Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 13527–13533, November 1996
Colloquium Paper

This paper was presented at a colloquium entitled "Memory: Recording Experience in Cells and Circuits," organized by Patricia S. Goldman-Rakic, held February 17–20, 1996, at the National Academy of Sciences in Irvine, CA.

Illusory memories: A cognitive neuroscience analysis

(memory distortion/false recognition/hippocampal formation/frontal lobes)

DANIEL L. SCHACTER

Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138

ABSTRACT Memory illusions and distortions have long been of interest to psychology researchers studying memory, but neuropsychologists and neuroscientists have paid relatively little attention to them. This article attempts to lay the foundation for a cognitive neuroscience analysis of memory illusions and distortions by reviewing relevant evidence from a patient with a right frontal lobe lesion, patients with amnesia produced by damage to the medial temporal lobes, normal aging, and healthy young volunteers studied with functional neuroimaging techniques. Particular attention is paid to the contrasting roles of prefrontal cortex and medial temporal lobe structures in accurate and illusory remembering. Converging evidence suggests that the study of illusory memories can provide a useful tool for delineating the brain processes and systems involved in constructive aspects of remembering.

Memory is essential to a wide variety of cognitive functions and everyday activities. Because our well-being and even survival may depend on access to reliable information about the past, it is not surprising that memory is often accurate. Nonetheless, memories are not always accurate and, under some conditions, may be grossly distorted. When people misremember past experiences the consequences can be serious, as in cases of mistaken eyewitness identification (1). Memory distortions and illusions have long been of interest to cognitive psychologists, dating to the classic 1932 study by Bartlett (2) on the reconstructive nature of memory (for historical review, see refs. 3 and 4). Three decades later, Neisser (5) put forward similar ideas. His monograph stimulated intensive interest on the part of cognitive psychologists in questions concerning memory distortions, resulting in many striking demonstrations of erroneous remembering in laboratory studies (e.g., refs. 6 and 7). Cognitive studies concerning memory distortion continued through the 1980s (e.g., refs. 8–10) and have grown dramatically during the 1990s, inspired in part by controversies over the accuracy of memories retrieved in psychotherapy (cf. refs. 11–14) and effects of suggestive questioning on the reliability of childrens' recollections (e.g., ref. 15).

In contrast to the intensive focus on memory distortions and illusions by psychologists, cognitive neuroscientists concerned with brain mechanisms of memory have paid relatively little attention to them. There are exceptions, of course, such as empirical and theoretical observations concerning confabulations about past events that are sometimes observed in patients with lesions to the ventromedial frontal lobes and nearby regions in the basal forebrain (16–18). But in contrast to the intensive focus on such issues as working memory (19, 20), brain mechanisms of encoding and retrieval (21–25), implicit versus explicit memory (26–29), emotional memories (e.g., refs. 30–32) and other major topics, the relative absence of cognitive neuroscience research on illusory remembering is notable.

The main thesis of this article is that a cognitive neuroscience analysis of illusory recollections, combining observations of memory disorders produced by neurological dysfunction and studies of normal remembering with recently developed functional neuroimaging techniques, can provide important insights into the constructive nature of memory processes in the brain. A key point of agreement between cognitive and biological theories is that memories do not preserve a literal representation of the world; memories are constructed from fragments of information that are distributed across different brain regions, and depend on influences operating in the present as well as the past (cf. refs. 5, 9, and 33–37). By studying memory distortions and illusions from a cognitive neuroscience perspective, it should be possible to gain useful insights into the neural underpinnings of this constructive process.

In this article, I sketch the outlines of a cognitive neuroscience approach to constructive memory processes by considering recent research conducted in our own and others' laboratories. I begin by discussing evidence from neuroimaging research that highlights a distinction between strategic effort and conscious recollection in memory retrieval, which in turn provides a foundation for understanding brain systems that are relevant to memory distortion. I then use this distinction to examine memory distortions in a case of frontal lobe dysfunction, patients with amnesia produced by damage to the medial temporal lobes, normal aging, and healthy young volunteers. Converging evidence from these studies highlights that illusory memories depend on a dynamic interplay among dissociable component processes that contribute to the constructive nature of remembering.

