The Cognitive Neuroscience of Memory Distortion

Review

Daniel L. Schacter* and Scott D. Slotnick
Department of Psychology
Harvard University
Cambridge, Massachusetts 02138

Memory distortion occurs in the laboratory and in everyday life. This article focuses on false recognition, a common type of memory distortion in which individuals incorrectly claim to have encountered a novel object or event. By considering evidence from neuropsychology, neuroimaging, and electrophysiology, we address three questions. (1) Are there patterns of neural activity that can distinguish between true and false recognition? (2) Which brain regions contribute to false recognition? (3) Which brain regions play a role in monitoring or reducing false recognition? Neuroimaging and electrophysiological studies suggest that sensory activity is greater for true recognition compared to false recognition. Neuropsychological and neuroimaging results indicate that the hippocampus and several cortical regions contribute to false recognition. Evidence from neuropsychology, neuroimaging, and electrophysiology implicates the prefrontal cortex in retrieval monitoring that can limit the rate of false recognition.

Introduction

The scientific analysis of memory would not be possible if memory were perfect. Experimental and theoretical analyses of memory typically depend on variations in the proportions of items or responses recalled, recognized, or otherwise produced across different conditions. Such variations, which provide the basis for drawing inferences about the processes that support memory function, necessarily involve some degree of forgetting. Beyond this ubiquitous sense in which imperfect retention is necessary to study memory, observations concerning the nature of forgetting have been conceptually central to the cognitive neuroscience of memory. For example, observations of pathological forgetting after damage to the medial temporal lobe (MTL) in human amnesic patients have greatly influenced our current understanding of the role of the MTL in various aspects of memory (Squire and Schacter, 2002).

Memory's imperfections, however, are not restricted to forgetting. Distortions and illusions of memory also occur—that is, there are instances in which some form of memory is present, but incorrect. Memory distortions can be considered in the context of the proposal that memory's imperfections can be divided into seven basic categories, or "sins." (Schacter, 1999, 2001). Three of the sins involve different types of forgetting (called *transience, absent-mindedness*, and *blocking*), and one refers to intrusive, unwanted memories of arousing or traumatic events (*persistence*). Most relevant to the topic of the present paper, three of the sins concern memory

distortions: misattribution, suggestibility, and bias. *Misattribution* occurs when retrieved information is assigned to the wrong source (e.g., mistaking a previously imagined event for a real one); *suggestibility* refers to the incorporation of inaccurate information from external sources, such as misleading questions, into one's own memories; and *bias* involves the distorting influences of present knowledge, beliefs, and feelings on recollection of previous experience.

Cognitive psychologists have long been interested in each of the distortion-related sins and have produced much research concerning their properties and implications (for reviews, see Johnson et al., 1993; Koriat and Goldsmith, 1996; Roediger and McDermott, 2000; Schacter, 2001; Schacter et al., 1998a). Such distortions have interested psychologists because they can provide insight into the constructive nature of memory, revealing how bits of information are patched together to form memories with varying degrees of accuracy. Cognitive neuroscientists have, until recently, shown little interest in examining memory distortions. During the past decade, however, there has been steadily increasing research on the topic, and a cognitive neuroscience of memory distortion is beginning to emerge. Much of this work has focused on the misattribution error called *false recognition*, which occurs when subjects incorrectly claim that a novel item has been encountered earlier in an experiment. False recognition is typically inferred from "old" responses to novel items that are conceptually or perceptually related to previously studied items—that is, false alarms above and beyond the "baseline" level of false alarms to unrelated novel items (cf. Roediger and McDermott, 1995; Underwood, 1965).

Cognitive neuroscience research during the past decade has used three different approaches to analyze false recognition and related forms of memory distortion: neuroimaging, electrophysiology, and neuropsychology. Several studies have used functional neuroimaging techniques—positron emission tomography (PET) and functional magnetic resonance imaging (fMRI)—in an attempt to identify specific brain regions associated with true and false recognition. Electrophysiological activity associated with true and false recognition has been examined using event-related potentials (ERPs), which measure changes in voltage topography on the scalp over time. Neuropsychological analyses have focused on brain-damaged patients with lesions to MTL and prefrontal regions that have been implicated previously in accurate remembering.

We consider evidence from each of these approaches in an attempt to address three main issues that have emerged in the literature during the past decade. First, we examine the evidence that brain activity can distinguish between true and false memories. From the subjective standpoint of an individual engaged in the act of remembering, true and false recognition are associated with an identical mnemonic outcome: in both cases, people believe they are remembering an event that occurred in their past. Therefore, it has been of considerable interest to determine whether, and in what sense,

*Correspondence: dls@wjh.harvard.edu

brain activity differs during true and false recognition. We focus in particular on the hypothesis, tested by several investigators, that true memory, more than false memory, is preferentially associated with the reactivation of sensory/perceptual processing that occurred during encoding. Second, we examine evidence that addresses the issue of what kind of brain activity is associated with, or responsible for, the generation of false memories. A number of studies have been particularly concerned with the role of MTL structures in the generation of false memories, but recent research has also begun to examine the role of other regions as well. Third, we will consider research that has examined what brain regions are involved in monitoring or reducing susceptibility to false memories. Much of this research has focused on the analysis of regions within prefrontal cortex.

Distinguishing between True and False Memories: The Sensory Reactivation Hypothesis

Although subjective responses are comparable during true and false recognition, cognitive studies have provided some evidence for differences in the nature of the representations that support the two forms of memory. In particular, several behavioral studies indicate that true memories are accompanied by retrieval of greater sensory/perceptual details than are false memories (e.g., Mather et al., 1997; Norman and Schacter, 1997; Schooler et al., 1986). Such findings make intuitive sense because true memories engage perceptual encoding processes that are presumably not involved in the creation of false memories. These findings have therefore led neuroscientists to ask whether neural activity accompanying true recognition, compared with false recognition, shows signs of sensory processing or reactivation.

To begin to address the question of whether brain activity can distinguish between true and false memories, experimental paradigms are needed that produce robust and reliable memory distortions in the laboratory. These considerations are especially important for neuroimaging studies, where it is necessary to obtain sufficient numbers of observations to yield a signal that can be statistically distinguished from noise. Roediger and McDermott (1995) described such a paradigm. They modified and extended a procedure developed initially by Deese (1959) in which subjects hear lists of associated words (e.g., *candy, sour, sugar, bitter, good, taste, tooth*, etc.) that all converge on a nonpresented "theme word" or false target (e.g., *sweet*). Roediger and McDermott reported extremely high levels of false recognition (e.g., 80%) to the theme words across a variety of word associate lists. Moreover, the level of false recognition responses was indistinguishable from the hit rate to studied items, and false recognition responses were accompanied by high confidence. Numerous subsequent studies using this Deese-Roediger/McDermott ("DRM") paradigm have delineated a variety of behavioral and cognitive properties of this powerful false recognition effect (e.g., Brainerd et al., 2003; Gallo et al., 2001; Mather et al., 1997; Neuschatz et al., 2003; Norman and Schacter, 1997; Robinson and Roediger, 1997; Schacter et al., 1999).

Neuroimaging of True and False Recognition

In the first neuroimaging study to compare true and false recognition, PET scans were carried out while subjects performed a recognition test that followed study of various DRM associate lists (Schacter et al., 1996a). After auditory presentation of lists containing 20 words that were associates of nonpresented theme words, subjects were scanned during separate test blocks in which they responded separately to true targets (words that had been studied previously), false targets (nonstudied semantic associates of previously studied items), or target controls (nonpresented words that were unrelated to previously studied words). Compared to a common baseline fixation condition, both true and false recognition were associated with blood flow increases in various regions that are commonly activated by episodic retrieval tasks, including dorsolateral/anterior prefrontal, medial parietal, and medial temporal regions. Direct comparisons indicated greater activation during true than false recognition in a left temporoparietal region previously associated with auditory processing and retention. Because subjects had heard true targets—but not false targets—during the auditory study phase of the experiment, this finding supports the possibility that true recognition is preferentially associated with an auditory sensory signature that reflects memory for auditory/ phonological aspects of previously studied words.

