



# fMRI Evidence for the Role of Recollection in Suppressing Misattribution Errors: The Illusory Truth Effect

## Citation

Mitchell, Jason Paul, Chad S. Dodson, and Daniel L. Schacter. 2005. fMRI evidence for the role of recollection in suppressing misattribution errors: The illusory truth effect. *Journal of Cognitive Neuroscience* 17(5): 800-810.

## Published Version

doi:10.1162/0898929053747595

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:3627267>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available. Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# fMRI Evidence for the Role of Recollection in Suppressing Misattribution Errors: The Illusory Truth Effect

Jason P. Mitchell<sup>1,2</sup>, Chad S. Dodson<sup>3</sup>, and Daniel L. Schacter<sup>1</sup>

## Abstract

■ Misattribution refers to the act of attributing a memory or idea to an incorrect source, such as successfully remembering a bit of information but linking it to an inappropriate person or time [Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology*, 56, 326–338; Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, 54, 182–203; Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and remembers*. Boston: Houghton Mifflin]. Cognitive studies have suggested that misattribution errors may occur in the absence of recollection for the details of an initial encounter with a stimulus, but little is known about the neural basis of

this memory phenomenon. Here we used functional magnetic resonance imaging (fMRI) to examine the hypothesized role of recollection in counteracting the illusory truth effect, a misattribution error whereby perceivers systematically overrate the truth of previously presented information. Imaging was conducted during the encoding and subsequent judgment of unfamiliar statements that were presented as true or false. Event-related fMRI analyses were conditionalized as a function of subsequent performance. Results demonstrated that encoding activation in regions previously associated with successful recollection—including the hippocampus and the ventrolateral prefrontal cortex (PFC)—correlated with the successful avoidance of misattribution errors, providing initial neuroimaging support for earlier cognitive accounts of misattribution. ■

## INTRODUCTION

Event-related fMRI has provided a new way to examine the neural correlates of episodic encoding by allowing examination of the brain regions in which activity predicts the memorial fate of individual items. In such subsequent memory designs (first developed in electrophysiological research; see Paller & Wagner, 2002, for discussion), brain activation during the encoding of subsequently remembered items is compared to activation during the encoding of subsequently forgotten items, allowing the identification of brain regions in which processing correlates with the successful formation of episodic memories on a trial-by-trial basis. This research has consistently identified regions of the ventrolateral prefrontal cortex (PFC) and medial-temporal lobe (MTL) as contributing to memory formation (Davachi, Mitchell, & Wagner, 2003; Davachi & Wagner, 2002; Paller & Wagner, 2002; Strange, Otten, Josephs, Rugg, & Dolan, 2002; Otten, Henson, & Rugg, 2001; Kirchoff, Wagner, Maril, & Stern, 2000; Henson, Rugg, Shallice, Josephs, & Dolan, 1999; Brewer, Zhao, Des-

mond, Glover, & Gabrieli, 1998; Wagner, Schacter, et al., 1998).

However, as pointed out by a number of cognitive theorists (Schacter, 1999, 2001; Greenwald & Banaji, 1995; Jacoby, Kelley, Brown, & Jasechko, 1989), memory is more than just a simple matter of remembering or forgetting past events. Schacter (1999, 2001) has recently discussed a number of additional ways in which memory can prove fragile, from various kinds of memory distortions to persistent memories of traumatic events. Although relatively little is currently known about the neural processes that either give rise to or prevent such memory failures, researchers have very recently begun to adapt the basic subsequent memory paradigm to investigate the functional neuroanatomy of these phenomena (Anderson, Ochsner, et al., 2004; Maril, Simons, Mitchell, Schwartz, & Schacter, 2003; Maril, Wagner, & Schacter, 2001).

In the current research, we extend the scope of these earlier investigations to another form of memory failure, that of *misattribution*. Misattribution refers to situations in which one erroneously attributes a memory or idea to an incorrect source, such as successfully remembering a bit of information but linking it to an inappropriate person, time, or place (Schacter, 1999, 2001; Johnson, Hashtroudi, & Lindsay, 1993; Jacoby et al., 1989). One

<sup>1</sup>Harvard University, <sup>2</sup>Dartmouth College, <sup>3</sup>University of Virginia

form of misattribution error occurs when participants alter their subjective judgments of a stimulus as a result of its initial presentation. Among the best studied examples of this form of misattribution is the “illusory truth effect,” whereby perceivers tend to increase their judgments of the truth of information that had been presented earlier (Rahhal, May, & Hasher, 2002; Begg, Robertson, Gruppuso, Anas, & Needham, 1996; Begg, Anas, & Farinacci, 1992; Gilbert, Krull, & Malone, 1990; Bacon, 1979; Hasher, Goldstein, & Toppino, 1977). For example, Begg, Anas, et al. (1992) presented participants with a series of statements that were identified as being true, as being false, or without reference to the truth value of the statement (because the statements referred to obscure facts about the world, participants were unlikely to know the actual truth value of each statement). Subsequently, participants indicated their subjective sense of the veracity of both these previously encountered as well as novel statements. The tendency to judge old statements to be true was demonstrated by two behavioral effects. First, participants more frequently judged the correct truth value of cued-true items than that of cued-false items. That is, whereas cued-true items were typically judged correctly as true, cued-false items were more often judged incorrectly to be true than to be false. Second, participants judged cued-neutral statements (i.e., those for which no truth information was provided during study) to be more true than novel statements; that is, mere exposure to a statement increased its truth value. These two behavioral results together mark the illusory truth effect.

Cognitive explanations for this first manifestation of the illusory truth effect (i.e., the tendency to misjudge false items as true) have underscored the importance of *recollection* to the successful avoidance of misattribution errors (Schacter, 1999; Johnson et al., 1993; Begg, Anas, et al., 1992; Jacoby, 1991; Jacoby et al., 1989). Specifically, these accounts have suggested that misattribution errors are particularly likely under conditions in which perceivers have little or no contextual recollection for the details surrounding their initial encounter with a stimulus. For example, the illusory truth effect is thought to occur when perceivers recognize that information was encountered previously, but fail to recollect whether it was presented as true or false. As a result, perceivers may experience an item as being familiar and attempt to link this subjective sense of familiarity to some other feature of the item, such as its veracity. In other words, in the absence of full-blown recollection, perceivers are thought to attribute the familiarity of an item to the truth value of the information it conveys.

The notion that recollection underlies the ability to avoid judgment-based misattribution errors such as the illusory truth effect is supported by observations that illusory truth increases as source memory decreases. For example, source memory is attenuated both by the introduction of a secondary task during encoding or by

decreasing the discriminability of sources (Johnson et al., 1993; Lindsay, 1990), and these manipulations have also been shown to augment misattribution errors on illusory truth paradigms (Begg, Anas, et al., 1992). Likewise, analytic techniques designed to separate out the contributions of recollection to memory performance—the process dissociation procedure (Jacoby, 1991)—have suggested that successful suppression of the illusory truth effect correlates with the strength of recollection (Begg, Anas, et al., 1992).

