonality may have resulted in similar levels of activity across the trial conditions in many MTL regions. Alternatively, it is quite possible that true and false recognition are not differentiated by the MTL. If the role of the MTL is to integrate pieces of information and later help recover these pieces via this reconstructive retrieval process, a misattributed memory may be indistinguishable at this level.

We should note that although large regions of the MTL were not active differentially across trial types, it was not the case that we observed no differential activity in the MTL at all. In an analysis optimized for the MTL, we did observe MTL activity that varied by trial type in posterior right parahippocampal gyrus, right entorhinal cortex, and bilateral temporopolar cortex. We should note that although the activity in the right parahippocampal gyrus extended to include some of the right parahippocampal cortex, it was largely posterior to what we define as parahippocampal cortex. Nonetheless, the pattern of activity in right parahippocampal gyrus and left temporopolar cortex was strikingly similar to the pattern observed in the occipital regions (Figure 3d). Such parahippocampal gyrus activations (often it is unclear whether the findings lie within the parahippocampal cortex) have been reported to occur during memory retrieval (see, e.g., Cabeza et al., 2001; Schacter et al., 1997;

Stark & Okado, 2003). Cabeza et al. (2001) suggested that this region responds to sensory details of the information retrieved and can therefore differentiate between memories of a perceived stimulus and memories of an internally generated stimulus, because a presented stimulus carries more sensory details than one that was not presented, which should, at best, seem familiar. In accordance with Cabeza et al.'s results, the parahippocampal gyrus activity from this study did show more activity for S–S than for I–S trials. However, it also showed more activity for S–S than for S–S (no MI) trials and equal activity for S–S, S–N, and U–N trials, which does not concur with the explanation offered by Cabeza et al. Although the tasks (DRM vs. reality monitoring) clearly differed, it is unclear how this could have affected the results given the hypothesis proposed by Cabeza et al., since the present task seems to provide ample opportunity to assess activity for externally perceived versus internally generated stimuli as well. The pattern of activity observed here (and, potentially, those reported by Cabeza et al., given the marked resemblance to the pattern of activity observed in the occipital region) suggests that the MTL activations observed in this study may not be the result of MTL memory processes but, rather, the result of visual cortex's feeding information into these regions (Suzuki & Eichenbaum, 2000). Logically, the converse—that is, that these MTL regions are driving the activity observed in visual cortex—could also be true. With its poor temporal resolution, fMRI cannot clearly differentiate between the two.

This study has demonstrated that occipital and right posterior parahippocampal gyrus showed greater activity for seen pictures than for imagined pictures believed to be seen. In contrast, the right anterior cingulate regions showed the opposite pattern. These data suggest a distinction between retrieval processes that yield true and false memories. Finally, the left parietal lobe, left frontal lobe, and MTL looked similar for seen pictures and for imagined pictures believed to be seen, suggesting an inability to detect differences between retrieval processes that yield true memories and those that yield false memories.

Thus, the data suggest that for true and false memories of pictures, the occipital and posterior parahippocampal gyrus regions show activity that distinguishes these memories, whereas the left frontal and parietal activity serve as an index of how much both true and false memories are believed to be true. The anterior cingulate gyrus shows activity that isolates memories that are associated with effort and conflict, which tend to be the false memories in this type of reality monitoring paradigm. How these various regions work together during retrieval to yield true and false memories is a topic for future research.