In a related study, Schacter et al. (1997a) performed a similar experiment using the then newly developed event-related fMRI, which allows stimuli from different conditions to be mixed together (as opposed to presented in separate test blocks according to condition, as in PET studies). Their results supported the general conclusion that many of the same or similar patterns of brain activity are observed during both true and false recognition (compared to a fixation baseline control). However, no regions were found that showed greater activation for true than false recognition, not even the left temporoparietal region observed by Schacter et al. (1996a).

These early neuroimaging studies are thus inconclusive regarding the question of whether brain regions involved in perceptual processing and representation are preferentially active during true versus false recognition. Cabeza et al. (2001) capitalized on prior research indicating that differences between true and false recognition can be increased when perceptual processing of target materials is increased during encoding (e.g., Schacter et al., 1999). Increased perceptual encoding, in turn, could be reflected in more discriminable patterns of brain activity during retrieval. To enhance perceptual encoding, Cabeza et al. instructed subjects (prior to scanning) to try to remember lists of semantically associated words and also to try to remember the source (a man or a woman) who presented the word lists; on a videotape, half the words were spoken by a man and half were spoken by a woman. They hypothesized that, during subsequent retrieval, previously studied words but not semantically associated false targets would activate regions initially involved in encoding perceptual (i.e., source) information, whereas regions involved in the encoding/retrieval of semantic information would show comparable activation during the two forms of recognition.
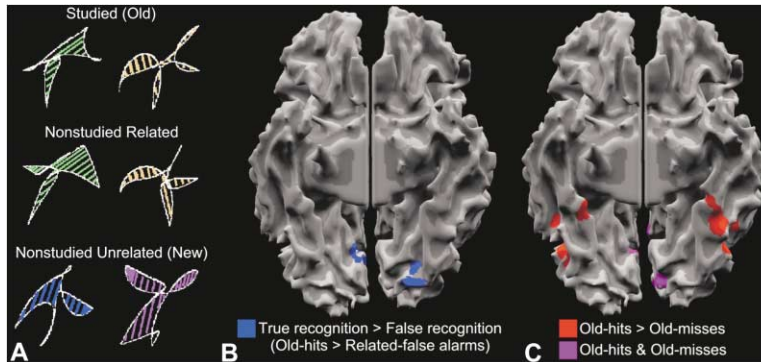
Figure 1. Stimuli and Results from Slotnick and Schacter, 2004

Using event-related fMRI, Cabeza et al. documented a dissociation between two regions within the MTL: parahippocampal gyrus, which showed greater activation during true than false recognition, and hippocampus, which showed significant activation during both true and false recognition (we discuss this finding further in the next section of the article). The true > false parahippocampal activation suggests greater recovery of sensory or contextual information during true than false recognition (see Bar and Aminoff, 2003).

Further relevant evidence comes from a study by Okado and Stark (2003) that used fMRI to compare true recognition of previously perceived events with false recognition of previously imagined events (cf. Johnson and Raye, 1981). Prior to scanning, subjects studied verbal labels of common objects followed either by a picture of the object or instructions to imagine the object. They were subsequently given a "lie test," intended to increase the number of source monitoring errors, in which they were asked to indicate whether or not they had previously seen an actual picture of the object, and were strongly encouraged to lie that they had seen a picture even when they had not. Subjects were then scanned during a more standard memory test in which they indicated whether they had actually seen a picture of an object during the study phase. Results revealed that a number of regions showed greater activity during true than false recognition, including bilateral occipital cortices and right parahippocampal gyrus, which the authors suggested reflects greater recovery of perceptual information for true than false memories. Okado and Stark's findings concerning parahippocampal gyrus resemble those reported previously by Cabeza et al. (2001).

To examine further the nature of the possible differential role for recovery of sensory/perceptual information during true versus false recognition, a recent event-related fMRI study by Slotnick and Schacter (2004) used abstract shapes that were expected to engage visual processing regions. All shapes in the study phase were related to nonpresented prototypes, and subjects later made old/new recognition decisions about previously studied shapes, nonstudied related shapes, and nonstudied unrelated shapes (see Figure 1A for examples). Earlier behavioral studies using similar paradigms have shown high levels of false recognition to nonstudied prototype shapes that are perceptually related to previously studied shapes (e.g., Koutstaal et al., 1999; see

also, Posner and Keele, 1968). Moreover, imaging studies of true recognition for pictorial stimuli, that appeared after the early studies of true-false recognition discussed above, have revealed reactivation during retrieval of some of the same visual processing regions that have been associated with perceptual processing during encoding (Wheeler and Buckner, 2003; Wheeler et al., 2000). Similar sensory reactivation effects have been observed during memory for sounds (Nyberg et al., 2000; Wheeler et al., 2000) and motor sequences (Nyberg et al., 2001). These observations, coupled with the suggestive findings considered above from previous imaging studies of true versus false recognition (Schacter et al., 1996a; Cabeza et al., 2001; Okado and Stark, 2003), led Slotnick and Schacter (2004) to hypothesize that true recognition of previously studied shapes, as compared to false recognition of nonstudied related shapes, would be accompanied by a sensory signature involving preferential activation of regions within the ventral visual processing stream (Mishkin et al., 1983; Felleman and Van Essen, 1991).

Slotnick and Schacter (2004) also inquired about the nature of the hypothesized sensory signature, specifically, whether observed sensory reactivation effects were accompanied by conscious recollection of sensory features of studied shapes. As noted earlier, previous behavioral studies of true versus false recognition have shown that true recognition is associated with increased sensory memory (e.g., Mather et al., 1997; Norman and Schacter, 1997; Schooler et al., 1986), and it has generally been assumed that such memory reflects conscious recollection of sensory/perceptual details of previously studied items. In contrast, a separate line of research concerning the phenomenon of *perceptual priming* suggests that sensory memory can sometimes be expressed as nonconscious or implicit memory. Perceptual priming occurs when exposure to a stimulus results in enhanced identification or production of that stimulus on a later test (Tulving and Schacter, 1990). Priming can occur in the absence of conscious recollection, as demonstrated by behavioral studies of healthy subjects and by neuropsychological studies showing intact priming in amnesic patients with MTL damage who have severe deficits in conscious memory (for reviews, see Roediger and McDermott, 1993; Schacter and Buckner, 1998). Importantly, priming can reflect memory for sensory attributes of studied items, such as modality of presentation, and even more specific sensory features

such as shape of an object or typefont of a word. Therefore, it is conceivable that sensory reactivation effects during true recognition could be expressed as priming rather than conscious recollection.

Using event-related fMRI, Slotnick and Schacter attempted to determine whether activity in visual processing regions is specifically related to conscious memory, which would be indicated by greater activity during "old" than "new" responses to studied items (see Wheeler and Buckner, 2003, 2004). Alternatively, if activation is comparable during "old" and "new" responses to studied items, then it is reasonable to infer that this activity is associated with a nonconscious or implicit form of memory (see Rugg et al., 1998, for elaboration of this logic).

Behavioral data revealed that participants made significantly more "old" responses to studied shapes than to related nonstudied shapes; although this difference was small in magnitude, it indicates a somewhat greater degree of memory for studied shapes. Critically, they also made significantly more "old" responses to related nonstudied shapes (e.g., prototypes) than to unrelated nonstudied shapes, confirming the presence of a strong false recognition effect. Analysis of the fMRI results in cortical regions within the ventral visual processing stream revealed that, consistent with the sensory reactivation hypothesis, in early visual areas there was significantly greater activity during true than false recognition (see Figure 1B). By contrast, later visual areas showed comparable levels of activity during true and false recognition.

