

Not All False Memories Are Created Equal: The Neural Basis of False Recognition

Rachel J. Garoff-Eaton¹, Scott D. Slotnick² and Daniel L. Schacter¹

¹Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA and ²Department of Psychology, Boston College, McGuinn Hall, Chestnut Hill, MA 02467, USA

False recognition, a type of memory distortion where one claims to remember something that never happened, can occur in response to items that are similar but not identical to previously seen items (i.e., related false recognition) or in response to novel items (i.e., unrelated false recognition). It is unknown whether these 2 types of memory errors arise from the same or distinct neural substrates. Using functional magnetic resonance imaging, we compared the neural activity associated with true recognition, related false recognition, and unrelated false recognition for abstract shapes. True recognition and related false recognition were associated with similar patterns of neural activity, including activity in the prefrontal cortex, the parietal cortex, and the medial temporal lobe. By contrast, unrelated false recognition was associated with activity in language-processing regions. These results indicate that false recognition is not a unitary phenomenon, but rather can reflect the operation of 2 distinct cognitive and neural processes.

Keywords: false memory, fMRI, memory, retrieval, visual memory

Introduction

Memory is sometimes prone to distortions and errors that can have great theoretical and practical importance (Roediger 1996; Schacter 2001; Loftus 2003). A rapidly growing body of research has attempted to illuminate a common memory error known as false recognition, where people mistakenly claim that a novel object, face, word, or other type of stimulus, has been encountered during a specific prior episode (e.g., Roediger and McDermott 1995). False recognition is important theoretically because analysis of such errors can provide insights into basic mechanisms of encoding and retrieval (Schacter and others 1998), and it is important practically because these errors are often implicated in cases of mistaken eyewitness memory that can result in wrongful convictions (Wells and Olson 2003).

In the laboratory, people sometimes show false recognition after they have studied words or pictures that are semantically or perceptually related to a new item that appears on a subsequent test (i.e., related false alarms). Other false recognition errors, by contrast, occur in response to new items that are unrelated in any systematic way to previously studied items (i.e., unrelated false alarms). For instance, in the Deese-Roediger-McDermott (DRM) paradigm (Deese 1959; Roediger and McDermott 1995), participants study associatively related word lists (e.g., note, sound, piano, sing, radio, band, melody, horn, concert, instrument, symphony, jazz, orchestra, art, rhythm). On a subsequent recognition test, related false alarms occur when an "old" response is given to a semantically similar lure (e.g., music) and unrelated false alarms occur when an "old" response is given to a novel word that has no associative or semantic relation to the previously studied words (e.g., spider).

In addition to DRM-like paradigms that rely on associative word lists, related and unrelated false alarms can also be explored through studies that utilize categorical stimuli (e.g., Koutstaal and Schacter 1997; Seamon and others 2000) or perceptually similar stimuli (e.g., Schacter, Verfaellie, and Anes 1997; Koutstaal and others 1999; Slotnick and Schacter 2004).

Little is known about the relationship between related and unrelated false alarms. In standard signal detection models of memory, these 2 types of false alarms can be assumed to arise from a single underlying process that gives rise to memory strength or familiarity that is sufficient to surpass a participant's criterion for calling an item "old" (Miller and Wolford 1999; Slotnick and Dodson 2005). However, neuropsychological studies of amnesic patients reveal that, compared with healthy controls, amnesics sometimes show reduced false recognition of related lures together with increased false recognition of unrelated lures (e.g., Schacter, Verfaellie, and Pradere 1996; for review, see Schacter and Slotnick 2004). Because amnesics also show reduced true recognition of studied items, the overall pattern of results suggests that related and unrelated false recognition are based on distinct underlying processes, whereas similar mechanisms support both true recognition (i.e., hits) and related false alarms. However, such data are not necessarily inconsistent with some signal detection approaches to false recognition (e.g., Wixted and Stretch 2000). Thus, it is ambiguous from behavioral and neuropsychological research whether false recognition reflects a single, unitary phenomenon or multiple, distinct processes.

Much research has compared the neural activity associated with true and false recognition (e.g., Schacter, Reiman, and others 1996; Düzel and others 1997; Johnson and others 1997; Schacter, Buckner, and others 1997; Endl and others 1999; Fabiani and others 2000; Walla and others 2000; Cabeza and others 2001; Curran and Cleary 2003; Goldmann and others 2003; Nessler and Mecklinger 2003; Kahn and others 2004; Slotnick and Schacter 2004; for review, see Schacter and Slotnick 2004). However, these studies have not distinguished between the neural activity associated with different types of false alarms, owing largely to methodological limitations such as insufficient number of observations for analysis and failure to include both related and unrelated lures. To determine whether related and unrelated false recognition depend on the same or different underlying neural processes, we used event-related functional magnetic resonance imaging (fMRI) to compare the neural activity associated with true recognition, related false alarms, and unrelated false alarms. Participants studied a series of novel shapes and then were given a recognition test that contained 3 different types of shapes: 1) shapes that were the *same* as those studied previously; 2) shapes that were *similar* but not identical

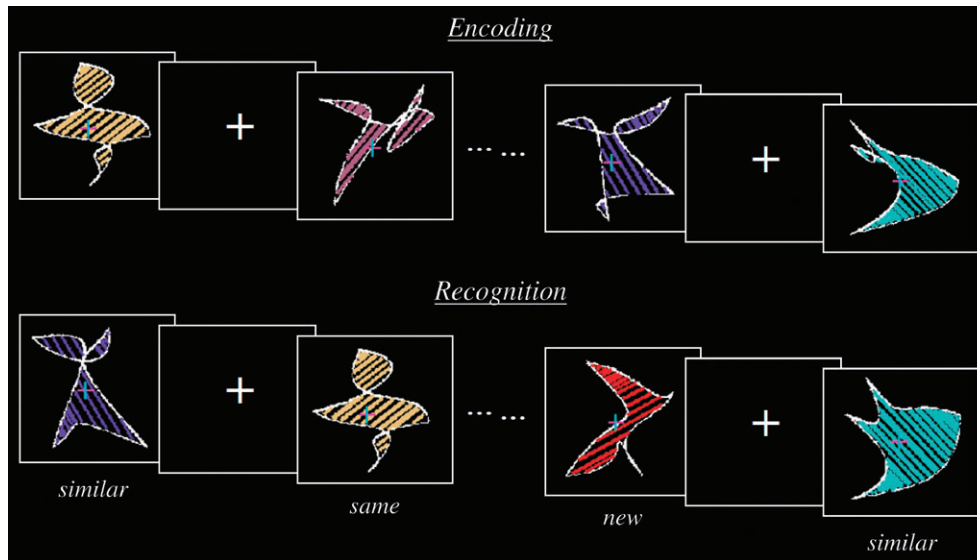


Figure 1. Schematic of experimental procedure. During the encoding phase, participants viewed a series of novel shapes and were instructed to decide whether the shapes were pleasant or unpleasant while also remembering the shapes for the subsequent recognition test. During the recognition phase, participants viewed 3 types of shapes: 1) shapes that were the *same* as shapes from the encoding phase; 2) shapes that were *similar* to shapes from the encoding phase; and 3) *new* shapes. Participants decided whether each shape was the “same,” “similar,” or “new.”