If the illusory truth effect can indeed be counteracted by recollection, both successful encoding and judgment of cued-false items should be associated with brain areas previously identified in neuroimaging and patient studies of contextual recollection. Recent subsequent memory studies using fMRI (Ranganath, Yonelinas, et al., 2004; Davachi et al., 2003; Henson, Rugg, et al., 1999) have converged on a consistent set of brain regions in which encoding activation correlates specifically with subsequent recollection-based memory. This “neural signature” of successful recollective encoding consists of both the ventrolateral PFC (inferior frontal gyrus [IFG] and operculum) as well as the hippocampus. For example, Davachi et al. (2003) oriented participants to a series of words in two different ways (either form an image related to the word or read the word backwards). Subsequently, participants performed a two-step recognition task in which they were first asked to decide whether each item was old or new; if the word was judged to be old, the participant was further asked to specify whether he or she had performed the “imagine” or “read backwards” task. This latter source memory discrimination is thought to require recollection of the study episode, whereas simple old/new discriminations can be performed on the basis of other information, such as the item’s relative familiarity. Subsequent memory analyses indicated that encoding activation in both the left ventrolateral PFC and the hippocampus correlated with later recollection of which orienting task had been performed, but did not differentiate between old/new recognition. That is, activation in the ventrolateral PFC and the hippocampus during encoding correlated specifically with subsequent recollection, and not simply overall memory for the items. Encoding activation in highly similar brain regions has been observed by other researchers to correlate with recollection, as indexed by correct source memory judgments (Ranganath, Yonelinas, et al., 2004), as well as with “remember” responses on the remember/know paradigm (Henson, Rugg, et al., 1999).

There is currently less consensus regarding the brain regions that accompany recollection during *retrieval*. However, across a number of studies, tasks that require access to recollective details have been associated with greater activation (relative to retrieval tasks that can be performed without recollection) in a distributed set of brain regions that includes the ventrolateral PFC, the

superior frontal gyrus (SFG), the hippocampus, and left-lateralized posterior parietal regions (Dobbins, Rice, Wagner, & Schacter, 2003; Dobbins, Foley, Schacter, & Wagner, 2002; Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000; Ranganath, Johnson, & D'Esposito, 2000; Henson, Rugg, et al., 1999; Henson, Shallice, & Dolan, 1999; Rugg, Fletcher, Chua, & Dolan, 1999; Nolde, Johnson, & D'Esposito, 1998; Wagner, Desmond, Glover, & Gabrieli, 1998; Wagner, Poldrack, et al., 1998; Petrides, Alivisatos, & Evans, 1995).

This earlier research supports specific predictions about the brain regions that should be important for successful avoidance of misattribution errors on the illusory truth paradigm. In the current study, participants were scanned during encoding and subsequent judgment of trivia statements that were cued as either true or false or were associated with no truth information. During the judgment phase of the experiment, participants considered each of the statements and indicated whether each should be considered true or false (or that there was no basis to decide). The event-related nature of the experimental design allowed us to isolate neural activity during both encoding and judgment as a function of cue type (false, true) as well as subsequent judgment success. The theoretical perspective we have sketched suggests that to the extent that (a) misattribution is attenuated by the presence of recollection for the study episode and (b) later access to recollection correlates with encoding activation in the ventrolateral PFC and the hippocampus, then (c) encoding activation in these two brain regions should correlate with whether an item is subsequently subject to the illusory truth effect (i.e., inappropriately judged to be true). Because misattribution in illusory truth paradigms is indexed by the tendency to misjudge cued-false information to be true, we specifically expected encoding activation in the ventrolateral PFC and the hippocampus would distinguish between cued-false items that were later correctly judged to be false and those that were subsequently misjudged to be true. In contrast, because cued-true items can be correctly judged to be true even in the absence of recollection (i.e., on the basis of other stimulus features, such as item familiarity), no such difference should be observed between correctly judged and misjudged cued-true items.

Relatively little is known about the cognitive processes that give rise to the second form of illusory truth, the tendency to judge old items to be true in the absence of any truth cues during encoding. Because participants more frequently judge old items to be true (relative to novel items) even when item recognition is near ceiling (Bacon, 1979), it seems unlikely that this effect results from the unavailability of recollection, and accordingly, the scope of our predictions is restricted to information cued as true or false. Nevertheless, to provide behavioral confirmation that participants did, in fact, demonstrate both forms of illusory truth, the current study included

items that were uncued during encoding (i.e., cued-neutral items), even though we did not have specific predictions about brain regions involved in processing such stimuli.<sup>1</sup>

Finally, we made a set of secondary predictions regarding the activations at retrieval. Because of the bias to call familiar items true, cued-false items were expected to involve greater activation of source information than cued-true items, and should therefore engage brain regions previously implicated in recollective retrieval. Therefore, in line with earlier research, we expected that the ventrolateral PFC, the SFG, the hippocampus, and posterior parietal regions would be more activated for items that required access to recollection (i.e., cued-false items) than for items that could be judged without recollection (cued-true items).

## RESULTS

### Behavioral Data

The illusory truth effect is typically defined by two patterns of behavioral results, both of which were observed in the current study (Table 1). First, the correct truth value was assigned to a greater proportion of cued-true items (0.80) than cued-false items (0.58),  $t(11) = 7.97, p = 10^{-6}$ . Second, despite the absence of veracity information, a greater proportion of cued-neutral statements ( $M = 0.53$ ) were judged to be true than novel statements ( $M = 0.29$ ),  $t(11) = 10.13, p = 10^{-7}$ . These behavioral results together confirm that the current design succeeded in producing the illusory truth effect.

### fMRI Data: Encoding Phase

fMRI encoding data were conditionalized on the basis of subsequent truth judgments. Specifically, each cued-false and cued-true item was considered either correct or incorrect as a function of whether it was later judged false or true, resulting in 4 types of encoding trials: *cued-false/judged-false* (FF), *cued-false/judged-incorrect* (Fx), *cued-true/judged-true* (TT), and *cued-true/judged-incorrect* (Tx). A statement was considered correct if it was judged to be either “definitely” or

**Table 1.** Mean Proportion of Response Judgments by Encoding Cue

Encoding Cue	Judgment Response		
	“False”	“Neutral”	“True”
False	0.58	0.16	0.26
Neutral	0.19	0.28	0.53
True	0.06	0.14	0.80
Novel	0.17	0.54	0.29