Recollection and Effort in Memory Retrieval

As noted earlier, both cognitive psychologists and neuroscientists agree that memory retrieval is constructive. What component processes contribute to constructive retrieval? This question is brought into sharp focus by neuroimaging studies of memory. Research on memory processes using positron-emission tomography (PET) and, to a lesser extent, functional MRI has progressed rapidly in recent years (for reviews, see refs. 38 and 39). In both PET and functional MRI studies, neuroanatomical correlates of memory are examined by comparing estimates of regional cerebral blood flow in different experimental conditions. Studies concerning retrieval of recently encountered words, pictures, and other kinds of episodic information have consistently revealed activations in the prefrontal cortex, particularly in right anterior

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: PET, positron-emission tomography.

prefrontal regions (25). These extensive retrieval-related frontal lobe activations were initially surprising, because frontal lobe lesions do not usually produce severe amnesia (40, 41), although they are associated with a variety of memory impairments that will be considered later. An important question concerns what exact role specific frontal regions play in retrieval of episodic memories. Similarly, questions concerning the role of the hippocampal formation have also arisen in neuroimaging research. Although the hippocampus and related medial temporal lobe structures have long been implicated in memory (36, 42), and early PET evidence showed hippocampal activation during retrieval (43), a number of subsequent PET studies of episodic memory retrieval failed to detect hippocampal activations, in contrast to extensive activity in prefrontal cortex and posterior cortical regions (e.g., refs. 21, 25, and 44).

When a blood flow increase is observed in a specific brain region during memory retrieval, it could be attributable either to the mental effort associated with searching memory or to the actual recollection of previously studied information (23, 45, 46). To tease apart these two contributors to retrieval activations, Schacter *et al.* (23) manipulated the manner in which subjects studied target words before scanning. Words in the *high recall condition* were presented four times, and during each presentation subjects engaged in a "deep" or elaborative encoding task (they rated the number of meanings associated with the target word; ref. 47). We reasoned that subjects would successfully recall many of these words on a subsequent test. Words in the low recall condition were presented only once, and subjects engaged in a "shallow" encoding task (rating the number "t-junctions" in the word). We reasoned that subjects would recall few of these words despite expending considerable retrieval effort as they attempted to think back to the study list. Subjects were later given cued recall tests during separate 1-min scans for high recall and low recall words in which the first three letters of target words were presented, and they tried to retrieve appropriate list items. In a separate baseline condition, three letter word beginnings were presented that could not be completed with study list items, and subjects responded with the first word that came to mind.

As expected, behavioral data revealed that subjects remembered many more words in the high recall condition than in the low recall condition. Analysis of blood flow changes in the different conditions revealed a consistent pattern of results. Compared with the baseline condition, the low recall condition was associated with extensive bilateral blood flow increases in the prefrontal cortex, but there were no blood flow increases in the vicinity of the hippocampal formation. In contrast, the high recall minus baseline comparison yielded bilateral flow increases in the hippocampal formation, but no significant increases in prefrontal regions. The high recall minus low recall comparison revealed blood flow increases in the right hippocampus, whereas the low recall minus high recall condition revealed blood flow increases in the left prefrontal cortex (see ref. 23 for discussion of findings in other brain regions).

We also conducted the same experiment with a group of elderly adults, and observed strong evidence of normal hippocampal blood flow increases during the high recall condition in our older subjects (48). In contrast, however, elderly adults exhibited abnormal patterns of blood flow in prefrontal cortex during the low recall condition. Specifically, elderly adults did not show significant anterior prefrontal blood flow increases in the low recall minus baseline comparison. In the low recall minus high recall comparison, the elderly failed to show blood flow increases in left anterior prefrontal regions, in contrast to young subjects; instead, the elderly exhibited posterior frontal activations in the vicinity of Broca's area. These findings are consistent with other data linking altered frontal lobe functioning with age-related memory changes (e.g., refs. 49 and 50), and suggest that elderly adults did not engage in the same kinds of strategic or effortful retrieval processes in the low recall condition that young subjects did.