to previously studied shapes; and 3) *new* shapes that were unrelated to previously studied shapes (Fig. 1). Participants were asked to decide whether each shape was the “same,” “similar,” or “new.” In this paradigm, true recognition was defined as a “same” response to a *same* shape, related false recognition was defined as a “same” response to a *similar* shape, and unrelated false recognition was defined as a “same” response to a *new* shape. If the 2 types of false alarms reflect the activity of a single underlying mechanism, then predominantly common brain regions should be active during related and unrelated false recognition. On the other hand, if the 2 types of false alarms depend on different underlying mechanisms, largely distinct patterns of brain activity should be associated with related and unrelated false recognition.

Materials and Methods

Participants

Twenty right-handed native English speakers with normal or corrected-to-normal vision and no history of neurological trauma participated in the study. Nine participants were excluded from the analysis due to an insufficient number of unrelated false recognition responses (i.e., fewer than 8). Thus, reported results are based on the data from the remaining 11 participants (7 females, age range 18.5–26.4 years, mean age 22.2 years). The behavioral protocol was approved by the Harvard Institutional Review Board, and the imaging protocol was approved by the Massachusetts General Hospital Institutional Review Board. Informed consent was acquired from all participants prior to their participation in the study.

Procedure

Participants completed multiple 4 min 40 s runs, each consisting of a study and a test phase. During the study phase, participants viewed 20 novel shapes (Fig. 1). Each shape was presented in the center of the screen for 2.5 s with an intertrial interval (ITI) ranging from 4 to 16 s. Participants were instructed to respond whether each shape was “pleasant” or “unpleasant” and to remember each shape for the subsequent recognition test. After the study phase, instructions for the test phase appeared on the screen for 10 s. During the test phase, participants viewed 25 shapes (Fig. 1). Of these, 10 shapes were

identical to shapes presented in the preceding study phase (*same* shapes), 10 shapes were similar but not identical to shapes presented in the preceding study phase (*similar* shapes), and 5 shapes were unrelated to any of the shapes shown in the preceding study phase (*new* shapes). Each shape was presented in the center of the screen for 2.5 s with an ITI ranging from 4 to 16 s. Participants were asked to decide whether each shape was the “same,” “similar,” or “new” and to respond as quickly as possible without sacrificing accuracy. To ensure that the definition of each shape category was clear (particularly for *similar* shapes), participants were shown shape exemplars and trained on 1 complete study-test run prior to scanning.

Stimuli were generated using the procedure described by Slotnick and Schacter (2004). Four hundred and fifty prototype shapes were created using custom software written in MATLAB (The MathWorks, Inc, Natick, MA, USA). For each prototype, a similar shape was constructed by adding 75% Gaussian distributed noise to the end points and control points of the 4 Bezier curves that defined the prototype, while the identical internal line color and orientation were maintained. To eliminate any confusion due to overlapping internal line colors between shape pairs, 25 distinct internal line colors were used for each set of 25 unrelated shapes tested in a given run.

During the imaging session, participants completed between 10 and 14 study-test runs. A unique study list stimulus order was created within each run, and runs were randomly ordered with a unique time sequence for each participant. Stimuli were not repeated across runs. The same set of prototype shapes were always tested, whereas the condition in which a given prototype was tested (i.e., *same*, *similar*, or *new*) was counter-balanced across participants. Furthermore, during the study phase, no more than 3 shapes in a given condition at recognition (i.e., *same* or *similar*) were shown sequentially, and during the test phase, no more than 4 shapes in a given condition (i.e., *same*, *similar*, or *new*) were shown sequentially.

Image Acquisition and Data Analysis

Scanning was performed on a 3-T head-only Siemens Allegra MRI scanner. High-resolution anatomic images were acquired using a multi-planar rapidly acquired gradient echo sequence (repetition time (TR) = 30 ms, echo time (TE) = 3.3 ms, 128 slices, 1 × 1 × 1.33-mm voxels), and functional images were acquired using a T_2^* -weighted echo planar imaging (EPI) sequence (TR = 2000 ms, TE = 30 ms, 30 slices, 4.5-mm isotropic voxels). Participants viewed stimuli through an angled mirror attached to the head coil, which reflected images that were back-projected on a screen at the superior end of the scanner bore.

Unless otherwise noted, data analysis was conducted using SPM99 (Wellcome Department of Cognitive Neurology). For each participant, standard preprocessing was performed on the functional data. First, images were slice-time corrected and motion corrected. Then, using custom software written in MATLAB, a temporal high-pass filter was applied to each run, removing linear, quadratic, cubic, and quartic components, and all runs were concatenated. Finally, images were normalized to the SPM99 EPI template by resampling at 3-mm isotropic resolution and spatially smoothed using a 4.5-mm full-width half-maximum (FWHM) Gaussian kernel.

Statistical analyses were conducted using a general linear model. For each participant, a canonical hemodynamic response function was used to model activity associated with each event type at encoding and retrieval. After entering these hemodynamic response models into a general linear model, a beta-weight (i.e., model amplitude) associated with each event type was computed for each voxel. To identify those voxels that showed statistically consistent event-related activity across participants, beta-weight differences associated with 2 event types of interest were computed and compared using a 1-sample *t*-test (i.e., a random-effects analysis). Reported coordinates reflect the most significant voxel proximal to the center of each region of activity in Talairach space (Talairach and Tournoux 1988). For viewing purposes, functional activity was projected onto either a template 3-dimensional brain (SPM99) or a 2-dimensional mean, anatomic image.