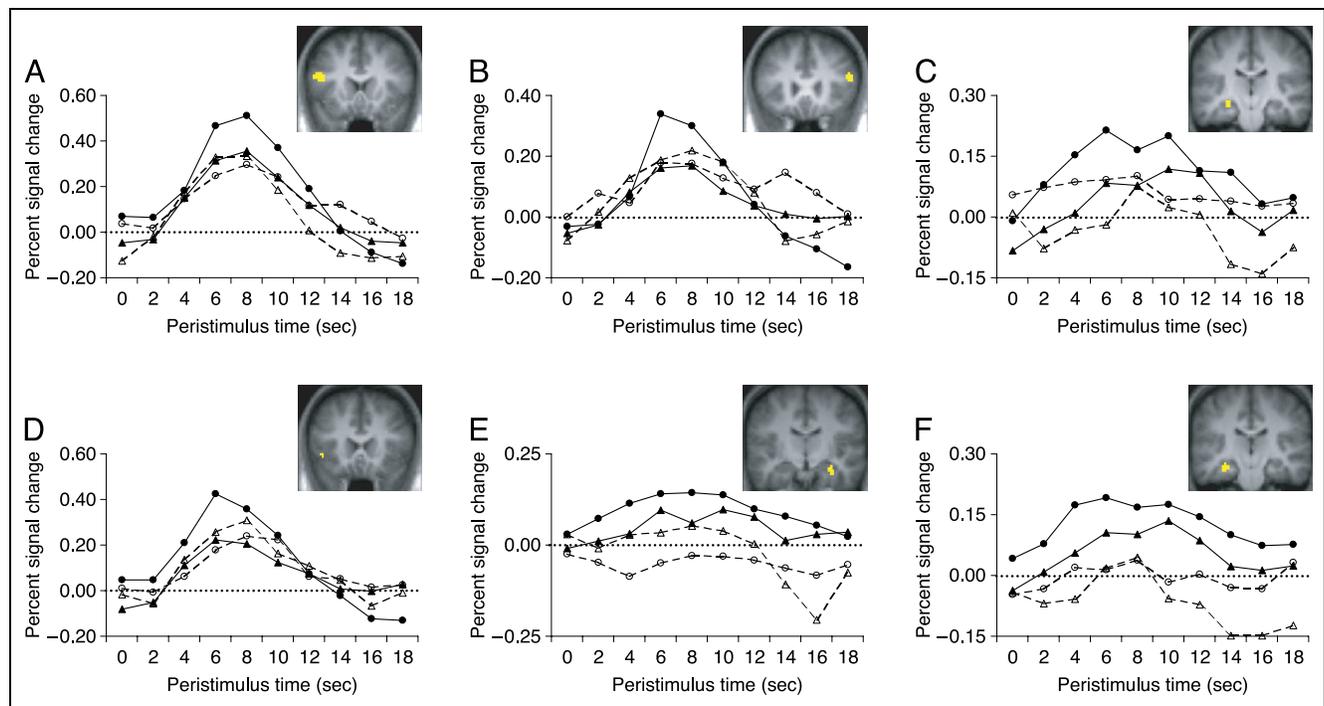
“probably” the same truth value as cued during encoding (e.g., Responses 1 or 2 for a cued-false item), and incorrect if it received any other response.

We adopted several complementary analytic strategies to examine differences in neural activation among conditions. First, we identified brain regions that were significantly activated by the encoding task relative to baseline, regardless of trial type (i.e., all encoding trials > fixation) and subsequently interrogated these regions of interest (ROIs) for differences between conditions. This analysis provides a conservative method for examining differences between conditions, as ROIs were defined in a manner that was unbiased with regard to differences among conditions. Brain areas that were significantly activated by the encoding task included a number of loci in the left ventrolateral PFC along the IFG, a more focused region in the right IFG, left-lateralized superior frontal gyrus, occipito-temporal cortex, MTL (including the left hippocampus), and the bilateral occipital cortex. ROI analyses demonstrated that encoding activation associated with FF trials was significantly greater than other trial types (i.e., FF > Fx ~ TT ~ Tx) in three regions: left hippocampus and bilateral IFG (Figure 1A–C). That is, encoding activation in the left hippocampus and the bilateral IFG differentiated subsequent correct judgments of cued-false statements (i.e., FF items) from other item types and did not differentiate among correct

judgments of cued-true (TT) or incorrect judgments of either kind (Fx or Tx items). No other difference between conditions was observed in this analysis in any other brain region.

However, by restricting analysis to regions that were activated by all or most conditions, this initial analysis strategy was fairly conservative. For example, regions in which some trial types were associated with “deactivations” below baseline would not be identified. To protect against Type II errors, we also compared conditions in a series of direct contrasts. A direct contrast of FF > TT yielded a number of regions in which encoding activation was differentially correlated with subsequent correct judgments of cued-false statements compared with cued-true statements: bilateral IFG, left operculum, left MTL, left fusiform gyrus, and left occipital cortex (Table 2). The reverse contrast, TT > FF, yielded no brain regions, even at a relaxed statistical threshold of  $p < .01$ , uncorrected. The contrast of FF > Fx yielded activations in the right IFG, right fusiform gyrus, left parahippocampal cortex, and right anterior hippocampus/amygdala. No regions were observed in the contrast of TT > Tx.

Finally, to confirm that predicted task differences were obtained in the exact regions previously associated with recollective encoding, we interrogated the same ROIs reported in an earlier study of successful source



**Figure 1.** ROIs identified in the encoding phase of the experiment. Activations are rendered onto coronal slices of participants’ mean normalized anatomical image. For each region, the hemodynamic time course of the BOLD response is displayed for encoding trials conditionalized by subsequent truth judgments. Three regions that demonstrated differences between conditions are presented: (A) left IFG; (B) right IFG; (C) left hippocampus. In addition, ROIs from an earlier study of recollective encoding (Davachi et al., 2003) were interrogated: (D) left operculum; (E) right hippocampus; (F) left hippocampus. In all six regions, greater activation was associated with successfully encoded cued-false items (FF, solid closed circles) than with incorrectly encoded cued-false items (Fx, dashed open circles), correctly encoded cued-true items (TT, solid closed triangles), or incorrectly encoded true items (Tx, dashed open triangles), i.e., a pattern of FF > Fx ~ TT ~ Tx.

**Table 2.** Peak Voxel and Number of Voxels for ROIs Demonstrating Time Course Differences between Conditions during Encoding

<i>Region</i>		<i>x</i>	<i>y</i>	<i>z</i>	<i>Voxels</i>
<i>FF &gt; Fx ~ TT ~ Tx</i>					
Inf. frontal gyrus	L	-54	27	-3	24
	L	-51	24	18	46
	L	-45	12	27	54
	L	-57	15	24	33
	R	57	27	18	26
Med. temporal lobe	L	-30	-15	-15	20
	L	-21	-33	0	19
	L	-27	-27	-21	39
<i>FF &gt; TT</i>					
Inf. frontal gyrus	L	-51	18	0	16
Med. temporal lobe	L	-30	-9	-15	13
<i>FF &gt; Fx</i>					
Inf. frontal gyrus	R	54	24	21	11
Parahipp. cortex	L	-24	-39	-21	30
Fusiform gyrus	R	42	-51	-21	21
Amygdala	R	21	-3	-21	13