To examine further the nature of responses in early and late visual processing regions, Slotnick and Schacter compared activation associated with "old" responses to studied shapes (old-hits) and "new" responses to studied shapes (old-misses). The contrast between old-hits and old-misses was employed to reveal those regions associated with conscious memory, as discussed above. As illustrated in Figure 1C, a striking dissociation was observed. Compared with the baseline of correct rejections to new unrelated items, *both* old-hits and old-misses were similarly associated with activity in early visual processing regions (BA17, BA18). By contrast, the old-hits greater than old-misses comparison, which is assumed to reflect conscious memory, was associated specifically with activity in late visual processing regions (BA19, BA37; we will discuss this finding later in the article). Thus, activity in early visual processing areas may reflect specific memory for a particular shape, regardless of participants' conscious judgments of whether they had seen the item.

This conclusion is consistent with the priming hypothesis noted earlier: true and false recognition may be distinguished by a nonconscious perceptual priming effect that occurs only for studied items. However, this idea could be questioned because some old-misses may have occurred when participants were unsure about whether they had seen a shape earlier. Thus, the indistinguishable activations for old-hits and old-misses in early visual processing regions might be attributable to "unsure" old-miss responses that were close to the "old" response threshold. That is, rather than representing a nonconscious priming effect, the old-miss-related activity observed by Slotnick and Schacter might reflect vague or weak explicit memory. However, the same pat-

tern of activation for old-hits and old-misses was observed in early visual regions even when subjects said they were "sure" that they not seen a shape previously, thus casting doubt on the "threshold" hypothesis. Although one must always be cautious in neuroimaging studies when interpreting an absence of a significant difference between two conditions, Slotnick and Schacter replicated these results in a follow-up study.

The priming hypothesis might also be questioned because repetition priming effects in neuroimaging studies are most commonly expressed by reduced activity in a particular region during a primed condition compared to an unprimed condition (for reviews, see Henson, 2003; Schacter and Buckner, 1998; Schacter et al., 2004; Wiggs and Martin, 1998). However, this conclusion is based primarily on studies in which familiar items are used as experimental stimuli, such as common words or pictures of everyday objects. In studies that have used unfamiliar materials more akin to the Slotnick and Schacter abstract shapes, such as faces (Henson et al., 2000) or drawings of structurally possible and impossible objects (Schacter et al., 1995), priming-related increases have been observed. Although more evidence is required to fully evaluate the priming hypothesis advanced by Slotnick and Schacter, it remains viable in light of known neuroimaging data concerning priming.

Electrophysiological Differences between True and False Recognition

Studies using event-related potentials (ERPs), which measure changes in voltage topography on the scalp over time, have been conducted to probe the underlying neural activity associated with true and false recognition in real time. Although ERPs do not allow the precise spatial localization achieved by PET and fMRI, they do permit more fine-grained analyses of the temporal properties of activity accompanying memory retrieval. Moreover, ERP studies have established a foundation of basic knowledge concerning properties of memory encoding and retrieval that are associated with specific aspects of the ERP waveform. One ERP component that seems especially relevant to the sensory reactivation hypothesis under consideration is a parietal effect (often left lateralized) that occurs within the 400–800 ms range after stimulus onset and typically shows greater activity during hits to old items than correct rejections to new items (Wilding and Rugg, 1996; Donaldson and Rugg, 1998; Curran and Cleary, 2003). This parietal effect has been associated previously with memory for contextual/perceptual details (Wilding, 2000).

Early studies of ERP effects associated with true and false recognition used the DRM semantic associates paradigm and measured ERP activity during retrieval. The main message from early ERP studies was that brain activity during true and false recognition is highly similar (Düzel et al., 1997; Johnson et al., 1997). Johnson et al. reported some differences under blocked testing conditions (i.e., old items and related lures were tested in separate blocks) that resembled those initially reported by Schacter et al. (1996a) with PET. However, those differences were not observed when trial types were intermixed, leading Johnson et al. to postulate that differences in retrieval strategy during blocked and intermixed testing accounted for observed true-false recognition ERP differences.

As with fMRI, however, more recent ERP studies have

revealed more consistent true-false recognition differences. Employing a DRM paradigm, Curran et al. (2001) measured ERPs at test to old words, related words, and new words. They found that the left parietal effect (400–800 ms) was greater for true recognition (old-hits) than for false recognition (related-false alarms) in a subset of their subjects (see Curran et al., 2001, for discussion). Fabiani et al. (2000) also used a DRM paradigm, but presented words at study to the left or right of fixation, whereas at test they were presented centrally. At central and posterior (including parietal) electrodes, a parietal effect (210–700 ms) was greater for true recognition as compared to false recognition; furthermore, this effect occurred on the scalp contralateral to where the word was presented at study (i.e., words previously presented in the right hemifield at study elicited effects on the left scalp, and vice versa). If this parietal effect reflects the same processes as those noted earlier, this result is consistent with the notion that the relatively early parietal effect reflects perceptual/contextual details associated with remembering (Wilding, 2000). Specifically, these contralateral parietal effects may reflect reactivation of contralateral visual processing regions during memory for the lateralized word from study, as in the previously discussed memory-related sensory reactivation reported with fMRI. If so, these true greater than false recognition memory related effects would parallel the differential sensory effects reported by Slotnick and Schacter (2004). Further supporting this suggestion, some evidence suggests that this differential ERP activity for true and false recognition may not have been accessible to conscious awareness, because participants in a similar study were not able to reliably identify the side of initial presentation (Gratton et al., 1997).

Similar to the Curran et al. (2001) and Fabiani et al. (2000) studies just described, the true greater than false recognition parietal effect has been shown to be consistent across a number of studies employing semantic associative word lists, categorized word lists, unrelated word lists, and unrelated picture sets (Nessler et al., 2001; Nessler and Mecklinger, 2003; Walla et al., 2000; Endl et al., 1999). To the extent that the parietal effect is a marker for sensory reactivation during memory retrieval, these differential true greater than false recognition sensory reactivation effects likely map directly onto the differential true greater than false recognition sensory reactivation effects discussed with fMRI (Schacter et al., 1996a; Cabeza et al., 2001; Slotnick and Schacter, 2004). Moreover, the ERP results suggest that these effects occur relatively early in time, likely before a manual response is made (which usually takes about 1 s in these tasks; e.g., Endl et al., 1999; Fabiani et al., 2000; Nessler et al., 2001; Nessler and Mecklinger, 2003). Thus, the differential true versus false recognition sensory effects appear to be relatively fast in nature. Such converging evidence from ERP and fMRI studies provides confidence that some form of the sensory reactivation hypothesis is viable.

Brain Regions that Support False Memories: Clues from Neuropsychology and Neuroimaging

One of the fundamental questions concerning false memories centers on why they occur at all: why should people claim to remember events that never happened?

Cognitive studies have delineated a range of explanations for different types of memory distortions (for discussion, see Alba and Hasher, 1983; Brainerd and Reyna, 1993; Bartlett, 1932; Johnson et al. 1993; Loftus, 1993; Roediger, 1996; Roediger et al., 2001; Schacter et al., 1998a; Schacter, 1999, 2001). For example, analyses of the DRM false recognition effect have typically appealed to the idea that false recognition reflects memory for the general sense or semantic gist of previously studied lists or that false recognition is based on previously generating a lure item during the study phase of the experiment and then failing at test to discriminate between the generated lure and studied items.

With respect to brain activity, because much evidence has implicated the hippocampus and related structures in the formation and retrieval of true memories, a number of studies have focused on the possible role of this region in the generation of false memories. Recent studies have begun to explore the possible contributions of other regions, as well as to examine how events that occur at the time of encoding can result in subsequent false memories. In this section, we begin by considering neuropsychological studies of amnesic patients with MTL damage. We then consider neuroimaging studies of memory retrieval, and conclude with recent studies that have focused on encoding.