All reported results were corrected for multiple comparisons with $P < 0.05$ via a voxel cluster extent threshold defined from Monte Carlo simulations (Forman and others 1995; Slotnick and others 2003; Slotnick and Schacter 2004; Garoff and others 2005). To conduct each simulation, activation at each voxel within the acquisition volume was modeled using normally distributed random noise (with zero mean and unit variance) and then thresholded to achieve the desired type I error rate (i.e., individual voxel P value). Spatial autocorrelation was modeled through convolution with a 4.5-mm FWHM Gaussian smoothing kernel (Aguirre and others 1997). The probability of observing each cluster extent size was computed across simulations (with smaller probabilities being associated with larger clusters). The cluster extent threshold was selected to yield the desired probability of observing that size cluster or any larger cluster by chance (i.e., to correct for multiple comparisons with $P < 0.05$). Thus, for single contrasts, we used an individual voxel threshold of $P < 0.01$, which corresponded to a cluster extent threshold of 14 resampled voxels (yielding results corrected for multiple comparisons with $P < 0.05$). When the conjunction of 2 contrasts is reported (i.e., activity jointly associated with 2 single contrasts), an individual voxel threshold of $P < 0.035$ and a cluster extent threshold of 14 resampled voxels were applied to each of the single contrasts. According to Fisher's technique of computing joint probability (Fisher 1973), by using an individual voxel threshold of $P < 0.035$ for each individual contrast, the joint probability of observing an active voxel is $P < 0.01$. Thus, the combination of the joint individual voxel threshold of $P < 0.01$ with a cluster extent threshold of 14 resampled voxels yields conjunction analysis results corrected for multiple comparisons (with $P < 0.05$) (Slotnick and Schacter 2004). In the case of the reported triple conjunction (i.e., activity associated with 3 single contrasts), an individual voxel threshold of $P < 0.06$ and a cluster extent threshold of 14 resampled voxels were applied to each of the 3 single contrasts used in the triple conjunction analysis, which again yielded results corrected for multiple comparisons (with $P < 0.05$).

To characterize the nature of the activity associated with the reported contrasts, event-related time courses were extracted from activated clusters of interest (for details, see Slotnick and Schacter 2004). Custom software written in MATLAB was used to remove linear trends from the event-related time courses, and the activity from 0 to 4 s before stimulus onset served as a measure of baseline activity. For clusters of activity that spanned multiple brain regions, event-related time courses were extracted from a 7-mm sphere within the activated region, and statistics were performed on the mean event-related activity from 4 to 8 s after stimulus onset, unless otherwise specified. Because these post hoc time course analyses were used to characterize, not define, significant event-related activity, 1-tailed paired *t*-tests were used to compare time course activity associated with different conditions.

Results

Behavioral Results

To address the question under investigation, we will focus on the results associated with the following 4 conditions: true recognition ("same"/*same*), related false recognition ("same"/*similar*), unrelated false recognition ("same"/*new*), and correct rejections ("new"/*new*). The rate of true recognition (0.59) was significantly higher than the rate of both related false recognition (0.31, $t = 6.56$, $P < 0.0001$, paired *t*-test) and unrelated false recognition (0.20, $t = 9.54$, $P < 0.0001$, paired *t*-test). Furthermore, the rate of related false recognition was significantly higher than unrelated false recognition ($t = 3.83$, $P < 0.005$, paired *t*-test); thus, differences between related, nonstudied shapes (i.e., *similar* shapes) and unrelated, nonstudied shapes (i.e., *new* shapes) did affect participants' responses, providing behavioral support for our distinction between related false recognition and unrelated false recognition. Finally, participants correctly rejected *new* shapes more often than they falsely accepted them as studied shapes (i.e., the proportion of *new* shapes labeled "new" [0.42] was greater than the proportion of *new* shapes labeled "same" [0.20, $t = 4.42$, $P < 0.005$, paired *t*-test]).

Although it could be argued that many "same" responses to *new* items (i.e., unrelated false recognition responses) reflected guessing in the current paradigm, potentially undermining the validity of the unrelated false recognition results, a pilot experiment that used a nearly identical protocol in addition to previously reported results (Garoff and others 2005) indicates that participants are highly confident of these responses. Specifically, in the behavioral pilot experiment, confidence ratings were collected from participants ($n = 9$), and the majority of "same" responses to *new* items was given with "high" rather than "low" confidence. Furthermore, in a previous fMRI study that used a similar method to study memory for visual objects rather than novel shapes (Garoff and others 2005), most unrelated false recognition responses were given with high confidence, again indicating that unrelated false memories do not merely reflect guessing in the current paradigm. Moreover, by giving participants the option of responding "similar," we aimed to minimize the number of "same" responses given to any shape (*same*, *similar*, or *new*) without a strong conviction that it had been studied previously.

Of particular relevance to interpreting the imaging results, the mean reaction times associated with related false recognition (mean = 1598 ms) and unrelated false recognition (mean = 1584 ms) did not differ ($t < 1$, $P > 0.2$, paired *t*-test), suggesting that these 2 types of memory errors did not require differential amounts of effort. The mean reaction time associated with true recognition (mean = 1415 ms) was significantly shorter than the mean reaction time associated with both related false recognition ($t = 5.31$, $P < 0.0005$, paired *t*-test) and unrelated false recognition ($t = 3.67$, $P < 0.005$, paired *t*-test). Furthermore, the mean reaction time associated with correct rejections (mean = 1610 ms) did not differ from the mean reaction time associated with unrelated false recognition ($t = 1.46$, $P = 0.18$, paired *t*-test).