Note: The table lists brain regions in which a significant difference was detected at encoding between any two conditions (cued-false/judged-false [FF], cued-false/judged-incorrect [Fx], cued-true/judged-true [TT], and cued-true/judged-incorrect [Tx] trials). For  $FF > Fx \sim TT \sim Tx$ , ROIs were defined from the contrast of *all encoding trials > baseline* and subsequently interrogated for differences between conditions.  $FF > TT$  and  $FF > Fx$  report the results of direct contrasts between those trial types. Coordinates are in MNI space. Inf. = inferior; med. = medial; parahipp. = parahippocampal.

memory encoding (Davachi et al., 2003). As reviewed above, Davachi et al. demonstrated that encoding activation in a number of regions—left ventrolateral PFC (both IFG and operculum regions), bilateral hippocampus, and left parahippocampal cortex—correlated specifically with recollection, rather than other components of recognition memory. Because these earlier results provide an independent “neural signature” of recollective encoding, we examined the neural response in each of these regions (i.e., precisely the same voxels) for the current data. In each, encoding activation was significantly greater for FF than for Fx, TT, or Tx trials (i.e.,  $FF > Fx \sim TT \sim Tx$ ). Figure 1 (panels D–F) displays hemodynamic time courses for the left operculum and bilateral hippocampus regions reported by Davachi et al.

### fMRI Data: Judgment Phase

As for the encoding phase, several complementary analyses were conducted on fMRI data from the judg-

ment phase of the experiment. We first identified brain regions that were significantly activated by the judgment task relative to baseline, regardless of trial type (i.e., all judgment trials > fixation). These regions included all those activated during encoding plus the bilateral parietal cortex, the anterior cingulate cortex (ACC), and the left occipito-temporal junction. We observed two patterns of differences among trial types. First, in the bilateral IFG, bilateral SFG, and ACC, an equivalent amount of activation was associated with FF and Fx items (all  $p$  values > .25), which were both greater than TT and Tx items (Table 3). In addition, Fx items showed significantly greater activation than FF, TT, and Tx items at two additional SFG loci. Interestingly, FF activation was not significantly greater than Fx activation in any region. This pattern of results was confirmed through a direct contrast of  $FF > Fx$ , which yielded no brain regions where activation was greater for FF than Fx items, even at a relaxed statistical threshold ( $p < .01$ , uncorrected).

Although the speed with which participants judged FF (2259 msec) and TT (2250 msec) items was very similar, Fx (2717 msec) trials were judged significantly more slowly than either of these other types of trials (both  $p$  values < .05). To examine the possibility that Fx activation might reflect the relatively long reaction times associated with these items, a subset of FF and Fx items was matched for reaction time within each participant ( $M$ s of 2168 and 2096 msec, respectively). Four participants had too few trials of at least one type (< 10) to permit meaningful analyses with this RT-matched set. For the remaining participants, no difference was

**Table 3.** Peak Voxel and Number of Voxels for ROIs Demonstrating Differences between Conditions during Judgment

<i>Region</i>		<i>x</i>	<i>y</i>	<i>z</i>	<i>Voxels</i>
<i>FF ~ Fx &gt; TT ~ Tx</i>					
Inf. frontal gyrus	L	-51	24	27	48
	L	-57	21	18	38
	R	51	24	18	59
	R	48	15	21	35
Sup. frontal gyrus	R	3	24	51	39
Ant. cingulate cortex	L	-9	21	39	39
<i>Fx &gt; FF ~ TT ~ Tx</i>					
Sup. frontal gyrus	L	-6	9	54	72
	L	-6	6	66	48

Note: The table lists ROIs in which a significant difference was detected between any two conditions during the judgment phase. All ROIs were identified from the contrast of *all judgment trials > baseline* and subsequently interrogated for differences between conditions. Inf. = inferior; sup. = superior; ant. = anterior.

observed between FF and Fx items in all frontal regions (SFG, bilateral IFG, ACC), including those regions originally demonstrating a pattern of Fx > FF. However, as displayed in Figure 2, both left posterior parietal loci (-36, -60, 45; -33, -72, 36) demonstrated a pattern consistent with successful retrieval of recollective source information, that is, significantly greater retrieval activation for RT-matched FF items than Fx items (i.e., FF > Fx ~ TT). Although Tx items were also associated with significantly longer RTs (2871 msec) than other item types, the small overall number of Tx items prevented a comparable analysis for cued-true statements.

## DISCUSSION

### Activations at Encoding

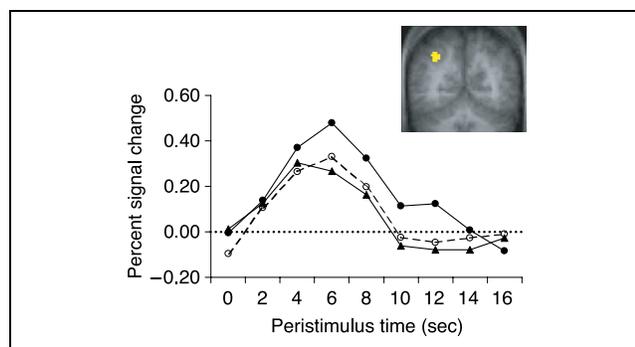
Cognitive research has suggested that misattribution errors can be avoided under conditions that support recollection of the contextual details surrounding one's initial encounter with a stimulus. In the current study, we indexed misattribution errors through the illusory truth effect, explanations for which have suggested that correct judgment of cued-false information (i.e., avoidance of misattribution) requires such access to recollective details. Accordingly, we predicted that the successful encoding of cued-false statements (i.e., FF items) would be associated with a pattern of neural activity previously linked to recollection-based memory encoding. Primary results demonstrated that FF items were associated with greater activation (relative to Fx, TT, or Tx items) in three brain regions: left and right ventrolateral PFC (corresponding to IFG) as well as left-lateralized hippocampus.

Together, these regions comprise the set of loci observed by previous event-related fMRI studies that have identified brain regions in which encoding activa-

tion correlates with successful memory (Davachi et al., 2003; Davachi & Wagner, 2002; Paller & Wagner, 2002; Strange et al., 2002; Otten, Henson, & Rugg, 2001; Kirchoff et al., 2000; Henson, Rugg, et al., 1999; Brewer et al., 1998; Wagner, Schacter, et al., 1998). Consistent with the notion that these regions contribute to encoding by specifically subserving recollection, almost all of these studies have restricted imaging analyses to items later associated with contextual recollection, whether indexed by free recall (Strange et al., 2002); items that are recognized with high confidence (Otten, Henson, & Rugg, 2001; Kirchoff et al., 2000; Wagner, Schacter, et al., 1998); "remember" responses in the remember/know procedure (Henson, Rugg, et al., 1999; Brewer et al., 1998); or by accurate source memory (Ranganath, Yonelinas, et al., 2004; Davachi et al., 2003). Indeed, recent work that has attempted to dissociate brain regions that contribute specifically to recollective encoding from those that play a role in other forms of memory (such as familiarity-based encoding) has suggested the selectivity of these regions for encoding that subserves subsequent recollection (Ranganath, Yonelinas, et al., 2004; Davachi et al., 2003). In the current work, secondary analyses confirmed that FF items were associated with greater encoding activation (relative to other item types) in the exact same regions linked by some of these earlier researchers to successful recollective encoding.