False Recognition and Amnesic Patients

Using the DRM semantic associates procedure described earlier, Schacter et al. (1996b) gave amnesic patients and matched controls lists of semantic associates and tested their recognition of studied and nonstudied words. Amnesic patients, as expected, showed reduced levels of veridical recognition memory compared to the matched controls: the amnesics attained fewer hits than controls to words that had actually appeared on the lists and also made more false alarms than controls to new words that were unrelated to the words from the study lists. These findings are typical of recognition memory experiments with amnesic patients. The more informative result involved false recognition of semantic associates. In comparison to normal controls, amnesic patients made significantly fewer false alarms to nonpresented theme words. Thus, amnesics were less susceptible to false recognition of associatively related words, even though they made more false alarms to unrelated words than did normal controls (see also Melo et al., 1999; Schacter et al., 1998b). Schacter et al. (1997b) compared false recognition of nonstudied words that were related either conceptually (e.g., "twister," "funnel") or perceptually (e.g., "hate," "mate") to previously studied words. In two experiments, amnesic patients showed similarly reduced levels of false recognition for perceptually and conceptually related nonstudied words, thus extending the earlier results. Koutstaal et al. (1999) examined false recognition of abstract visual patterns in amnesic patients, using a prototype recognition procedure that is similar in some respects to the paradigm used in Slotnick and Schacter's (2004) fMRI study described earlier. Koutstaal et al. found robust false recognition to prototypes in control subjects, together with significantly reduced false recognition of prototypes in amnesic patients.

A parallel series of studies on false recognition has been reported in patients with Alzheimer's disease (AD), who are generally characterized by neuropathology that

includes, but is not limited to, MTL regions. AD patients typically exhibit patterns of episodic memory performance that are generally similar to those seen in amnesic patients. Results thus far from AD patients are highly similar to those from amnesics: compared with age-matched controls, AD patients show reduced false recognition of lure items that are either semantically or perceptually related to previously studied items (Balota et al., 1999; Budson et al., 2000, 2001, 2003).

One interpretation of the foregoing results is that healthy controls form and retain a well-organized representation of the semantic or perceptual gist of a list of related study items (cf. Payne et al., 1996; Reyna and Brainerd, 1995; Schacter et al., 2001). Lure items that match semantic or perceptual features of this representation are likely to be falsely recognized; unrelated words that do not match it are likely to be correctly rejected. As a result of MTL damage, amnesic and AD patients may form and retain only a weak or degraded gist representation. Thus, patients are less likely than controls to make false alarms to semantic associates or perceptually similar items (because there is little information for theme words to match) and are more likely than controls to make false alarms to unrelated new items (because the degraded gist representation does not allow amnesic and AD patients to easily exclude words that do not match the theme or gist of the list). These findings therefore suggest that the MTL structures that are damaged in amnesic and AD patients may participate in storing the gist information that underlies false recognition in the foregoing paradigms.

An alternative possibility is that on a standard recognition memory test, which probes memory for specific individual items, amnesic and AD patients fail to retrieve or gain access to relatively intact gist representations. Perhaps gist information is available and could be accessed by patients if a test were to probe gist memory directly. Evidence bearing on this point comes from a study by Verfaellie et al. (2002) that used a modified version of the DRM semantic associates procedure in which participants are instructed to call "old" any item that is semantically related to the theme or gist of a previously studied list, even if the item itself had not appeared on the list. Evidence from healthy control participants suggests that such a task provides a more direct probe of gist information than a standard old/new recognition task (Brainerd and Reyna, 1998; Koutstaal, 2003; Schacter et al., 2001). Verfaellie et al. (2002) reported that, even on this "meaning test," amnesic patients provide fewer "old" responses to semantically related lure words than do controls, thereby supporting the idea of a degraded gist representation.

There are, however, conditions in which amnesic and AD patients show *greater* false recognition than control subjects. When DRM lists are repeatedly studied and tested, allowing control subjects to build up detailed memories of the specific words they encountered, their false recognition rates drop considerably, to a level below that of amnesic and AD patients (Budson et al., 2000; Schacter et al., 1998b). With repeated study and test, patients build up their degraded gist representation but do not develop specific memories of list items, resulting in increased false recognition. This finding highlights that when experimental conditions promote high

levels of specific recollection, amnesic and AD patients are likely to show greater false recognition than controls, because only the controls can use specific information to oppose or reduce gist-based false alarms (see also, Budson et al., 2002; Verfaellie et al., 2004).

Neuroimaging of True and False Recognition

Although the foregoing studies are consistent with the idea that the MTL can contribute to the generation of some types of false memories, there are limits on the neuroanatomical specificity of the conclusions that can be drawn from such studies, since lesions in amnesic patients are rarely limited to specific regions within the MTL and brain damage in AD patients typically extends to other regions. Functional neuroimaging studies hold the promise for more precise spatial localization.

Early neuroimaging evidence suggested comparable activation of MTL structures during true and false recognition (Schacter et al., 1996a, 1997a). However, these findings emerged from contrasting true and false recognition with a low-level fixation baseline and thus provide relatively weak evidence for the hypothesis that activation of MTL structures contributes significantly to false recognition. More compelling evidence was provided by Cabeza et al. (2001), who found significant activation of the hippocampus during both true and false recognition in the DRM paradigm compared with a control condition in which subjects responded to unrelated lure words. Moreover, whereas hippocampal activations were indistinguishable during true and false recognition, as noted earlier, parahippocampal gyrus showed greater activation during true than false recognition, pointing toward a dissociation among MTL structures. More recently, Slotnick and Schacter (2004) also reported some evidence for hippocampal activity during both true and false recognition compared with responses to unrelated control items. Thus, while further neuroimaging data are required to settle the issue, there is evidence consistent with the idea that hippocampal activation during retrieval may contribute to the generation of false memories.

Slotnick and Schacter (2004) reported stronger evidence implicating other regions in the generations of false memories. As noted earlier, they found evidence for activation of early and late visual processing regions during both true recognition of studied shapes and false recognition of related shapes, compared with correct rejection of unrelated new shapes (but only the late visual processing regions showed comparable activation during true and flase recognition). Moreover, when they further examined responses to studied shapes by contrasting old-hits with old-misses (a significant difference in this contrast was assumed to reflect conscious memory), they found greater activity in the late visual processing regions (BA19, BA37) during old-hits than old misses (Figure 1C). Slotnick and Schacter found a similar pattern of activity in a left parietal region that has recently been implicated in endorsing items as "old" (Wheeler and Buckner, 2003). These findings raise the possibility that a network of regions, perhaps a subset of those activated during true recognition, become active during false recognition and contribute to the generation of an "old" response. However, this hypothesis must be regarded as tentative at the present time. Further studies using the type of old-hit versus old-miss contrast re-

ported by Slotnick and Schacter will be needed before the issue can be settled (for additional fMRI findings concerning subjective experience and retrieval of episodic memories, see Eldridge et al., 2000).

Encoding Studies and the Formation of False Memories

Further clues concerning how brain activity contributes to the generation of false memories are provided by studies that have examined the role of encoding processes. Upon first consideration, the notion that encoding processes are related to subsequent false memories may seem odd or even nonsensical: false memories refer to inaccurate recollections of events that never happened, so it makes little sense to talk about how encoding processes influence such memories. However, it is known that events during encoding can influence false recognition rates in the DRM paradigm (e.g., Arndt and Reder, 2003; Schacter et al., 1999). Moreover, in reality monitoring paradigms where false recognition occurs because subjects incorrectly claim to have seen an item they only imagined (e.g., Johnson and Raye, 1981), subjects do actually encode a specific event (perceiving or imagining objects), so it is possible to examine how brain activity during this event is related to subsequent recognition performance.