Imaging Results

To identify neural activity common to both true and related false recognition, but not unrelated false recognition, we examined the conjunction of the true recognition > unrelated

false recognition and the related false recognition > unrelated false recognition contrasts. This analysis revealed widespread neural activity (Table 1, Fig. 2), including prefrontal cortex, anterior cingulate cortex, inferior parietal cortex, medial parietal cortex, medial temporal lobe, ventral temporal cortex, and occipital cortex activity. To characterize the nature of this activity, time courses were extracted from regions that have previously been implicated in memory retrieval (for reviews, see Buckner and Wheeler 2001; Slotnick and others 2003) (Fig. 2). For instance, the right parahippocampal gyrus revealed significant increases in activity associated with both true recognition ($t = 2.13$, $P < 0.03$) and related false recognition ($t = 2.44$, $P < 0.02$) as well as a significant decrease in activity associated with unrelated false recognition ($t = 3.45$, $P < 0.005$). Accordingly, significantly greater activity was associated with true recognition and related false recognition than with unrelated false recognition in this region (true recognition > unrelated false recognition, $t = 3.80$, $P < 0.005$; related false recognition > unrelated false recognition, $t = 4.19$, $P < 0.001$), whereas there was no difference in activity between true and related false recognition ($t < 1$, $P > 0.2$; Fig. 2, bottom right). The same patterns of activity were observed in the left middle frontal gyrus (Fig. 2, top left; true recognition increase, $t = 2.33$, $P < 0.03$; related false recognition increase, $t = 2.81$, $P < 0.02$; unrelated false recognition decrease, $t = 2.04$, $P < 0.04$; true recognition > unrelated false recognition, $t = 2.83$, $P < 0.01$; related false recognition > unrelated false recognition, $t = 4.44$, $P < 0.001$; true recognition > related false recognition, $t < 1$, $P > 0.2$) and right hippocampus (Fig. 2; top right; true recognition increase, $t = 2.79$, $P < 0.01$; related false recognition increase,

$t = 3.98$, $P < 0.005$; unrelated false recognition decrease, $t = 3.27$, $P < 0.005$; true recognition > unrelated false recognition, $t = 3.68$, $P < 0.005$; related false recognition > unrelated false recognition, $t = 5.02$, $P < 0.0005$; true recognition > related false recognition, $t < 1$, $P > 0.1$). The time courses from these 3 regions of interest (ROIs) reflect the general pattern of results observed; however, there were a few regions where the significant difference between true or related false recognition and unrelated false recognition was driven by a significant decrease in unrelated false recognition activity coupled with no significant change in activity associated with true and related false recognition (e.g., the precuneus, Fig. 2, bottom left; true recognition increase, $t < 1$, $P > 0.2$; related false recognition increase, $t < 1$, $P > 0.2$; unrelated false recognition decrease, $t = 4.31$, $P < 0.001$; true recognition > unrelated false recognition, $t = 4.15$, $P < 0.005$; related false recognition > unrelated false recognition, $t = 5.27$, $P < 0.0005$; true recognition > related false recognition, $t < 1$, $P > 0.2$). The exact cause of event-related decreases is uncertain (cf., Gusnard and Raichle 2001; McKiernan and others 2003); still, such activity relates to the hypotheses under scrutiny in that it can be taken to reflect the common neural substrates associated with true and related false recognition as distinct from unrelated false recognition. It should also be mentioned that a similar pattern of results (i.e., widespread neural activity in areas including the prefrontal cortex, parietal cortex, and medial temporal lobe) was observed in the direct contrast between related and unrelated false recognition.

Next, to identify neural activity associated with unrelated false recognition, but neither true nor related false recognition,

Table 1
Neural regions associated with both true recognition (labeling a *same* shape "same") and related false recognition (labeling a *similar* shape "same") as compared with unrelated false recognition (labeling a *new* shape "same")

Region	BA	(x, y, z)	Region	BA	(x, y, z)
Right superior frontal gyrus	8	23, 23, 51	Left precuneus	7	-11, -40, 49
Left superior frontal gyrus	8	-9, 48, 44	Left precuneus	19	-36, -71, 41
Left middle frontal gyrus	8	-24, 28, 41	Left cuneus	18	-19, -67, 13
Right medial frontal gyrus	8	3, 38, 41	Left cuneus	17	-5, -79, 11
Left medial frontal gyrus	8	-3, 38, 39	Right supramarginal gyrus	40	53, -57, 35
Left superior frontal gyrus	6	-24, 11, 51	Left supramarginal gyrus	40	-56, -51, 40
Right middle frontal gyrus	6	25, 14, 54	Left angular gyrus	39	-47, -61, 36
Left middle frontal gyrus	6	-20, 11, 58	Right hippocampus	—	30, -21, -10
Right medial frontal gyrus	6	3, 23, 47	Right parahippocampal gyrus	35	21, -18, -10
Right precentral gyrus	6	33, -12, 60	Right parahippocampal gyrus	19	30, -49, -6
Left precentral gyrus	6	-41, -15, 58	Right fusiform gyrus	19	25, -53, -8
Right precentral gyrus	4	38, -19, 48	Left superior temporal gyrus	39	-56, -61, 23
Left postcentral gyrus	3	-40, -23, 55	Left superior temporal gyrus	38	-37, 5, -16
Left medial frontal gyrus	9	-13, 37, 21	Left middle temporal gyrus	21	-63, -46, -2
Left inferior frontal gyrus	9	-47, 11, 28	Right middle temporal gyrus	22	52, -43, 1
Right inferior frontal gyrus	44	57, 9, 14	Right middle temporal gyrus	37	53, -62, 6
Left inferior frontal gyrus	44	-50, 11, 21	Right inferior temporal gyrus	37	59, -56, -7
Right inferior frontal gyrus	47	37, 20, -13	Left middle occipital gyrus	37	-53, -70, 3
Left inferior frontal gyrus	47	-36, 15, -9	Left superior occipital gyrus	19	-36, -75, 26
Right anterior cingulate	32	3, 34, 25	Right middle occipital gyrus	19	51, -71, -5
Left anterior cingulate	32	-8, 37, 22	Left middle occipital gyrus	19	-49, -74, 3
Right anterior cingulate	24	1, 23, 19	Right caudate (head)	—	18, 19, 4
Left anterior cingulate	24	-2, 23, 19	Left caudate (head)	—	-14, 13, -1
Right cingulate gyrus	24	5, 4, 35	Right caudate	—	21, -7, 21
Left cingulate gyrus	24	-2, -12, 36	Left caudate	—	-15, -5, 19
Left cingulate gyrus	31	-12, -34, 41	Right thalamus	—	14, -25, 6
Left posterior cingulate	30	-21, -64, 10	Left thalamus	—	-12, -19, 5
Right insula	13	32, 17, -4	Left globus pallidus	—	-13, -5, 4
Left insula	13	-27, 18, -6	Right putamen	—	31, -12, 6
Right paracentral lobule	5	9, -38, 50	Left putamen	—	-25, -10, 3
Left paracentral lobule	5	-2, -38, 62	Right cerebellum	—	12, -57, -28
Left precuneus	31	-8, -46, 35	Left cerebellum	—	-8, -61, -20

Note: Coordinates (x, y, z) are reported in Talairach space.