REFERENCES

- ALLAN, K., WILDING, E. L., & RUGG, M. D. (1998). Electrophysiological evidence for dissociable processes contributing to recollection. *Acta Psychologica*, **98**, 231-252.
- BOTVINICK, M., NYSTROM, L. E., FISELL, K., CARTER, C. S., & COHEN, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, **402**, 179-181.
- BUCKNER, R. L., KOUTSTAAL, W., SCHACTER, D. L., DALE, A. M., ROTTE, M., & ROSEN, B. R. (1998). Functional-anatomic study of episodic retrieval: II. Selective averaging of event-related fMRI trials to test the retrieval success hypothesis. *NeuroImage*, **7**, 163-175.
- BUCKNER, R. L., KOUTSTAAL, W., SCHACTER, D. L., WAGNER, A. D., & ROSEN, B. R. (1998). Functional-anatomic study of episodic retrieval using fMRI: I. Retrieval effort versus retrieval success. *NeuroImage*, **7**, 151-162.
- CABEZA, R., & NYBERG, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, **12**, 1-47.
- CABEZA, R., RAO, S. M., WAGNER, A. D., MAYER, A. R., & SCHACTER, D. L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences*, **98**, 4805-4810.
- CARTER, C. S., BRAVER, T. S., BARCH, D. M., BOTVINICK, M. M., NOLL, D., & COHEN, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, **280**, 747-749.
- COX, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computational Biochemical Research*, **29**, 162-173. Available at <http://afni.nimh.nih.gov>.
- DEESE, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58**, 17-22.
- FABIANI, M., STADLER, M. A., & WESSELS, P. M. (2000). True but not false memories produce a sensory signature in human lateralized brain potentials. *Journal of Cognitive Neuroscience*, **12**, 941-949.
- FARAH, M. J. (1989). The neural basis of mental imagery. *Trends in Neurosciences*, **12**, 395-399.
- GONSALVES, B., & PALLER, K. A. (2000). Neural events that underlie remembering something that never happened. *Nature Neuroscience*, **3**, 1316-1321.
- HENSON, R. N. A., RUGG, M. D., SHALLICE, T., JOSEPHS, O., & DOLAN, R. J. (1999). Recollection and familiarity in recognition memory: An event-related functional magnetic resonance imaging study. *Journal of Neuroscience*, **19**, 3962-3972.
- INSAUSTI, R., JUOTTONEN, K., SOININEN, H., INSAUSTI, A. M., PARTANEN, K., VAINIO, P., LAAKSO, M. P., & PITKANEN, A. (1998). MR volumetric analysis of the human entorhinal, perirhinal, and temporo-parietal cortices. *American Journal of Neuroradiology*, **19**, 659-671.
- JOHNSON, M. K. (1997). Source monitoring and memory distortion. *Philosophical Transactions of the Royal Society of London: Series B*, **352**, 1733-1745.
- JOHNSON, M. K., HASHTRUDI, 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.
- JOHNSON, M. K., & RAYE, C. L. (2000). Cognitive and brain mechanisms of false memories and beliefs. In D. L. Schacter & E. Scarry (Eds.), *Memory and belief* (pp. 36-86). Cambridge, MA: Harvard University Press.
- JOHNSON, M. K., RAYE, C. L., WANG, A. Y., & TAYLOR, T. H. (1979). Fact and fantasy: The roles of accuracy and variability in confusing imaginations with perceptual experiences. *Journal of Experimental Psychology: Human Learning & Memory*, **5**, 229-240.
- JOHNSON, M. K., TAYLOR, T. H., & RAYE, C. L. (1977). Fact and fantasy: The effects of internally generated events on the apparent frequency of externally generated events. *Memory & Cognition*, **5**, 116-122.
- KOSSLYN, S. M., PASCUAL-LEONE, A., FELICIAN, O., CAMPOSANO, S., KEENAN, J. P., THOMPSON, W. L., GANIS, G., SUKEL, K. E., & ALPERT, N. M. (1999). The role of Area 17 in visual imagery: Convergent evidence from PET and rTMS. *Science*, **284**, 167-170.
- LOFTUS, E. F., & LOFTUS, G. R. (1980). On the permanence of stored information in the human brain. *American Psychologist*, **35**, 409-420.
- LOFTUS, E. F., MILLER, D. G., & BURNS, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 19-31.