Gonsalves and Paller (2000) recorded ERPs while subjects viewed pictures of common objects or generated images of objects. On a later source monitoring test, when subjects were required to indicate whether an object had earlier been perceived or imagined, they sometimes claimed incorrectly that they had perceived an object that they had only imagined. Posterior ERPs at encoding were more positive when subjects falsely recognized pictures that they had only imagined, perhaps reflecting more vivid visual imagery at encoding for items later subject to false recognition. In a more recent fMRI version of this study, Gonsalves et al. (2004) again scanned subjects during encoding while viewing pictures or generating images of objects. Gonsalves et al. found that false memories of having seen imagined objects were predicted by activation, at the time of encoding, in regions associated with visual imagery and spatial attention, including precuneus and inferior parietal cortex. Presumably, the engagement of such regions during encoding reflected greater perceptual detail associated with the visual mental image, making it more difficult for subjects to distinguish true from false memories on the later test.

Monitoring and Reduction of False Memories: A Role for Prefrontal Cortex?

The preceding section examined some of the brain regions that contribute to the formation or generation of false memories. However, it is important to keep in mind that memory is often accurate; false memories, though important when they occur, are by no means the dominant outcome of memory retrieval. From this perspective, we can ask about regions that may contribute to reducing or avoiding false memories. Much attention has been focused on regions within prefrontal cortex, in part because a good deal of recent neuroimaging work has implicated a subset of prefrontal regions in strategic retrieval activities, such as specifying cues for

retrieval and monitoring or verifying the source of memories (e.g., Dobbins et al., 2002, 2003; Henson et al., 1999; Rugg et al., 1999). Here we first consider evidence from neuropsychological studies that have documented increased false recognition in patients with damage to various regions within prefrontal cortex. Then we consider relevant findings from neuroimaging studies.

False Recognition and Frontal Lobe Damage

Although neuropsychologists have long been interested in memory distortions such as confabulation, where patients generate detailed recollections of events that never happened (e.g., Burgess and Shallice, 1996; Dab et al., 1999; Johnson, 1991; Moscovitch, 1995; Talland, 1965), during the past decade several groups have reported detailed studies of patients with damage to regions in the frontal lobes, who exhibit heightened false recognition.

One of the first such cases was reported by Delbecq-Derouesne et al. (1990). They described a patient, RW, who suffered a ruptured aneurysm of the anterior communicating artery (which typically results in damage to the frontal lobes as well as basal forebrain). RW made large numbers of false alarms on recognition tests and expressed considerable confidence in his erroneous memories. Schacter et al. (1996c) reported pathological false recognition in a patient (BG) who had suffered an infarction restricted to the right frontal lobe. For example, after BG studied a list of common words, memory for previously studied words (i.e., hit rate) was relatively normal, but he made many more false alarms (claiming to recognize a word that did not actually appear in the study list) than did any of the eight matched control subjects. Schacter et al. found that BG exhibited increased false recognition relative to matched controls when he was tested with pseudowords, environmental sounds, or words that are either associatively related or unrelated to previously studied words. However, they also found a way to reduce BG's false recognition responses. After showing BG a study list consisting of pictures of inanimate objects from various categories (e.g., furniture, articles of clothing), on the subsequent recognition test, he was presented with items that were on the study list as well as new items that did not appear in the study list. Some of these new items were members of the categories used in the study list, others came from miscellaneous categories of inanimate objects that were not represented on the study list, and still others were animate objects (i.e., animals). BG claimed to recognize many of the new items that were drawn from previously studied categories of inanimate objects, but he almost never said he recognized nonstudied items that were not members of previously studied categories. False recognition was thus restricted to items associated with the previously studied categories. Thus, Schacter et al. (1996c) suggested that when making a recognition decision BG relies excessively on the global correspondence between a test item and previously studied words. Follow up experiments with BG (Curran et al., 1997) suggested that impaired retention of specific information about individual items on the list contributed to his overreliance on general features of the study episode.

We must be cautious about drawing general conclusions from this single case study, especially regarding

the role of the right frontal lobe in false recognition. For instance, Parkin et al. (1996) reported elevated false recognition in a patient, JB, who had suffered a ruptured anterior communicating artery aneurysm that produced left frontal lobe damage. A series of experiments revealed a performance profile that in many respects resembled that of right frontal patient BG: false recognition of nonpresented materials was consistent across various types of materials and accompanied by high confidence but was nearly eliminated when lure items were perceptually dissimilar to studied items. Subsequent research with this same patient (Parkin et al., 1999) suggested that JB's problems result from a poorly focused retrieval description that is normally used to guide memory search (Burgess and Shallice, 1996; Schacter et al., 1998a). The poorly focused retrieval description, in turn, was linked to encoding deficits that resulted in generic representations of target items that lack specific details. In any event, the overall pattern of results suggests that, when making recognition judgments, patients JB and BG both rely too much on global similarities between study and test items, probably reflecting deficits in strategic retrieval/monitoring operations combined with some encoding problems.

Rapcsak and his collaborators have provided additional evidence for strategic retrieval deficits in studies that focused on false recognition of faces in patients with frontal lobe damage. Early studies by Rapcsak's group demonstrated that patients with various types of frontal lobe lesions were prone to high levels of false recognition when presented with unfamiliar faces (e.g., Rapcsak et al., 1996). Subsequent experiments by Rapcsak et al. (1999) indicated that frontal patients often failed to engage in strategic or effortful retrieval monitoring operations that are required when erroneous recognition decisions can be based on a general sense of familiarity with a face. This misleading influence of familiarity can be triggered by the automatic activation of contextual information that matches the previously encountered faces. However, when patients were specifically instructed to use stricter decision criteria that produced more careful retrieval monitoring and were thus less likely to be unduly influenced by misleading familiarity, false recognition of faces was greatly reduced. In fact, false recognition in the frontal patients now no longer differed from that of control subjects. Ward and collaborators have documented a similar type of false facial familiarity in a detailed case study of a patient, MR, whose facial recognition problems were associated with damage to the left frontal lobe (Ward et al., 1999; Ward and Jones, 2003).

In addition to these case reports, group studies have also been reported. Swick and Knight (1999) reported increased false alarms to unrelated new items in a group of patients with damage to dorsolateral prefrontal cortex (10 left, 1 right), which they attributed to an excessively lenient response criterion. Budson et al. (2002) used the DRM paradigm in a study of 13 patients with lesions to dorsolateral frontal cortex (7 right, 5 left, 1 bilateral). These frontal patients showed similar levels of false recognition to controls after a single study and test session but showed significantly higher levels of false recognition than controls after five study/test repetitions. This pattern occurred because controls, but not frontal patients, reduced their false alarms after five repetitions. Verfaellie et al. (2004) reported on a mixed group of patients with unilateral left, right, or bilateral frontal damage. Patients studied categorized lists of words, with either one, four, or eight category members presented on the study list. Although both patients and controls showed increased false alarms to related lures with increasing category size, there were no overall group differences. However, subsequent analyses revealed that a subset of frontal patients did show excessive false alarming; analysis of lesion data indicated no clear difference between these patients and those who showed normal false alarm responses.

Thus, while the literature clearly establishes that some patients with frontal lobe damage do show heightened false recognition under some conditions, the phenomenon is not observed in all such patients in all conditions. Further work will be necessary to decide between theoretical hypotheses that have been proposed to account for the observed deficits.

Neuroimaging of Frontal Function during True and False Recognition

As noted earlier, it is well established that a variety of prefrontal regions show increased activation during standard recognition memory tests. Several of the previously reviewed PET and fMRI studies of false recognition during the DRM and related paradigms have also reported evidence for activation of several prefrontal regions during false recognition, and in some cases, greater activation during false than true recognition. Schacter et al. (1996a) reported that a region in the dorsolateral/anterior prefrontal cortex, which has been associated with retrieval monitoring (e.g., Rugg et al., 1996; Dobbins et al., 2002, 2003), showed greater activity during false than true recognition, perhaps reflecting the need for evaluation or monitoring of the strong sense of familiarity produced by false targets. Schacter et al. (1997a) found some trends indicating greater activation of right anterior prefrontal cortex during false than true recognition, again suggesting a role for retrieval monitoring processes during false recognition. Consistent with this suggestion, analyses of event-related time courses indicated a delayed onset for anterior prefrontal activity during both true and false recognition compared with other brain regions. Although various interpretations of this observation are possible (see Schacter et al., 1997a), it is consistent with the idea that anterior prefrontal activity reflects a late-occurring evaluation or monitoring of the products of retrieval. Cabeza et al. (2001) and Slotnick and Schacter (2004) provided additional evidence for greater activation of right prefrontal cortex during false than true recognition (see also, von Zerssen et al., 2001).