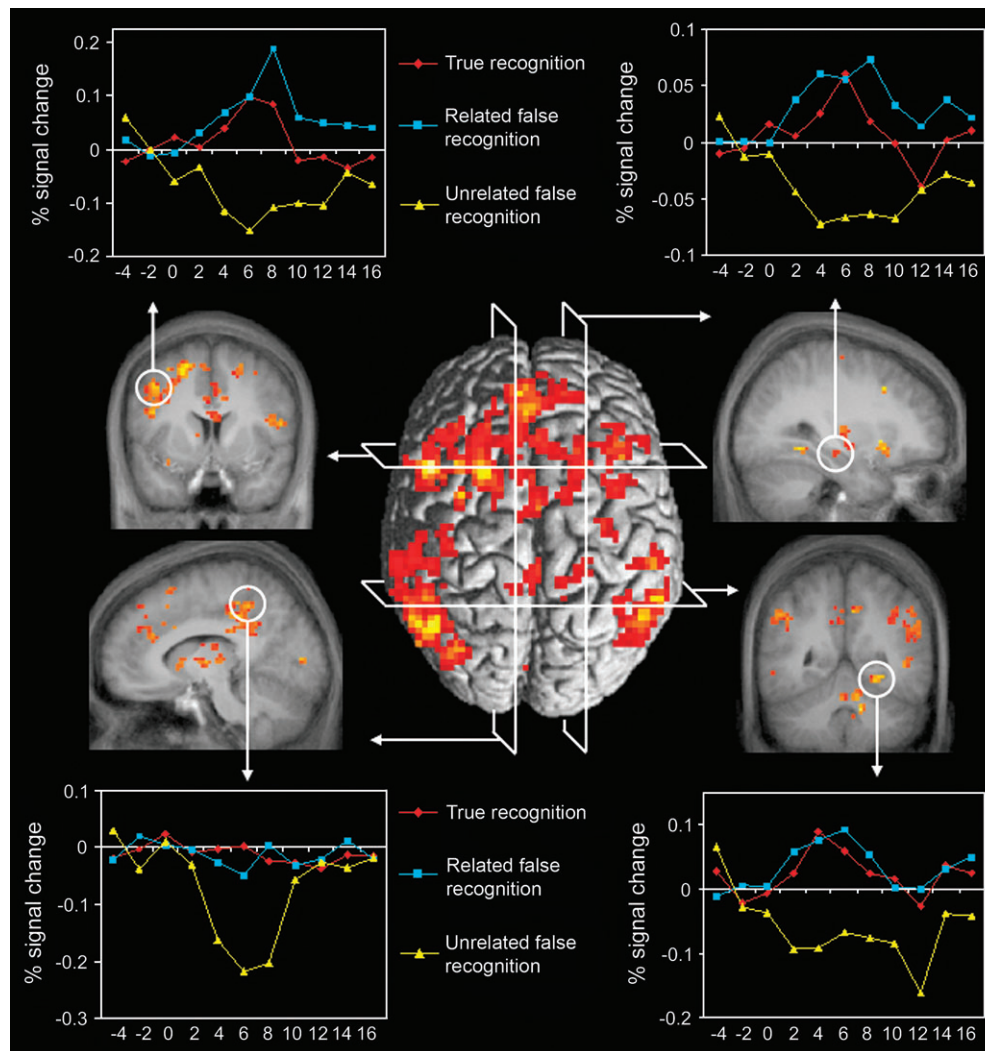


Figure 2. Neural activity associated with both true recognition (labeling a *same* shape “same”) and related false recognition (labeling a *similar* shape “same”) as compared with unrelated false recognition (labeling a *new* shape “same”). In the center of the figure, functional activity associated with this conjunction analysis (true recognition > unrelated false recognition and related false recognition > unrelated false recognition) is projected onto the dorsal surface of a 3-dimensional template brain. In each corner, functional activity is projected onto either a coronal or a sagittal slice of the group mean anatomic image, where an ROI is demarcated by a circle on each slice, and the corresponding event-related activity time courses associated with that ROI are shown (upper-left corner, left middle frontal gyrus; lower-left corner, left precuneus; upper-right corner, right hippocampus; lower-right corner, right parahippocampal gyrus).

we conducted the reverse of the above conjunction analysis, examining the common activity associated with the unrelated false recognition > true recognition and the unrelated false recognition > related false recognition contrasts (Table 2, Fig. 3). Only 2 clusters of activity were associated with this conjunction analysis, both spanning the left superior and middle temporal gyri, with 1 cluster in Brodmann area (BA) 22 and the other in BA38; these regions have previously been associated with language processing (Price 2000). Time courses extracted from the cluster of activity in BA22 revealed a significant increase in activity associated with unrelated false recognition ($t = 2.59$, $P < 0.02$), a significant decrease in activity associated with true recognition ($t = 4.19$, $P < 0.001$), and no significant change in activity associated with related false recognition ($t < 1$, $P > 0.2$). Furthermore, unrelated false recognition was associated with greater activity in this region than both true recognition ($t = 3.65$, $P < 0.005$) and related false recognition ($t = 2.82$, $P < 0.01$). In order to ensure that the observed activity increases associated with unrelated false recognition were not

Table 2

Neural regions associated with unrelated false recognition (labeling a *new* shape “same”) as compared with both true recognition (labeling a *same* shape “same”) and related false recognition (labeling a *similar* shape “same”)

Region	BA	(x, y, z)
Left superior temporal gyrus	22	-60, -34, 7
Left middle temporal gyrus	22	-60, -36, 5
Left superior temporal gyrus	38	-39, 7, -29
Left middle temporal gyrus	38	-40, 6, -27

Note: Coordinates (x, y, z) are reported in Talairach space.

merely common to all *new* shapes regardless of response type (e.g., due to novel shape processing), we compared the activity in this region associated with unrelated false recognition (“same”/*new*) with that associated with correct rejections (“new”/*new*). Whereas unrelated false recognition was associated with an activity increase, no increase in activity was associated with correct rejections ($t < 1$, $P > 0.2$); the difference