The results from both young and old suggest that blood flow increases in the hippocampal formation during the stem cued recall test are primarily associated with the actual recollection of previously studied words. It remains to be determined whether and to what extent these observations generalize to other tasks and materials. However, Nyberg *et al.* (51) have recently reported strong positive correlations between retrieval success and blood flow increases in left medial temporal/hippocampal regions on a recognition memory task. Schacter *et al.* (52) reported blood flow increases in the right hippocampal formation during episodic recognition of novel possible objects, which were well-remembered, but not during episodic recognition of novel impossible objects, which were more poorly remembered. Taken together, these studies provide converging evidence that blood flow increases in the hippocampal formation during episodic memory retrieval are associated with some aspect of the conscious recollection of a recent event. Failures to detect hippocampal activations in conditions that yield high levels of remembering may be related to difficulties in imaging this region or to the possibility that the hippocampal formation is engaged to some degree in all experimental conditions, thereby making it difficult to detect blood flow increases in particular conditions (for discussion, see refs. 23, 38, and 39).

The pattern of frontal lobe activations in the Schacter *et al.* (23) study suggest a rather different role for prefrontal regions, and are consistent with the idea that specific regions within the left and right prefrontal cortices, respectively, play different roles in memory retrieval (25, 53). We found that right anterior prefrontal cortex (area 10) was activated in the low recall minus baseline comparison, but not in the low recall minus high recall comparison. This observation suggests that right anterior prefrontal cortex may play an important role in the effortful processes that are involved in switching from lexical retrieval to episodic retrieval. One problem with this suggestion is that we failed to observe significant right prefrontal blood flow increases in the high recall minus baseline comparison. This finding is especially puzzling because Squire *et al.* (43) reported right prefrontal activation in a condition that is quite similar to our high recall condition. However, the high recall minus baseline comparison did indeed yield a notable ($P < 0.01$) trend for a blood flow increase in right anterior prefrontal cortex, although it failed to meet the statistical threshold for significance used in the study ($P < 0.005$).

Further consideration of the findings of Squire *et al.* (43) and of the possible role of right prefrontal cortex in memory retrieval provides some insight into why we observed only a trend for right prefrontal activation in the high recall minus baseline condition. In one condition of their experiment, Squire *et al.* (43) examined blood changes when subjects were instructed to provide the first word that came to mind in response to a three letter cue. Although this condition was designed to examine the implicit form of memory known as priming (54), subjects produced almost as many words from the study list when they were instructed to write down the first word that came to mind as when they were instructed to try to remember study list words. By contrast, in most studies levels of priming are considerably lower than levels of explicit memory. When considered in light of the fact that the subjects in Squire *et al.*'s study were shown short study lists, engaged in deep or semantic encoding, and viewed the study lists twice, the nearly equivalent levels of priming and recall suggest that priming was "contaminated" by some form of explicit memory. Although subjects were probably following instructions and writing down the first word that came to mind, they may have involuntarily remembered the study list words.

Consistent with the preceding ideas and with data described earlier linking the hippocampal formation with conscious recollection, Squire *et al.* (43) observed hippocampal activation during primed stem completion [Schacter *et al.* (23) demonstrated that the priming-related hippocampal blood flow increase could be abolished by eliminating explicit memory contamination]. Importantly, however, Squire *et al.* did not observe right anterior frontal lobe activation in this "involuntary" explicit memory condition. This finding suggests that right prefrontal activation is associated with switching into a voluntary or intentional retrieval mode. Perhaps Schacter *et al.* (23) did not observe a robust right prefrontal activation in the high recall condition (even though intentional memory instructions were given) because after four exposures to the study list, subjects often relied on an automatic, involuntary retrieval process to complete test stems. These suggestions are consistent with the results of two other PET studies of recognition memory in which right anterior prefrontal activity was associated with entering the voluntary or intentional retrieval mode (45, 46). However, in a related study of recognition memory, Rugg *et al.* (M. D. Rugg, P. C. Fletcher, C. D. Frith, R. S. J. Frackowiak, and R. J. Dolan, unpublished work) found that right anterior frontal activations were also associated with increasing numbers of successful retrievals.