In a different type of false recognition paradigm, Treyer et al. (2003) used PET to scan subjects during a continuous recognition paradigm in which participants viewed a series of pictures, some of which were repeated, and were instructed to respond "no" to new pictures and "yes" to repeated ones. After the initial run, the same pictures were presented again in a separate run. Subjects were instructed to disregard whether a picture appeared in an earlier run and respond "yes" only to picture recurrences *within* the same run. Thus, in this paradigm, false recognition reflects a failure to

remember exactly when an item was encountered previously. The key finding of the study was that a left orbitofrontal region showed increased activation under task conditions that required subjects to make this difficult source discrimination to avoid false recognition. This finding was deemed especially important because this region is often damaged (along with MTL structures) in confabulating amnesic patients who are especially susceptible to source memory confusions on this task (Schnider, 2001; Schnider and Ptak, 1999).

ERP Evidence for a Late Frontal Component

We have discussed studies suggesting that a relatively early parietal component of the ERP waveform is related to sensory reactivation of studied items. In addition, ERP studies of recognition memory have revealed a later (often right lateralized) frontal effect that occurs within 1000–1600 ms after stimulus onset and is thought to be related to postretrieval monitoring. Several ERP studies have provided evidence that the late-occurring frontal component is similar for true and false recognition (Düzel et al., 1997; Curran et al., 2001; Nessler et al., 2001; Nessler and Mecklinger, 2003). However, there is also evidence that false recognition can evoke greater frontal activity than true recognition (Fabiani et al., 2000; Goldmann et al., 2003). As noted above, if some frontal regions do support a postretrieval monitoring function, such monitoring may be increased during false recognition, with a concomitant relatively greater level of activity (Goldmann et al., 2003). Curran et al. (2001) performed a median split to compare high performing subjects, who discriminated well between studied and nonstudied related items, with low performing subjects, who did not. The high performing subjects were characterized by more positive late right frontal ERPs than were low performing subjects, perhaps reflecting retrieval monitoring operations that were more likely to be engaged by high than low performing subjects. Overall, it seems likely that these later frontal effects map onto specific regions of the prefrontal cortex regions that are associated with both true and false recognition (Cabeza et al., 2001; Slotnick and Schacter, 2004).

Conclusion

The evidence reviewed in this paper indicates that, over the past decade, a cognitive neuroscience of memory distortion has begun to emerge. With respect to the three issues that we considered, we have seen that fMRI and ERP evidence converge on the conclusion that there is a neural sensory signature that can distinguish between true and false recognition, and that neuroimaging and neuropsychological evidence converge on the conclusions that regions within the medial temporal lobe are related to the generation of false recognition, whereas regions within prefrontal cortex are related to memory monitoring activities.

As the field moves forward and new questions arise, we are hopeful that approaches emphasizing converging evidence will continue to develop. For example, Simons et al. (2004) recently reported evidence concerning false recognition in patients with semantic dementia who have cognitive deficits largely restricted to semantic memory, primarily as a result of damage to the anterolateral temporal lobe. Such patients can show normal

episodic memory. Using a categorized pictures paradigm in which healthy controls frequently show false recognition of related lures (nonstudied pictures from studied categories), Simons et al. found reduced false recognition of related lures in semantic dementia patients despite normal memory for studied pictures. A further experiment indicated normal false recognition to related lures (and intact true recognition) when meaningless shapes were used. Taken together, the results suggest that anterolateral temporal cortex is necessary for some aspect of memory for semantic gist, but not perceptual gist (or item-specific memory). Given these results, it would be highly desirable to now initiate neuroimaging studies with similar paradigms that examine the distinction between semantic versus perceptual gist, further probing the role of anterolateral temporal cortex. Using this type of approach, we are hopeful that the cognitive neuroscience of memory distortion will progress even more rapidly during the coming decade than it has during the last.

References

Alba, J.W., and Hasher, L. (1983). Is memory schematic? Psychol. Bull. *93*, 203–231.

Arndt, J., and Reder, L.M. (2003). The effect of distinctive visual information on false recognition. J. Mem. Lang. *48*, 1–15.

Balota, D.A., Watson, J.M., Duchek, J.M., and Ferraro, F.R. (1999). Cross-modal semantic and homograph priming in healthy young, healthy old, and in Alzheimer's disease individuals. J. Int. Neuropsychol. Soc. *5*, 626–640.

Bar, M., and Aminoff, E. (2003). Cortical analysis of visual context. Neuron *38*, 347–358.

Bartlett, F.C. (1932). Remembering: A Study in Experimental and Social Psychology (London: Cambridge University Press).

Brainerd, C.J., and Reyna, V.F. (1993). Memory independence and memory interference in cognitive development. Psychol. Rev. *100*, 42–67.

Brainerd, C.J., and Reyna, V.F. (1998). When things that were never experienced are easier to "remember" than things that were. Psychol. Sci. *9*, 484–489.

Brainerd, C.J., Payne, D.G., Wright, R., and Reyna, V.F. (2003). Phantom recall. J. Mem. Lang. *48*, 445–467.

Budson, A.E., Daffner, K.R., Desikan, R., and Schacter, D.L. (2000). When false recognition is unopposed by true recognition: gist-based memory distortion in Alzheimer's disease. Neuropsychology *14*, 277–287.

Budson, A.E., Desikan, R., Daffner, K.R., and Schacter, D.L. (2001). Perceptual false recognition in Alzheimer's disease. Neuropsychology *15*, 230–243.

Budson, A.D., Sullivan, A.L., Mayer, E., Daffner, K.R., Black, P.M., and Schacter, D.L. (2002). Suppression of false recognition in Alzheimer's disease and in patients with frontal lobe lesions. Brain *125*, 2750–2765.

Budson, A.E., Sullivan, A.L., Daffner, K.R., and Schacter, D.L. (2003). Semantic versus phonological false recognition in aging and Alzheimer's disease. Brain Cogn. *51*, 251–261.

Burgess, P.W., and Shallice, T. (1996). Confabulation and the control of recollection. Memory *4*, 359–411.

Cabeza, R., Rao, S.M., Wagner, A.D., Mayer, A.R., and Schacter,

D.L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. Proc. Natl. Acad. Sci. USA 98, 4805–4810.

Curran, T., and Cleary, A.M. (2003). Using ERPs to dissociate recollection and familiarity in picture recognition. Brain Res. Cogn. Brain Res. 15, 191–205.

Curran, T., Schacter, D.L., Norman, K.A., and Galluccio, L. (1997). False recognition after a right frontal lobe infarction: memory for general and specific information. Neuropsychologia 35, 1035–1049.

Curran, T., Schacter, D.L., Johnson, M.K., and Spinks, R. (2001). Brain potentials reflect behavioral differences in true and false recognition. J. Cogn. Neurosci. 13, 201–216.

Dab, S., Claes, T., Morais, J., and Shallice, T. (1999). Confabulation with a selective descriptor process impairment. Cogn. Neuropsychol. 16, 215–242.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. J. Exp. Psychol. 58, 17–22.

Delbecq-Derouesne, J., Beauvois, M.F., and Shallice, T. (1990). Preserved recall versus impaired recognition. A case study. Brain 113, 1045–1074.