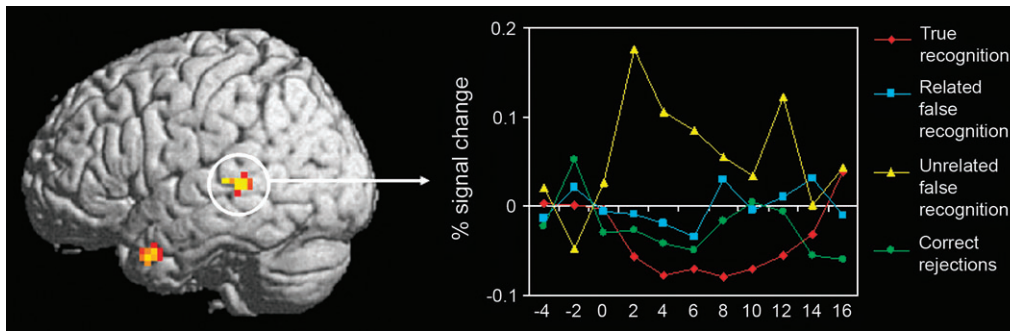


Figure 3. Neural activity associated with unrelated false recognition (labeling a *new* shape “same”) as compared with both true recognition (labeling a *same* shape “same”) and related false recognition (labeling a *similar* shape “same”). To the left, functional activity associated with this conjunction analysis (unrelated false recognition > true recognition and unrelated false recognition > related false recognition) is projected onto the left hemisphere of a 3-dimensional template brain. A cluster of activity in the left superior/middle temporal gyrus (BA22) is demarcated by a circle, and to the right, the event-related activity time courses associated with that region are shown.

in activity between these 2 conditions was significant ($t = 2.16$, $P < 0.03$). These results indicate that the activity increases associated with unrelated false recognition were not simply due to the presentation of *new* items. It should be noted that the direct comparison between unrelated and related false recognition yielded the same pattern of results (i.e., activity in the superior/middle temporal cortex only).

To determine whether there were any brain regions associated with both types of memory errors (i.e., related and unrelated false recognition), but not true recognition, we examined the union of the related false recognition > true recognition and the unrelated false recognition > true recognition contrasts. No activity was associated with this conjunction analysis. To examine whether such activity might be subthreshold, the cluster extent threshold was relaxed to reflect nonsignificant activity ($P = 0.2$); still, we failed to find common activity associated with these 2 types of memory errors.

Finally, we conducted a triple conjunction analysis in order to determine whether any activity was associated with giving a “same” response to any shape (i.e., true recognition, related false recognition, and unrelated false recognition), as compared with correct rejections. This analysis revealed only 2 small clusters of activity in the left cerebellum ($x = -7$, $y = -39$, $z = -10$; $x = -22$, $y = -48$, $z = -20$). Again, we relaxed the cluster extent threshold to reflect nonsignificant activity ($P = 0.2$), and only 1 additional cluster of activity in the right cerebellum ($x = 10$, $y = -41$, $z = -9$) was revealed.

Discussion

Results from the current study indicate that distinct patterns of neural activity are associated with related and unrelated false recognition. Whereas numerous brain regions were active during both true and related false recognition, as compared with unrelated false recognition, no neural activity was commonly associated with related and unrelated false recognition as compared with true recognition. Importantly, there was no difference in the response latencies associated with the 2 types of false alarms, which indicates that the observed neural differences were not due to differential effort associated with these 2 memory errors. Our results complement previous neuropsychological studies which have suggested that related and unrelated false recognition may rely on distinct underlying processes because amnesics show similar decreases in true and

related false recognition rates but increases in unrelated false recognition rates (e.g., Schacter, Verfaellie, and Pradere 1996; for review, see Schacter and Slotnick 2004). The present study is the first, to our knowledge, to identify the precise neural substrates supporting these 2 types of false recognition.

To identify the specific neural substrates underlying unrelated false recognition, we compared unrelated false recognition with both true and related false recognition which revealed only 2 clusters of activity in the left superior/middle temporal cortex (BA22 and BA38; Table 2, Fig. 3). Thus, the only neural activity differentially associated with unrelated false recognition occurred in brain regions commonly associated with language processing (Price 2000). Although the stimuli used in this study were designed to be novel nonverbalizable shapes, some participants reported applying verbal labels to those stimuli that reminded them of an object or animal (e.g., “flower” or “bird”). It is plausible that unrelated false recognition responses (i.e., “same” responses to *new* shapes) could result from the application of a verbal label to a *new* shape on the recognition test that matched a verbal label that had been applied to a different shape during encoding. Thus, many “same” responses to *new* shapes could be due to this verbal matching between labels applied at study and test, which could explain the activity observed in regions often associated with language processing. Moreover, this process would be distinct from the types of processes that most likely lead to true and related false recognition (e.g., perceptual feature matching, familiarity for visual aspects of the shapes).

To evaluate the generalizability of our results and explore our verbal labeling hypothesis, we conducted a post hoc analysis that included all 20 participants (i.e., those included and those excluded from the random-effects analyses). It is important to note, though, that this post hoc analysis should be considered with caution because of the small number of unrelated false recognition responses made by several participants. Across participants, a activity associated with unrelated false recognition within the left superior/middle temporal gyrus (BA22; a region often associated with language processing) was correlated with the proportion of unrelated false recognition responses ($r = 0.37$, $P = 0.05$, 1-tailed Pearson correlation). This correlation suggests that our results may generalize across the entire sample of subjects. Furthermore, this relationship illustrates that participants who made a larger number of unrelated false alarms also recruited a greater amount of activity from regions commonly associated with language processing, which

could reflect greater reliance on verbal strategies during unrelated false recognition for those participants who make many unrelated false alarms. Although the regions of left superior/middle temporal cortex associated with unrelated false recognition in the present study are commonly associated with verbal processing, it should be noted that activity in these regions can be observed in cases where it is not clear that language is involved. Thus, further research will be needed to test our hypothesis that unrelated false recognition for novel visual information reflects a verbal matching strategy.