The exact nature of the right anterior prefrontal contribution to retrieval is not yet known. However, one promising possibility is that this region is involved in reconstructing the general context of a recent event—focusing in on the target episode and filtering out or inhibiting irrelevant information ([see ref. 56; see M. D. Rugg, P. C. Fletcher, C. D. Frith, R. S. J. Frackowiak, and R. J. Dolan, for a related suggestion]). Thus, activation of the right anterior prefrontal cortex may depend in part on the extent to which a particular task requires subjects to filter irrelevant information and focus on specific episodes. By this view, the right prefrontal region should be activated whenever subjects initiate voluntary or intentional retrieval and, depending on task demands and materials, may also show additional activation associated with successful retrieval (see M. D. Rugg, P. C. Fletcher, C. D. Frith, R. S. J. Frackowiak, and R. J. Dolan, personal communication). In contrast, the left prefrontal region that was activated in the low recall minus high recall comparison (in which there was no sign of activation of the right prefrontal region) by Schacter *et al.* (23) may be involved in such strategic processes as generating candidate responses when it is difficult to recall target items. Future studies will be necessary to explore these and other possibilities. However, the general emphasis on a link between specific prefrontal regions and various components of "effortful" retrieval processes fits well with other studies have implicated specific prefrontal regions in different aspects of short-term or working memory, in which information must be actively maintained across a delay as other processing activities proceed (cf. refs. 57–59).

Despite the fact that much remains to be learned about the generality of the foregoing observations, the contrasting roles of prefrontal cortex and hippocampal formation in episodic memory retrieval can help to conceptualize a variety of memory distortions and illusions that are mentioned below.

False Recognition and the Right Frontal Lobe: Patient BG

A number of neuropsychologists have noted that damage to the ventromedial aspects of the frontal lobes and basal forebrain are often associated with the memory distortion known as confabulation, where patients describe detailed recollections of events that never happened (cf. refs. 16–18). Moscovitch (18) has contended that confabulation arises as a result of impairment to strategic retrieval and monitoring processes that depend on frontal regions.

We have recently studied a 65-year-old man, BG, who suffered an infarction restricted to the right frontal lobe (60). BG does not spontaneously generate extensive confabulations, shows no signs of amnesia, and is generally alert, attentive, and cooperative. He does, however, exhibit a striking pattern of false recognitions that provides useful clues concerning the role of prefrontal regions in illusory memories.

To investigate recognition memory in BG, we showed him to a list of familiar words, and later gave him a yes/no recognition test for old and new words. In addition to asking BG to indicate whether a word had appeared previously on a study list, we also probed the nature of his recollective experience by using the remember/know technique (61, 62). Subjects are instructed to indicate that they "remember" having encountered a word on the study list when they possess a specific recollection of something that they perceived or thought when they studied the word, whereas subjects say they "know" that the word appeared on the study list when it seems familiar, but they do not recollect any specific details about it. Although "remember" and "know" responses are to some extent correlated with high-confidence and low-confidence responses, respectively, the evidence also shows that remember/know judgments are not entirely accounted for by differences in confidence (see ref. 61). Results revealed that even though BG's memory for previously studied words (i.e., hit rate) was relatively normal, he made many more false alarms than did any of the eight control subjects. Moreover, BG claimed to "remember" nearly 40% of new words that had not been on the study list, whereas control subjects made "remember" responses to approximately 5% of new words. Both BG and control subjects provided "know" responses to about 10% of new words.

Inspection of BG's performance revealed that many of the new words he claimed to "remember" were associatively related to words that had appeared in the study list. For example, BG claimed to remember seeing the new word "cellar" on the study list, when in fact he had studied "basement." Cognitive research has shown that normal individuals are more prone to false recognition when new words are associatively related to previously studied words than when they have no associative relation to studied items (63, 64). BG's high rate of false recognitions may indicate that he was unduly influenced by semantic or associative similarity when deciding whether he remembered a particular item.

In a follow-up experiment, some of the lure items on the recognition test were associatively related to a word that had appeared on the study list, and others were unrelated to the study list words. The key result was that BG provided many more "remember" responses to both related and unrelated lure items than did control subjects, although he did so more frequently for related than unrelated lures. Thus, BG exhibited false recollections even when a nonstudied lure item had no relation to a previously studied item.