Dobbins, I.G., Foley, H., Schacter, D.L., and 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., and Schacter, D.L. (2003). Memory orientation and success: separable neurocognitive components underlying episodic recognition. Neuropsychologia 41, 318–333.

Donaldson, D.I., and Rugg, M.D. (1998). Recognition memory for new associations: electrophysiological evidence for the role of recollection. Neuropsychologia 36, 377–395.

Düzel, E., Yonelinas, A.P., Mangun, G.R., Heinze, H.J., and Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. Proc. Natl. Acad. Sci. USA 94, 5973–5978.

Eldridge, L.L., Knowlton, B.T., Furmanski, C.S., Bookeheimer, S.Y., and Engel, S.A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. Nat. Neurosci. 3, 1149–1152.

Endl, W., Walla, P., Lindinger, G., Deecke, L., and Lang, W. (1999). Event-related potential correlates of false recognitions of faces. Neurosci. Lett. 265, 115–118.

Fabiani, M., Stadler, M.A., and Wessels, P.M. (2000). True but not false memories produce a sensory signature in human lateralized brain potentials. J. Cogn. Neurosci. 12, 941–949.

Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–47.

Gallo, D.A., Roediger, H.L., III, and McDermott, K.B. (2001). Associative false recognition occurs without strategic criterion shifts. Psychon. Bull. Rev. 8, 579–586.

Goldmann, R.E., Sullivan, A.L., Droller, D.B.J., Rugg, M.D., Curran, T., Holcomb, P.J., Schacter, D.L., Daffner, K.R., and Budson, A.E. (2003). Late frontal brain potentials distinguish true and false recognition. Neuroreport 14, 1717–1720.

Gonsalves, B., and Paller, K.A. (2000). Neural events that underlie remembering something that never happened. Nat. Neurosci. 3, 1316–1321.

Gonsalves, B., Reber, P.J., Gitelman, D.R., Parrish, T.B., Mesulam, M.M., and Paller, K.A. (2004). Neural evidence that vivid imagining can lead to false remembering. Psychol. Sci., in press.

Gratton, G., Corballis, P.M., and Jain, S. (1997). Hemispheric organization of visual memories. J. Cogn. Neurosci. 9, 92–104.

Henson, R.N. (2003). Neuroimaging studies of priming. Prog. Neurobiol. 70, 53–81.

Henson, R.N., Shallice, T., and Dolan, R.J. (1999). Right prefrontal cortex and episodic memory retrieval: a functional MRI test of the monitoring hypothesis. Brain 122, 1367–1381.

Henson, R., Shallice, T., and Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. Science 287, 1269–1272.

Johnson, M.K. (1991). Reality monitoring: evidence from confabulation in organic brain disease patients. In Awareness of Deficit After Brain Injury: Clinical and Theoretical Issues, G.P. Prigatano and D.L. Schacter, eds. (New York: Oxford University Press), pp. 176–197.

Johnson, M.K., and Raye, C.L. (1981). Reality monitoring. Psychol. Rev. 88, 67–85.

Johnson, M.K., Hashtroudi, S., and Lindsay, D.S. (1993). Source monitoring. Psychol. Bull. 114, 3–28.

Johnson, M.K., Nolde, S.F., Mather, M., Kounios, J., Schacter, D.L., and Curran, T. (1997). The similarity of brain activity associated with true and false recognition memory depends on test format. Psychol. Sci. 8, 250–257.

Koriat, A., and Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. Psychol. Rev. 103, 490–517.

Koutstaal, W. (2003). Older adults encode—but do not always use—perceptual details: intentional versus unintentional effects of detail on memory judgments. Psychol. Sci. 14, 189–193.

Koutstaal, W., Schacter, D.L., Verfaellie, M., Brenner, C., and Jackson, E.M. (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, E.F. (1993). Desperately seeking memories of the first few years of childhood: the reality of early memories. J. Exp. Psychol. Gen. 122, 274–277.

Mather, M., Henkel, L.A., and Johnson, M.K. (1997). Evaluating characteristics of false memories: remember/know judgments and memory characteristics questionnaire compared. Mem. Cognit. 25, 826–837.

Melo, B., Winocur, G., and Moscovitch, M. (1999). False recall and false recognition: an examination of the effects of selective and combined lesions to the medial temporal lobe/diencephalon and frontal lobe structures. Cogn. Neuropsychol. 16, 343–359.

Mishkin, M., Ungerleider, L.G., and Macko, K.A. (1983). Object vision and spatial vision: two cortical pathways. Trends Neurosci. 6, 414–417.

Moscovitch, M. (1995). Confabulation. In Memory Distortions: How Minds, Brains, and Societies Reconstruct the Past, D.L. Schacter, ed. (Cambridge, MA: Harvard University Press), pp. 226–251.

Nessler, D., and Mecklinger, A. (2003). ERP correlates of true and false recognition after different retention delays: stimulus- and response-related processes. Psychophysiology 40, 146–159.

Nessler, D., Mecklinger, A., and Penney, T.B. (2001). Event related brain potentials and illusory memories: the effects of differential encoding. Brain Res. Cogn. Brain Res. 10, 283–301.

Neuschatz, J.S., Benoit, G.E., and Payne, D.G. (2003). Effective warnings in the Deese-Roediger-McDermott false-memory paradigm: the role of identifiability. J. Exp. Psychol. Learn. Mem. Cogn. 29, 35–41.

Norman, K.A., and Schacter, D.L. (1997). False recognition in younger and older adults: exploring the characteristics of illusory memories. Mem. Cognit. 25, 838–848.

Nyberg, L., Habib, R., McIntosh, A.R., and Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. Proc. Natl. Acad. Sci. USA 97, 11120–11124.

Nyberg, L., Petersson, K.M., Nilsson, L.G., Sandblom, J., Åberg, C., and Ingvar, M. (2001). Reactivation of motor brain areas during explicit memory for actions. Neuroimage 14, 521–528.

Okado, Y., and Stark, C. (2003). Neural processing associated with true and false retrieval. Cogn. Affect. Behav. Neurosci. 3, 323–334.

Parkin, A.J., Bindschaedler, C., Harsent, L., and Metzler, C. (1996). Pathological false alarm rates following damage to the left frontal cortex. Brain Cogn. 32, 14–27.

Parkin, A.J., Ward, J., Bindschaedler, C., Squires, E.J., and Powell, G. (1999). Recognition following frontal lobe damage: the role of encoding factors. Cogn. Neuropsychol. 16, 243–265.

Payne, D.G., Elie, C.J., Blackwell, J.M., and Neuschatz, J.S. (1996). Memory illusions: recalling, recognizing, and recollecting events that never occurred. J. Mem. Lang. 35, 261–285.

Posner, M.I., and Keele, S.W. (1968). On the genesis of abstract ideas. J. Exp. Psychol. *77*, 353–363.

Rapcsak, S.Z., Polster, M.R., Glisky, M.L., and Comer, J.F. (1996). False recognition of unfamiliar faces following right hemisphere damage: neuropsychological and anatomical observations. Cortex *32*, 593–611.

Rapcsak, S.Z., Reminger, S.L., Glisky, E.L., Kaszniak, A.W., and Comer, J.F. (1999). Neuropsychological mechanisms of false facial recognition following frontal lobe damage. Cogn. Neuropsychol. *16*, 267–292.

Reyna, V.F., and Brainerd, C.J. (1995). Fuzzy-trace theory: an interim synthesis. Learn. Individ. Differ. *7*, 1–75.

Robinson, K.J., and Roediger, H.L., III. (1997). Associative processes in false recall and false recognition. Psychol. Sci. *8*, 231–237.

Roediger, H.L., III. (1996). Memory illusions. J. Mem. Lang. *35*, 76–100.

Roediger, H.L., III, and McDermott, K.B. (1993). Implicit memory in normal human subjects. In Handbook of Neuropsychology, Volume 8, F. Boller and J. Grafman, eds. (Amsterdam: Elsevier), pp. 63–131.