The contributions of the left ventrolateral PFC to memory encoding are thought to derive primarily from its role in semantic processing. On the basis of extant neuroimaging work, a number of theorists have suggested that the processes subserved by the left PFC might collectively be thought of as "semantic working memory," consisting of the selection, manipulation, and integration of an item's semantic meaning (Wagner, Pare-Blagoev, Clark, & Poldrack, 2001; Buckner & Koutstaal, 1998; Gabrieli, Poldrack, & Desmond, 1998; Wagner, Schacter, et al., 1998), and that this region may support elaborative memory encoding by contributing to deep semantic analysis and selection of appropriate verbal information. The contribution of the left ventrolateral PFC to this kind of semantic analysis may account both for the frequent observations that the response in this region is greater during deep, semantic encoding tasks than more superficial ones (Cabeza & Nyberg, 2000) and that encoding activation in this region typically correlates with subsequent memory success for verbal stimuli (Paller & Wagner, 2002).

In addition to the left ventrolateral PFC regions typically observed to correlate with recollective encoding, we also obtained a region in the right ventrolateral PFC that was differentially engaged for items that were not subsequently subject to misattribution errors (FF items). Encoding activation in similar right PFC regions has been shown to correlate with subsequent memory for nonverbal materials, such as pictures (Kirchoff et al.,



**Figure 2.** In two regions of the left parietal cortex, greater activation was observed for FF trials (solid, closed circles) than for Fx trials (dashed open circles), demonstrating a retrieval success effect for RT-matched trials. Fx trials did not differ from TT trials (solid closed triangles). Data are displayed for the parietal region at -36, -60, 45. Because of the small number of Tx trials, these items could not be RT-matched to other trial types.

2000; Brewer et al., 1998), but the right PFC is not generally observed during the encoding of verbal materials of the sort used in the current study. However, statements used in the current study were substantially more complex stimuli than the single words that have served as stimuli in many previous subsequent memory studies, and the recruitment of PFC regions in both hemispheres has been observed during sentence comprehension (e.g., Constable et al., 2004; Just, Newman, Keller, McEleney, & Carpenter, 2004). Accordingly, given recent observations that the processes engaged during encoding critically determine which brain regions will correlate with subsequent memory (Mitchell, Macrae, & Banaji, 2004; Otten, Henson, & Rugg, 2002), right PFC regions might be expected to contribute to successful encoding for complex verbal materials, such as sentences. An alternative possibility is that the statements used in the current study conveyed information about a wide range of topics, some of which may have prompted pictorial encoding processes supported by the right hemisphere (Kirchhoff et al., 2000; Brewer et al., 1998; Kelley et al., 1998).

A critical role for the PFC in recollective memory has also been suggested by the study of neuropsychological patients. For example, patients with damage to the frontal cortex have shown relative sparing of item recognition, but impaired use of source information that is critical for recollective memory (Shimamura, Janowsky, & Squire, 1990; Janowsky, Shimamura, & Squire, 1989). These impairments are generally thought to reflect the disruption of the strategic aspects of memory encoding and retrieval, such as semantic elaboration or initiation of appropriate retrieval searches (Shimamura et al., 1990; Janowsky et al., 1989), and can result in heightened susceptibility to memory distortion (Schacter, Curran, Galluccio, Milberg, & Bates, 1996).

In addition to the ventrolateral PFC, neuropsychological, neuroimaging, and nonhuman animal research has also suggested an important role for the hippocampus in recollection. Although considerable controversy exists regarding the precise contributions of the hippocampus to recognition memory (Manns & Squire, 1999; Reed, Hamann, Stefanacci, & Squire, 1997; Reed & Squire, 1997), patients with damage limited to the hippocampus show particularly pronounced impairment when successful performance requires recollection; the magnitude of their deficits is less clear when performance requires only familiarity (Holdstock et al., 2002; Mayes, Holdstock, Isaac, Hunkin, & Roberts, 2002; Yonelinas et al., 2002; Baddeley, Vargha-Khadem, & Mishkin, 2001). Recent neuroimaging work (Jackson & Schacter, 2004; Sperling, Chua, et al., 2003; Sperling, Bates, et al., 2001; Henke, Weber, Kneifel, Wieser, & Buck, 1999) has complemented these findings by demonstrating a role for the hippocampus in binding together contextual details into an associative recollection. Finally, although hippocampal contributions to recognition memory in

nonhuman primates is still a matter of some debate (Baxter & Murray, 2001a, 2001b; Zola & Squire, 2001), lesion studies in rodents have generally demonstrated the role of the hippocampus in contextual binding processes of the kind that support recollection (Agster, Fortin, & Eichenbaum, 2002; Fortin, Agster, & Eichenbaum, 2002; Aggleton & Brown, 1999).

Given the strong consensus from earlier research that the ventrolateral PFC and the hippocampus contribute to memory by subserving encoding operations that specifically support later recollection, results of the current study support earlier psychological theories regarding the cognitive processes required to avoid misattribution errors. The present results indicated that the “neural signature” of successful recollective encoding—the ventrolateral PFC and the hippocampus—also accompanied the encoding of items that were not later subject to misattribution. That is, the illusory truth effect was effectively minimized for items that were accompanied by relatively stronger activation in regions known to contribute to recollective encoding. In contrast, because recollection does not necessarily accompany Fx, TT, or Tx items (all of which could reflect the use of item familiarity at test), such statements were, on average, associated with less activation in the ventrolateral PFC and the hippocampus.

### Activations at Judgment

During the judgment phase, cued-false items were associated with a pattern of neural activations previously observed during recollection/source memory tasks. In particular, we observed that activations in the bilateral IFG, left SFG, and left posterior parietal regions were greater for both FF and Fx items than for TT or Tx items. These results overlap to a considerable extent with several recent studies that have examined brain regions in which activation was greater during attempts to recollect specific source information than during item recognition or other noncontextual retrieval tasks (Dobbins, Foley, et al., 2002; Eldridge et al., 2000; Ranganath, Johnson, et al., 2000; Henson, Rugg, et al., 1999; Henson, Shallice, et al., 1999; McDermott, Buckner, Petersen, Kelley, & Sanders, 1999; Rugg et al., 1999; Nolde, Johnson, & D’Esposito, 1998; Nolde, Johnson, & Raye, 1998; Wagner, Desmond, et al., 1998; Wagner, Poldrack, et al., 1998; Petrides et al., 1995). Most notably, studies by Dobbins and colleagues reported activations in the left-lateralized SFG and the posterior parietal cortex and along the extent of the left IFG when comparing source recognition to item recognition (Dobbins, Foley, et al., 2002) or to judgments of item recency, a task that does not require reinstatement of the study context (Dobbins, Rice, et al., 2003). A number of other studies have also observed ventrolateral PFC activity during retrieval (Ranganath, Johnson, et al., 2000; Rugg et al., 1999; Nolde, Johnson, & D’Esposito, 1998; Wagner, Desmond,

et al., 1998), and have suggested that, whereas the right PFC accompanies a wide range of retrieval tasks (including simple old/new recognition), left PFC activity is likely to be observed during more complex or demanding retrieval tasks (Ranganath, Johnson, et al., 2000; Nolde, Johnson, & D'Esposito, 1998; Nolde, Johnson, & Raye, 1998).