In contrast to the unique activity associated with unrelated false recognition, the neural correlates of related false recognition were strikingly similar to those associated with true recognition. Specifically, neural activity common to both true and related false recognition (but not unrelated false recognition) was observed in numerous regions previously associated with successful retrieval (e.g., medial temporal lobe, prefrontal cortex, parietal cortex, and occipital cortex; Buckner and others 1998; Buckner and Wheeler 2001; Slotnick and others 2003). In the case of related false recognition, this pattern of activity probably reflects successful retrieval of the general characteristics of the originally studied shape (e.g., basic shape outline) and/or of the specific perceptual features that overlapped between the studied shape and the *similar*, tested shape (e.g., line-color/orientation). Previous studies have also reported similarities in the neural correlates of true and false recognition (e.g., Schacter, Reiman, and others 1996; Cabeza and others 2001; Slotnick and Schacter 2004; for review, see Schacter and Slotnick 2004); however, as discussed earlier, these studies did not distinguish between different types of false recognition. Furthermore, to identify neural activity common to both true and related false recognition, past neuroimaging studies have contrasted each of these memory processes with either passive fixation (e.g., Schacter, Reiman, and others 1996; Schacter, Buckner, and others 1997) or correct rejections (e.g., Düzel and others 1997; Cabeza and others 2001; Slotnick and Schacter 2004). Thus, both behavioral response type (i.e., no response or “new”) and item type (i.e., fixation or *new* item) differed across the conditions being contrasted. In the current paradigm, behavioral response was held constant; in all compared conditions, participants responded “same,” whereas item type varied (i.e., *same* for true recognition, *similar* for related false recognition, and *new* for unrelated false recognition). Thus, our results extend previous reports of similarities between true and false recognition by illustrating that the observed commonalities are not merely associated with responding “same” (or “old”; i.e., true and all false recognition), but rather are specific to true and related false recognition (and not unrelated false recognition).

Previous research has suggested that certain patterns of neural activity, specifically activity increases in the lateral parietal cortex, are associated with responding “old” on a recognition test (i.e., perceiving an item as old regardless of whether it has actually been studied; Wheeler and Buckner 2003, 2004). Thus, one might assume that, in the current study, there should be some similarity in the neural processes associated with giving a “same” response regardless of the item type (e.g., *same*, *similar*, or *new*). However, we found no common cortical activity associated with giving a “same” response to a shape of any type (i.e., true recognition, related false recognition, and unrelated false recognition) as compared with correctly rejecting a shape (i.e., giving a “new” response to

a *new* shape). To examine activity in the region previously associated with “perceived oldness” by Wheeler and Buckner (2003, 2004), we conducted a post hoc analysis by extracting event-related time course activity 6 s after stimulus onset from a 9-mm sphere centered on the reported central coordinate of their left lateral parietal ROI (BA 40/39, $x = -39$, $y = -55$, $z = 36$). Within this ROI, both true and related false recognition were associated with significant increases in activity ($t = 5.12$, $P < 0.0005$, and $t = 2.83$, $P < 0.01$, respectively), whereas there was no change in activity associated with unrelated false recognition ($t < 1$, $P > 0.2$). Furthermore, although there was no difference in the activity associated with true and related false recognition in this ROI ($t < 1$, $P > 0.2$), true recognition was associated with significantly greater activity than unrelated false recognition ($t = 1.96$, $P < 0.04$), and related false recognition was also associated with greater activity than unrelated false recognition, although this effect was only marginally significant ($t = 1.51$, $P < 0.09$). Thus, in our paradigm, activity in this left lateral parietal cortex ROI was not broadly associated with responding “same,” as no significant activity in this region accompanied unrelated false recognition responses (i.e., “same” responses to *new* shapes). Rather, activity in this region may reflect perceived perceptual overlap with studied shapes (or a combination of perceived oldness and perceptual overlap), which would only lead to a “same” response in the case of a *same* or a *similar* shape. By contrast, it seems that the neural processes that trigger a “same” response in the case of *new* shapes are distinct and may be related to processes associated with verbal matching (as discussed above) rather than perceived perceptual overlap.

Previous neuropsychological work has suggested that varying types of false recognition may be supported by different underlying cognitive processes. In the current study, we have provided the first evidence that related false recognition and unrelated false recognition can rely on distinct neural substrates. Specifically, related false recognition was associated with patterns of neural activity that were largely similar to those associated with true recognition, which may reflect successful retrieval of features that overlap with a previously studied shape in this paradigm. These results are consistent with previous studies comparing true and related false recognition for other classes of stimuli (e.g., Schacter, Reiman, and others 1996; Cabeza and others 2001; for review, see Schacter and Slotnick 2004), suggesting that these neural similarities are not unique to memory for novel shapes. By contrast, unrelated false recognition appeared to result from the employment of verbal rather than perceptual retrieval strategies in this study, as evidenced by activity increases in regions commonly associated with language processing. It should be noted that this particular relationship between unrelated false recognition and language-processing-related activity may be unique to visual stimuli; further research will be needed to explore the neural activity accompanying unrelated false alarms for other types of stimuli (e.g., words). In our study, unrelated false recognition showed no neural similarity to true or related false recognition, suggesting that related and unrelated false recognition are supported by distinct patterns of neural activity. Thus, the present neuroimaging results, along with the findings from previous neuropsychological studies, indicate that false recognition should not be considered a unitary process; rather, it can reflect the operation of at least 2 qualitatively different cognitive and neural processes.

Notes

We thank Jacqueline O'Brien and Jennifer Lai for their help with data management. This research was supported by grants NIMH MH60941 and NIA AG08441 (to DLS) and a National Science Foundation Graduate Research Fellowship (to RJG-E).

Address correspondence to Rachel J. Garoff-Eaton, Department of Psychology, William James Hall, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA. Email: garoff@fas.harvard.edu.