Roediger, H.L., III, and McDermott, K.B. (1995). Creating false memories: remembering words not presented in lists. J. Exp. Psychol. Learn. Mem. Cogn. *21*, 803–814.

Roediger, H.L., III, and McDermott, K.B. (2000). Tricks of memory. Curr. Dir. Psychol. Sci. *9*, 123–127.

Roediger, H.L., III, Watson, J.M., McDermott, K.B., and Gallo, D.A. (2001). Factors that determine false recall: a multiple regression analysis. Psychon. Bull. Rev. *8*, 385–407.

Rugg, M.D., Fletcher, P.C., Frith, C.D., Frackowiak, R.S.J., and Dolan, R.J. (1996). Differential activation of the prefrontal cortex in successful and unsuccessful memory retrieval. Brain *119*, 2073–2083.

Rugg, M.D., Mark, R.E., Walla, P., Schloerscheidt, A.M., Birch, C.S., and Allen, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. Nature *392*, 595–598.

Rugg, M.D., Fletcher, P.C., Chua, P.M., and Dolan, R.J. (1999). The role of the prefrontal cortex in recognition memory and memory for source. Neuroimage *10*, 520–529.

Schacter, D.L. (1999). The seven sins of memory. Insights from psychology and cognitive neuroscience. Am. Psychol. *54*, 182–203.

Schacter, D.L. (2001). Seven Sins of Memory: How the Mind Forgets and Remembers (New York: Houghton Mifflin Company).

Schacter, D.L., and Buckner, R.L. (1998). Priming and the brain. Neuron *20*, 185–195.

Schacter, D.L., Reiman, E., Uecker, A., Polster, M.R., Yun, L.S., and Cooper, L.A. (1995). Brain regions associated with retrieval of structurally coherent visual information. Nature *376*, 587–590.

Schacter, D.L., Reiman, E., Curran, T., Yun, L.S., Bandy, D., McDermott, K.B., and Roediger, H.L., III. (1996a). Neuroanatomical correlates of veridical and illusory recognition memory: evidence from positron emission tomography. Neuron *17*, 267–274.

Schacter, D.L., Verfaellie, M., and Pradere, D. (1996b). The neuropsychology of memory illustions: false recall and recognition in amnesic patients. J. Mem. Lang. *35*, 319–334.

Schacter, D.L., Curran, T., Galluccio, L., Milberg, W.P., and Bates, J.F. (1996c). False recognition and the right frontal lobe: a case study. Neuropsychologia *34*, 793–808.

Schacter, D.L., Buckner, R.L., Koutstaal, W., Dale, A.M., and Rosen, B.R. (1997a). Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. Neuroimage *6*, 259–269.

Schacter, D.L., Verfaellie, M., and Anes, M.D. (1997b). Illusory memories in amnesic patients: conceptual and perceptual false recognition. Neuropsychology *11*, 331–342.

Schacter, D.L., Norman, K.A., and Koutstaal, W. (1998a). The cognitive neuroscience of constructive memory. Annu. Rev. Psychol. *49*, 289–318.

Schacter, D.L., Verfaellie, M., Anes, M.D., and Racine, C. (1998b).

When true recognition suppresses false recognition: evidence from amnesic patients. J. Cogn. Neurosci. *10*, 668–679.

Schacter, D.L., Israel, L., and Racine, C. (1999). Suppressing false recognition in younger and older adults: the distinctiveness heuristic. J. Mem. Lang. *40*, 1–24.

Schacter, D.L., Cendan, D.L., Dodson, C.S., and Clifford, E.R. (2001). Retrieval conditions and false recognition: testing the distinctiveness heuristic. Psychon. Bull. Rev. *8*, 827–833.

Schacter, D.L., Dobbins, I.G., and Schnyer, D.M. (2004). Specificity of priming: A cognitive neuroscience perspective. Nat. Rev. Neurosci., in press.

Schnider, A. (2001). Spontaneous confabulation, reality monitoring, and the limbic system—a review. Brain Res. Brain Res. Rev. *36*, 150–160.

Schnider, A., and Ptak, R. (1999). Spontaneous confabulators fail to suppress currently irrelevant memory traces. Nat. Neurosci. *2*, 677–681.

Schooler, J.W., Gerhard, D., and Loftus, E.F. (1986). Qualities of the unreal. J. Exp. Psychol. Learn. Mem. Cogn. *12*, 171–181.

Simons, J.S., Lee, A.C.H., Graham, K.S., Verfaellie, M., Koutstaal, W., Hodges, J.R., Schacter, D.L., and Budson, A.E. (2004). Failing to get the gist: Reduced false recognition of semantic associates in semantic dementia. Neuropsychology, in press.

Slotnick, S.D., and Schacter, D.L. (2004). A sensory signature that distinguishes true from false memories. Nat. Neurosci. *7*, 664–672.

Squire, L.R., and Schacter, D.L. (2002). Neuropsychology of Memory, Third Edition. (New York: Guilford Press).

Swick, D., and Knight, R.T. (1999). Contributions of prefrontal cortex to recognition memory: electrophysiological and behavioral evidence. Neuropsychology *13*, 155–170.

Talland, G.A. (1965). An amnesic patient's disavowal of his own recall performance, and its attribution to the interviewer. Psychiatr. Neurol. (Basel) *149*, 67–76.

Treyer, V., Buck, A., and Schnider, A. (2003). Subcortical loop activation during selection of currently relevant memories. J. Cogn. Neurosci. *15*, 610–618.

Tulving, E., and Schacter, D.L. (1990). Priming and human memory systems. Science *247*, 301–306.

Underwood, B.J. (1965). False recognition produced by implicit verbal responses. J. Exp. Psychol. *70*, 122–129.

Verfaellie, M., Schacter, D.L., and Cook, S.P. (2002). The effect of retrieval instructions on false recognition: exploring the nature of the gist memory impairment in amnesia. Neuropsychologia *40*, 2360–2368.

Verfaellie, M., Rapcsak, S.Z., Keane, M.M., and Alexander, M.P. (2004). Elevated false recognition in patients with frontal lobe damage is neither a general nor a unitary phenomenon. Neuropsychology *18*, 94–103.

von Zerssen, G.C., Mecklinger, A., Opitz, B., and von Cramon, D.Y. (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., and Lang, W. (2000). False recognition in a verbal memory task: an event-related potential study. Brain Res. Cogn. Brain Res. *9*, 41–44.

Ward, J., and Jones, L. (2003). Inappropriate association of semantics and context to novel stimuli can give rise to the false recognition of unfamiliar people. Neuropsychologia *41*, 538–549.

Ward, J., Parkin, A.J., Powell, G., Squires, E.J., Townshend, J., and Bradley, V. (1999). False recognition of unfamiliar people: seeing film stars everywhere. Cogn. Neuropsychol. *16*, 293–315.

Wheeler, M.E., and Buckner, R.L. (2003). Functional dissociation among components of remembering: control, perceived oldness, and content. J. Neurosci. *23*, 3869–3880.

Wheeler, M.E., and Buckner, R.L. (2004). Functional-anatomic correlates of remembering and knowing. Neuroimage *21*, 1337–1349.

Wheeler, M.E., Petersen, S.E., and Buckner, R.L. (2000). Memory's

echo: vivid remembering reactivates sensory-specific cortex. Proc. Natl. Acad. Sci. USA *97*, 11125–11129.

Wiggs, C.L., and Martin, A. (1998). Properties and mechanisms of perceptual priming. Curr. Opin. Neurobiol. *8*, 227–233.

Wilding, E.L. (2000). In what way does the parietal ERP old/new effect index recollection? Int. J. Psychophysiol. *35*, 81–87.

Wilding, E.L., and Rugg, M.D. (1996). An event-related potential study of recognition memory with and without retrieval of source. Brain *119*, 889–905.