In addition, left-lateralized posterior parietal activations have also been reported at retrieval by several authors (Wheeler & Buckner, 2004; Henson, Rugg, et al., 1999) in comparisons of “remember” responses (which are presumed to reflect recollective memory) and “know” responses (which are presumed to reflect familiarity). Similar findings have been reported for a source memory task versus simple old/new recognition (Henson, Shallice, et al., 1999). Likewise, “remember” responses have also been associated with increased activity in the left SFG (Henson, Rugg, et al., 1999).

Interestingly, the full set of fMRI data did not yield evidence of a difference for FF > Fx. The lack of this difference in the current study is consistent with most event-related fMRI retrieval studies, which suggest that, although cortical activity may be associated with retrieval effort or mode, it does not typically track differences between hits and misses (Dobbins, Rice, et al., 2003; Dobbins, Foley, et al., 2002; Wagner, Desmond, et al., 1998). However, when reaction times for FF and Fx trials were equated, we effectively observed a retrieval success effect (FF > Fx) located in the posterior parietal cortex. Because most other retrieval studies have not controlled for reaction time differences between hits and misses, the generality of this parietal success effect is unknown. However, it is consistent with the formulation of Petrides and colleagues, who have suggested that, whereas the frontal lobes initiate and maintain an active retrieval mode, posterior cortical areas may actually subserve the stored representations or preexisting associations that comprise a successful recollection (Petrides et al., 1995). Future research will be needed to test this possibility directly.

## Conclusions

In the current study, we identified a pattern of neural activity associated with the successful encoding and judgment of the truth of information. Our results are consistent with earlier neuroimaging studies that have identified brain regions in which activity was associated either with (a) encoding of events in a manner that supports later recovery of recollective details of the study episode or (b) retrieval of such contextual information. Against the backdrop of extant research, the current study provides initial neuroimaging evidence that the successful avoidance of misattribution—here measured by suppression of the illusory truth effect—relies on the use of contextual recollection that may oppose other memory processes (e.g., item familiarity).

These kinds of misattribution errors have clear implications for day-to-day cognition. The difficulty posed by attempts to ignore or disregard information is well known outside the laboratory setting; indeed, trial lawyers and judges often go to great lengths to instruct juries to discount evidence or testimony inappropriately admitted into a case. Notwithstanding, a sizeable body of psychological research demonstrates that jury decisions can nevertheless be heavily influenced by inadmissible evidence even after repeated instructions to disregard such information (Kassin & Studebaker, 1998). Similarly, perceivers have been shown to persist in using experimentally provided information to make causal attributions or judgments about their own abilities (Anderson, Lepper, & Ross, 1980; Ross, Lepper, & Hubbard, 1975), even after debriefing procedures expose the fictitious nature of the information. In the current work, we investigate the neural basis of this class of memory failure and, in doing so, provide converging evidence for earlier, cognitive accounts of misattribution.

## METHODS

### Participants

Participants were 12 (6 women; mean age: 21 years, range: 18–27) right-handed, native English speakers with no history of neurological problems. Informed consent was obtained in a manner approved by the Human Studies Committee of the Massachusetts General Hospital.

### Stimuli and Behavioral Procedure

Stimuli consisted of 296 “trivia” statements (Bacon, 1979) with which participants were unlikely to be familiar, for instance, “It takes 4 hours to hard boil an ostrich egg.” Half of the statements were factually true; the remaining half were rendered false by changing one piece of information, for example, “It takes 6 hours to hard boil an ostrich egg.” On each encoding trial, a statement was presented for 5500 msec below one of the cue words, *True*, *False*, or *Neutral*. In each of two functional runs, an equal number of the statements were cued as true, false, or neutral (111 total statements per run). Across participants, statements were randomly cued as true, false, or neutral, regardless of their actual truth value. Participants were instructed to memorize each statement and whether it was true or false (for cued-neutral statements, participants were instructed to encode the statement only). To ensure they were attending to the task, participants were required to respond to each statement by pressing a different button for each of the three cue words.

During the truth judgment phase of the experiment, participants viewed 222 old statements presented randomly among 74 novel statements. On each trial, par-

Participants saw a statement presented below a 5-point scale that read: 1 = definitely false; 2 = probably false; 3 = neither true nor false; 4 = probably true; 5 = definitely true. Participants judged the statements for how true or false they believed each to be (or pressed 3 if they did not have a basis for judging its veracity). Instructions underscored that participants should use their feelings about how true or false each statement seemed but did not explicitly warn against intentionally revisiting the study phase. The scale and statement remained onscreen together for 5500 msec. Trials on which participants failed to respond within this window were excluded from all statistical analyses (< 5% of trials). During both phases, trials were separated by a variable interstimulus interval (500–5500 msec) that was used to optimize estimation of the event-related fMRI response (Dale, 1999). During interstimulus intervals, participants passively viewed a fixation crosshair.

### Imaging Procedure

Imaging was conducted using a 3-T Siemens Allegra scanner. Functional scanning used a gradient-echo echo-planar pulse sequence (TR, 2 sec; TE, 40 msec;  $3.75 \times 3.75$  in-plane resolution). The encoding (433 acquisitions per run) and judgment (544 acquisitions) phases of the experiment each consisted of two functional runs (21 slices aligned to AC–PC; 5 mm thick; 1 mm skip). Stimuli were projected onto a screen at the end of the magnet bore that participants viewed by way of a mirror mounted on the head coil. A pillow and foam cushions were placed inside the head coil to minimize head movements.

fMRI data were preprocessed and analyzed using SPM99 (Wellcome Department of Cognitive Neurology, London, UK). First, functional data were time-corrected for differences in acquisition time between slices for each whole-brain volume and realigned to correct for head movement. Functional data were then transformed into a standard anatomical space (3-mm isotropic voxels) on the basis of the ICBM 152 brain template (Montreal Neurological Institute), which approximates Talairach and Tournoux atlas space. Normalized data were then spatially smoothed (8 mm full width at half maximum) using a Gaussian kernel.

Statistical analyses were performed using the general linear model in which the event-related design was modeled using a canonical hemodynamic response function and its temporal derivative (temporal derivatives were included to account for variability associated with differences in the onset of the hemodynamic response, but were not analyzed further). Comparisons of interest were implemented using a random-effects model. A voxel-based statistical threshold of  $p < .001$  was used for all comparisons; ROIs were required to exceed 23 contiguous voxels in extent (providing an alpha level of  $p < .05$ , corrected) for contrasts of all

experimental task trials over fixation baseline, (e.g., all encoding trials > fixation) and 10 contiguous voxels (i.e.,  $p < .001$ , uncorrected) for direct comparisons between conditions (e.g., FF > TT). Peristimulus hemodynamic time courses for each of these ROIs were extracted on a subject-by-subject basis using a selective averaging procedure (R. A. Poldrack, University of California Los Angeles). Statistical comparisons between conditions were conducted using analysis of variance (ANOVA) procedures on the parameter estimates associated with each trial type ( $p < .05$ ).