- LOFTUS, E. F., & PALMER, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning & Verbal Behavior*, **13**, 585-589.
- LOFTUS, E. F., & PICKRELL, J. E. (1995). The formation of false memories. *Psychiatric Annals*, **25**, 720-725.
- MCCLOSKEY, M., & ZARAGOZA, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, **114**, 1-16.
- MILLER, A. R., BARATTA, C., WYNVEEN, C., & ROSENFELD, J. P. (2001). P300 latency, but not amplitude or topography, distinguishes between true and false recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 354-361.
- MILNER, B., SQUIRE, L. R., & KANDEL, E. R. (1998). Cognitive neuroscience and the study of memory. *Neuron*, **20**, 445-468.
- NESSLER, D., MECKLINGER, A., & PENNEY, T. B. (2001). Event related brain potentials and illusory memories: The effects of differential encoding. *Cognitive Brain Research*, **10**, 283-301.
- PRUESSNER, J. C., KOHLER, S., CRANE, J., PRUESSNER, M., LORD, C., BYRNE, A., KABANI, N., COLLINS, D. L., & EVANS, A. C. (2002). Volumetry of temporopolar, perirhinal, entorhinal, and parahippocampal cortex from high-resolution MR images: Considering the variability of the collateral sulcus. *Cerebral Cortex*, **12**, 1342-1353.
- RANGANATH, C., JOHNSON, M. K., & D'ESPOSITO, M. (2000). Left anterior prefrontal activation increases with demands to recall specific perceptual information. *Journal of Neuroscience*, **20**, 1-5.
- ROEDIGER, H. L., III (1996). Memory illusions. *Journal of Memory & Language*, **35**, 76-100.
- ROEDIGER, H. L., III, & MCDERMOTT, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 803-814.
- RUGG, M. D., FLETCHER, P. C., FRITH, C. D., FRACKOWIAK, R. S. J., & DOLAN, R. J. (1996). Differential activation of the prefrontal cortex in successful and unsuccessful memory retrieval. *Brain*, **119**, 2073-2083.
- SCHACTER, D. L. (1999). The seven sins of memory. *American Psychologist*, **54**, 182-203.
- SCHACTER, D. L., BUCKNER, R. L., KOUTSTAAL, W., DALE, A. M., & ROSEN, B. R. (1997). Late onset of anterior prefrontal activity during true and false recognition: An event-related fMRI study. *NeuroImage*, **6**, 259-269.
- SCHACTER, D. L., NORMAN, K. A., & KOUTSTAAL, W. (1998). The cognitive neuroscience of constructive memory. *Annual Review of Psychology*, **49**, 289-318.
- SCHACTER, D. L., REIMAN, E., CURRAN, T., YUN, L. S., BANDY, D., MCDERMOTT, K. B., & ROEDIGER, H. L., III (1996). Neuroanatomical correlates of veridical and illusory recognition memory: Evidence from positron emission tomography. *Neuron*, **17**, 267-274.
- STARK, C. E. L., & OKADO, Y. (2003). Making memories without trying: Medial temporal lobe activity associated with incidental memory formation during recognition. *Journal of Neuroscience*, **23**, 6748-6753.
- STARK, C. E. L., & SQUIRE, L. R. (2000). Functional magnetic resonance imaging (fMRI) activity in the hippocampal region during recognition memory. *Journal of Neuroscience*, **20**, 7776-7781.
- STARK, C. E. L., & SQUIRE, L. R. (2001a). Simple and associative recognition memory in the hippocampal region. *Learning & Memory*, **8**, 190-197.
- STARK, C. E. L., & SQUIRE, L. R. (2001b). When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences*, **98**, 12760-12766.
- SUZUKI, W. A., & EICHENBAUM, H. (2000). The neurophysiology of memory. *Annals of the New York Academy of Sciences*, **911**, 175-191.
- TALAIRACH, J., & TOURNOUX, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.
- TULVING, E. (1985). Memory and consciousness. *Canadian Psychology*, **26**, 1-12.
- WAGNER, A. D., DESMOND, J. E., GLOVER, G. H., & GABRIELI, J. D. E. (1998). Prefrontal cortex and recognition memory: Functional-MRI evidence for context-dependent retrieval processes. *Brain*, **121**, 1985-2002.
- WARD, D. (2000). *Deconvolution analysis of FMRI time series data*. Available at <http://afni.nimh.nih.gov/afni/docpdf/3dDeconvolve.pdf>.
- WHEELER, M. E., & BUCKNER, R. L. (2003). Functional dissociation among components of remembering: Control, perceived oldness, and content. *Journal of Neuroscience*, **23**, 3869-3880.
- WHEELER, M. E., PETERSEN, S. E., & BUCKNER, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences*, **97**, 11125-11129.

(Manuscript received June 19, 2003;
revision accepted for publication December 12, 2003.)