References

- Aguirre GK, Zarahn E, D'Esposito M. 1997. Empirical analyses of BOLD fMRI statistics II. Spatially smoothed data collected under null-hypothesis and experimental conditions. *Neuroimage* 5:199-212.
- Buckner RL, Koutstaal W, Schacter DL, Wagner AD, Rosen BR. 1998. Functional-anatomic study of episodic retrieval using fMRI. *Neuroimage* 7:151-162.
- Buckner RL, Wheeler ME. 2001. The cognitive neuroscience of remembering. *Nat Rev Neurosci* 2:624-634.
- Cabeza R, Rao SM, Wagner AD, Mayer AR, Schacter DL. 2001. Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proc Natl Acad Sci USA* 98:4805-4810.
- Curran T, Cleary AM. 2003. Using ERPs to dissociate recollection from familiarity in picture recognition. *Cogn Brain Res* 15:191-205.
- Curran T, Schacter DL, Johnson MK, Spinks R. 2001. Brain potentials reflect behavioral differences in true and false recognition. *J Cogn Neurosci* 13:201-216.
- Deese J. 1959. On the prediction of occurrence of particular verbal intrusions in immediate recall. *J Exp Psychol* 58:17-22.
- Düzel E, Yonelinas AP, Mangun GR, Heinze HJ, Tulving E. 1997. Event-related brain potential correlates of two states of conscious awareness in memory. *Proc Natl Acad Sci USA* 94:5973-5978.
- Endl W, Walla P, Lindinger G, Deecke L, Lang W. 1999. Event-related potential correlates of false recognitions of faces. *Neurosci Lett* 265:115-118.
- Fabiani M, Stadler MA, Wessels PM. 2000. True but not false memories produce a sensory signature in human lateralized brain potentials. *J Cogn Neurosci* 12:941-949.
- Fisher RA. 1973. *Statistical methods for research workers*, 14th edn. New York: Hafner.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med* 33:636-647.
- Garoff RJ, Slotnick SD, Schacter DL. 2005. The neural origins of specific and general memory: the role of the fusiform cortex. *Neuropsychologia* 43:847-859.
- Goldmann RE, Sullivan AL, Droller DB, Rugg MD, Curran T, Holcomb PJ, Schacter DL, Daffner KR, Budson AE. 2003. Late frontal brain potentials distinguish true and false recognition. *Neuroreport* 14:1717-1720.
- Gusnard DA, Raichle ME. 2001. Searching for a baseline: functional imaging and the resting human brain. *Nat Rev Neurosci* 2:685-694.
- Johnson MK, Nolde SF, Mather M, Kounios J, Schacter DL, Curran T. 1997. The similarity of brain activity associated with true and false recognition memory depends on test format. *Psychol Sci* 8:250-257.
- Kahn I, Davachi L, Wagner AD. 2004. Functional-neuroanatomic correlates of recollection: implications for models of recognition memory. *J Neurosci* 24:4172-4180.
- Koutstaal W, Schacter DL. 1997. Gist-based false recognition of pictures in older and younger adults. *J Mem Lang* 37:555-583.
- Koutstaal W, Schacter DL, Verfaellie M, Brenner C, Jackson EM. 1999. Perceptually based false recognition of novel objects in amnesia: effects of category size and similarity to category prototypes. *Cogn Neuropsychol* 16:317-341.
- Loftus EL. 2003. Make-believe memories. *Am Psychol* 58:518-537.
- McKiernan KA, Kaufman JN, Kucera-Thompson J, Binder JR. 2003. A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *J Cogn Neurosci* 15:394-408.
- Miller MB, Wolford GL. 1999. The role of criterion shift in false memory. *Psychol Rev* 106:398-405.
- Nessler D, Mecklinger A. 2003. ERP correlates of true and false recognition after different retention delays: stimulus- and response-related processes. *Psychophysiology* 40:146-159.
- Price CJ. 2000. The anatomy of language: contributions from functional neuroimaging. *J Anat* 197:335-359.
- Roediger HL. 1996. Memory illusions. *J Mem Lang* 35:76-100.
- Roediger HL III, McDermott KB. 1995. Creating false memories: remembering words not presented on lists. *J Exp Psychol Learn Mem Cogn* 21:803-814.
- Schacter DL. 2001. *The seven sins of memory: how the mind forgets and remembers*. Boston: Houghton Mifflin.
- Schacter DL, Buckner RL, Koutstaal W, Dale AM, Rosen BR. 1997. Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. *Neuroimage* 6:259-269.
- Schacter DL, Norman KA, Koutstaal W. 1998. The cognitive neuroscience of constructive memory. *Annu Rev Psychol* 49:289-318.
- Schacter DL, Reiman E, Curran T, Yun LS, Bandy D, McDermott KB, Roediger HL III. 1996. Neuroanatomical correlates of veridical and illusory recognition memory: evidence from positron emission tomography. *Neuron* 17:267-274.
- Schacter DL, Slotnick SD. 2004. The cognitive neuroscience of memory distortion. *Neuron* 44:1-20.
- Schacter DL, Verfaellie M, Anes MD. 1997. Illusory memories in amnesic patients: conceptual and perceptual false recognition. *Neuropsychology* 11:331-342.
- Schacter DL, Verfaellie M, Pradere D. 1996. The neuropsychology of memory illusions: false recall and recognition in amnesic patients. *J Mem Lang* 35:319-334.
- Seamon JG, Luo CR, Schlegel SE, Greene SE, Goldenberg AB. 2000. False memory for categorized pictures and words: the category associates procedure for studying memory errors in children and adults. *J Mem Lang* 42:120-146.
- Slotnick SD, Dodson CS. 2005. Support for a continuous (single-process) model of recognition memory and source memory. *Mem Cognit* 33:151-170.
- Slotnick SD, Moo LR, Segal JB, Hart J Jr. 2003. Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Brain Res Cogn Brain Res* 17:75-82.
- Slotnick SD, Schacter DL. 2004. A sensory signature that distinguishes true from false memories. *Nat Neurosci* 7:664-672.
- Talairach J, Tournoux P. 1988. *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.
- von Zerssen GC, Mecklinger A, Opitz B, von Cramon DY. 2001. Conscious recollection and illusory recognition: an event-related fMRI study. *Eur J Neurosci* 13:2148-2156.
- Walla P, Endl W, Lindinger G, Deecke L, Lang W. 2000. False recognition in a verbal memory task: an event-related potential study. *Cogn Brain Res* 9:41-44.
- Wells GL, Olson EA. 2003. Eyewitness testimony. *Annu Rev Psychol* 54:277-295.
- Wheeler ME, Buckner RL. 2003. Functional dissociation among components of remembering: control, perceived oddness, and content. *J Neurosci* 23:3869-3880.
- Wheeler ME, Buckner RL. 2004. Functional-anatomic correlates of remembering and knowing. *Neuroimage* 21:1337-1349.
- Wixted JT, Stretch V. 2000. The case against a criterion-shift account of false memory. *Psychol Rev* 107:368-376.