### Acknowledgments

We thank L. Davachi, I. Dobbins, D. Gilbert, O. Jackson, M. Johnson, A. Maril, S. Prince, S. Slotnick, and A. Wagner for advice and assistance. This research was supported by a grant from the National Institute of Health (R01 MH60941-01). J.P.M. was funded by a National Research Service Award predoctoral fellowship.

Reprint requests should be sent to Jason Mitchell, PhD, Department of Psychology, Harvard University, William James Hall, 33 Kirkland Street, Cambridge, MA 02138, or via e-mail: jmitchel@wjh.harvard.edu.

The data reported in this experiment have been deposited with the fMRI Data Center archive ([www.fmridc.org](http://www.fmridc.org)). The accession number is 2-2004-117XF.

### Note

1. Moreover, the current design did not permit an unambiguous determination of whether a cued-neutral item was subsequently judged correctly. Participants could indicate that a cued-neutral item was “neither true nor false” either because they successfully recollected the original cue or, conversely, because of the complete absence of memory (i.e., believing that the item was being seen for the first time). As such, cued-neutral items that were judged “neither true nor false” comprised a heterogeneous set that precluded appropriate neuroimaging comparisons. We reiterate that, because earlier cognitive research has suggested that recollection should not play a critical role in misattributions for cued-neutral items, fMRI analysis of these items was not relevant to the predictions outlined in the current study. The inclusion of cued-neutral statements was only critical for ensuring correspondence with previous work by confirming that we obtained both behavioral markers of the illusory truth effect.

### REFERENCES

- Aggleton, J. P., & Brown, M. W. (1999). Episodic memory, amnesia, and the hippocampal–anterior thalamic axis. *Behavioral and Brain Sciences*, *22*, 425–444; discussion 444–489.
- Agster, K. L., Fortin, N. J., & Eichenbaum, H. (2002). The hippocampus and disambiguation of overlapping sequences. *Journal of Neuroscience*, *22*, 5760–5768.
- Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance in social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology*, *39*, 1037–1049.
- Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., Glover, G. H., & Gabrieli, J. D.

- (2004). Neural systems underlying the suppression of unwanted memories. *Science*, *303*, 232–235.
- Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Learning and Memory*, *5*, 241–252.
- Baddeley, A., Vargha-Khadem, F., & Mishkin, M. (2001). Preserved recognition in a case of developmental amnesia: Implications for the acquisition of semantic memory? *Journal of Cognitive Neuroscience*, *13*, 357–369.
- Baxter, M. G., & Murray, E. A. (2001a). Effects of hippocampal lesions on delayed nonmatching-to-sample in monkeys: A reply to Zola and Squire (2001). *Hippocampus*, *11*, 201–203.
- Baxter, M. G., & Murray, E. A. (2001b). Opposite relationship of hippocampal and rhinal cortex damage to delayed nonmatching-to-sample deficits in monkeys. *Hippocampus*, *11*, 61–71.
- Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, *121*, 446–458.
- Begg, I. M., Robertson, R. K., Gruppuso, V., Anas, A., & Needham, D. R. (1996). The illusory-knowledge effect. *Journal of Memory and Language*, *35*, 410–433.
- Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1998). Making memories: Brain activity that predicts how well visual experience will be remembered. *Science*, *281*, 1185–1187.
- Buckner, R. L., & Koutstaal, W. (1998). Functional neuroimaging studies of encoding, priming, and explicit memory retrieval. *Proceedings of the National Academy of Sciences, U.S.A.*, *95*, 891–898.
- Cabeza, R., & Nyberg, L. (2000). Imaging cognition: II. An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*, 1–47.
- Constable, R. T., Pugh, K. R., Berroya, E., Mencl, W. E., Westerveld, M., Ni, W., & Shankweiler, D. (2004). Sentence complexity and input modality effects in sentence comprehension: An fMRI study. *Neuroimage*, *22*, 11–21.
- Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, *8*, 109–114.
- Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences, U.S.A.*, *100*, 2157–2162.
- Davachi, L., & Wagner, A. D. (2002). Hippocampal contributions to episodic encoding: Insights from relational and item-based learning. *Journal of Neuroscience*, *88*, 982–990.
- Dobbins, I. G., Foley, H., Schacter, D. L., & Wagner, A. D. (2002). Executive control during episodic retrieval: Multiple prefrontal processes subserve source memory. *Neuron*, *35*, 989–996.
- Dobbins, I. G., Rice, H. J., Wagner, A. D., & Schacter, D. L. (2003). Memory orientation and success: Separable neurocognitive components underlying episodic recognition. *Neuropsychologia*, *41*, 318–333.
- Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience*, *3*, 1149–1152.
- Fortin, N. J., Agster, K. L., & Eichenbaum, H. B. (2002). Critical role of the hippocampus in memory for sequences of events. *Nature Neuroscience*, *5*, 458–462.
- Gabrieli, J. D. E., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences, U.S.A.*, *95*, 906–913.
- Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology*, *59*, 601–613.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.
- Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Behavior*, *16*, 107–112.
- Henke, K., Weber, B., Kneifel, S., Wieser, H. G., & Buck, A. (1999). Human hippocampus associates information in memory. *Proceedings of the National Academy of Sciences, U.S.A.*, *96*, 5884–5889.
- 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.
- Henson, R. N. A., Shallice, T., & Dolan, R. J. (1999). Right prefrontal cortex and episodic memory retrieval: A functional MRI test of the monitoring hypothesis. *Brain*, *122*, 1367–1381.
- Holdstock, J. S., Mayes, A. R., Roberts, N., Cezayirli, E., Isaac, C. L., O'Reilly, R. C., & Norman, K. A. (2002). Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus*, *12*, 341–351.
- Jackson, O., & Schacter, D. L. (2004). Encoding activity in anterior medial temporal lobe supports subsequent associative recognition. *Neuroimage*, *21*, 456–462.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541.
- Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology*, *56*, 326–338.
- Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia*, *27*, 1043–1056.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.
- Just, M. A., Newman, S. D., Keller, T. A., McEleney, A., & Carpenter, P. A. (2004). Imagery in sentence comprehension: An fMRI study. *Neuroimage*, *21*, 112–124.
- Kassin, S. M., & Studebaker, C. A. (1998). Instructions to disregard and the jury: Curative and paradoxical effects. In J. M. Golding & M. D. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 413–434). Mahwah, NJ: Erlbaum.
- Kelley, W. M., Miezin, F. M., McDermott, K. B., Buckner, R. L., Raichle, M. E., Cohen, N. J., Ollinger, J. M., Akbudak, E., Conturo, T. E., Snyder, A. Z., & Petersen, S. E. (1998). Hemispheric specialization in human dorsal frontal cortex and medial temporal lobe for verbal and nonverbal memory encoding. *Neuron*, *20*, 927–936.
- Kirchhoff, B. A., Wagner, A. D., Maril, A., & Stern, C. E. (2000). Prefrontal-temporal circuitry for episodic encoding and subsequent memory. *Journal of Neuroscience*, *20*, 6173–6180.
- Lindsay, D. S. (1990). Misleading suggestions can impair eyewitnesses' ability to remember event details. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *16*, 1077–1083.
- Manns, J. R., & Squire, L. R. (1999). Impaired recognition memory on the doors and people test after damage limited to the hippocampal region. *Hippocampus*, *9*, 495–499.

- Maril, A., Simons, J. S., Mitchell, J. P., Schwartz, B. L., & Schacter, D. L. (2003). Feeling-of-knowing in episodic memory: An event-related fMRI study. *Neuroimage*, *18*, 827–836.
- Maril, A., Wagner, A. D., & Schacter, D. L. (2001). On the tip of the tongue: An event-related fMRI study of semantic retrieval failure and cognitive conflict. *Neuron*, *31*, 653–660.
- Mayes, A. R., Holdstock, J. S., Isaac, C. L., Hunkin, N. M., & Roberts, N. (2002). Relative sparing of item recognition memory in a patient with adult-onset damage limited to the hippocampus. *Hippocampus*, *12*, 325–340.
- McDermott, K. B., Buckner, R. L., Petersen, S. E., Kelley, W. M., & Sanders, A. L. (1999). Set- and code-specific activation in frontal cortex: An fMRI study of encoding and retrieval of faces and words. *Journal of Cognitive Neuroscience*, *11*, 631–640.
- Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2004). Encoding specific effects of social cognition on the neural correlates of subsequent memory. *Journal of Neuroscience*, *24*, 4912–4917.
- Nolde, S. F., Johnson, M. K., & D'Esposito, M. (1998). Left prefrontal activation during episodic remembering: An event-related fMRI study. *NeuroReport*, *9*, 3509–3514.
- Nolde, S. F., Johnson, M. K., & Raye, C. L. (1998). The role of prefrontal cortex during tests of episodic memory. *Trends in Cognitive Sciences*, *2*, 399–406.
- Otten, L. J., Henson, R. N. A., & Rugg, M. D. (2001). Depth of processing effects on neural correlates of memory encoding: Relationship between findings from across- and within-task comparisons. *Brain*, *124*, 399–412.
- Otten, L. J., Henson, R. N. A., & Rugg, M. D. (2002). State-related and item-related neural correlates of successful memory encoding. *Nature Neuroscience*, *5*, 1339–1344.
- Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Sciences*, *6*, 93–102.
- Petrides, M., Alivisatos, B., & Evans, A. C. (1995). Functional activation of the human ventrolateral frontal cortex during mnemonic retrieval of verbal information. *Proceedings of the National Academy of Sciences, U.S.A.*, *92*, 5803–5807.
- Rahhal, T. A., May, C. P., & Hasher, L. (2002). Truth and character: Sources that older adults can remember. *Psychological Science*, *13*, 101–105.
- 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*, RC108.
- Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia*, *42*, 2–13.
- Reed, J. M., Hamann, S. B., Stefanacci, L., & Squire, L. R. (1997). When amnesic patients perform well on recognition memory tests. *Behavioral Neuroscience*, *111*, 1163–1170.
- Reed, J. M., & Squire, L. R. (1997). Impaired recognition memory in patients with lesions limited to the hippocampal formation. *Behavioral Neuroscience*, *111*, 667–675.
- Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, *32*, 880–892.
- Rugg, M. D., Fletcher, P. C., Chua, P. M., & Dolan, R. J. (1999). The role of the prefrontal cortex in recognition memory and memory for source: An fMRI study. *Neuroimage*, *10*, 520–529.
- Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, *54*, 182–203.
- Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and remembers*. Boston: Houghton Mifflin.
- Schacter, D. L., Curran, T., Galluccio, L., Milberg, W. P., & Bates, J. F. (1996). False recognition and the right frontal lobe: A case study. *Neuropsychologia*, *34*, 793–808.
- Shimamura, A. P., Janowsky, J. S., & Squire, L. R. (1990). Memory for the temporal order of events in patients with frontal lobe lesions and amnesic patients. *Neuropsychologia*, *28*, 803–813.
- Sperling, R. A., Bates, J. F., Cocchiarella, A. J., Schacter, D. L., Rosen, B. R., & Albert, M. S. (2001). Encoding novel face–name associations: A functional MRI study. *Human Brain Mapping*, *14*, 129–139.
- Sperling, R. A., Chua, E., Cocchiarella, A., Rand-Giovannetti, E., Poldrack, R., Schacter, D. L., & Albert, M. (2003). Putting names to faces: Successful encoding of associative memories activates the anterior hippocampal formation. *Neuroimage*, *20*, 1400–1410.
- Strange, B. A., Otten, L. J., Josephs, O., Rugg, M. D., & Dolan, R. J. (2002). Dissociable human perirhinal, hippocampal, and parahippocampal roles during verbal encoding. *Journal of Neuroscience*, *22*, 523–528.
- Wagner, A. D., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1998). Prefrontal cortex and recognition memory. Functional-MRI evidence for context-dependent retrieval processes. *Brain*, *121*, 1985–2002.
- Wagner, A. D., Pare-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron*, *31*, 329–338.
- Wagner, A. D., Poldrack, R. A., Eldridge, L. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1998). Material-specific lateralization of prefrontal activation during episodic encoding and retrieval. *NeuroReport*, *9*, 3711–3717.
- Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., Rosen, B. R., & Buckner, R. L. (1998). Building memories: Remembering and forgetting of verbal experience as predicted by brain activity. *Science*, *281*, 1188–1191.
- Wheeler, M. E., & Buckner, R. L. (2004). Functional–anatomic correlates of remembering and knowing. *Neuroimage*, *21*, 1337–1349.
- Yonelinas, A. P., Kroll, N. E., Quamme, J. R., Lazzara, M. M., Sauve, M. J., Widaman, K. F., & Knight, R. T. (2002). Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nature Neuroscience*, *5*, 1236–1241.
- Zola, S. M., & Squire, L. R. (2001). Relationship between magnitude of damage to the hippocampus and impaired recognition memory in monkeys. *Hippocampus*, *11*, 92–98.

**This article has been cited by:**

1. Sam Harris, Sameer A. Sheth, Mark S. Cohen. 2008. Functional neuroimaging of belief, disbelief, and uncertainty. *Annals of Neurology* **63**:2, 141-147. [[CrossRef](#)]
2. Jason P. Mitchell, Alison L. Sullivan, Daniel L. Schacter, Andrew E. Budson. 2006. Misattribution errors in Alzheimer's disease: The illusory truth effect. *Neuropsychology* **20**:2, 185-192. [[CrossRef](#)]