Another possibility is that BG's tendency to "remember" nonstudied words is attributable to the fact that all of the words in our experiments were familiar to him on the basis of preexperimental knowledge—that is, all of the words were represented in his long-term memory before the experiment. Perhaps BG mistakenly took the preexperimental familiarity of a word as evidence that it appeared on the study list, and hence claimed to "remember" nonstudied words. To test this hypothesis, we exposed BG to novel pseudowords (e.g., "brap" and "spafe") that would not have been familiar to him on the basis of preexperimental knowledge. Results indicated that BG still exhibited the same false recognition phenomenon as in previous experiments, claiming to "remember" nonstudied pseudowords much more often than did control subjects. In follow-up experiments in which we asked BG to write down what he remembers about words that he believes were on the list (65), we found that BG's explanations of his pseudoword false alarms almost always made reference to real words he believed were on the list. Thus, preexperimental familiarity likely plays some role in BG's illusory recollections of pseudowords.

Schacter *et al.* (60) were able to nearly eliminate BG's false recognition responses with a simple manipulation. We showed BG pictures of inanimate objects from various categories (e.g., furniture, articles of clothing). On a subsequent recognition test, some of the nonstudied lure items came from these categories, other lure items came from miscellaneous categories of inanimate objects that were not represented on the study list, and still other lure items were animate objects (i.e., animals). BG claimed to "remember" many of the lure items that were drawn from previously studied categories of inanimate objects. But he almost never claimed to "remember" the lure items that were *not* members of previously studied categories.

BG's pattern of false alarms provides an interesting puzzle. On the one hand, BG exhibits considerable false recognition even for lure words that have no associative relationship to words that appeared on the study list. On the other hand, he does not exhibit false recognition to lures from nonstudied categories. To understand this puzzle, Schacter *et al.* (60) suggested that BG relies excessively on information about the general correspondence between a test item and previously studied words when making a recognition decision. Control subjects typically claimed to "remember" that a word or picture had appeared on a study list only when they retrieved specific information about a particular word or picture. BG, by contrast, relied inappropriately on a match between a test item and general characteristics of the study episode when making his recognition decisions.

In follow-up experiments with BG, Curran *et al.* (65) used signal detection analyses to separate out sensitivity and bias in BG's recognition performance. Not surprisingly, we found that BG consistently used excessively liberal response criteria compared with control subjects. In addition, however, we also found evidence of impaired sensitivity. BG's impaired retention of specific information about individual items on the list may have contributed to his over-reliance on general features of the study episode. Indeed, in one experiment, we increased BG's ability to recollect specific details about presented words via the use of a deep or semantic encoding task. Under these conditions, BG made predominantly "know" false alarms instead of "remember" false alarms, as in previous experiments. These observations suggest when BG has access to "high quality" recollective information about specific items he has actually studied, he can use this information to oppose his usual overreliance on general similarity between study and test items.

While overreliance on general similarity can account for BG's high false alarm rate, it does not explain why he claims to "remember" nonstudied items. As noted earlier, in the recent experiments by Curran *et al.* (65), we analyzed exactly what BG claims to recall when he makes a "remember" false alarm. We found that he tends to provide associations to other words or to events in his life—specific information from an inappropriate context. Given previous evidence implicating the frontal lobes in memory for source or contextual information (66, 67), it seems likely that a source memory deficit contributes to the character of BG's false recollections.

We do not yet know whether and to what extent the processes that are defective in BG are related to the processes subserved by the right anterior prefrontal regions activated in PET studies of intact subjects. For example, BG's propensity for false recognition could be related to the fact that he sustained damage to the right hemisphere, as opposed to the prefrontal cortex in particular. Indeed, our findings concerning BG are consistent with data from split-brain patients indicating that the left hemisphere is more prone to false recognition based on general similarity between study and test items than is the right hemisphere (68, 69). BG's pathological false alarm rate might be attributable, in part, to overreliance on left hemisphere processes because of defective right hemisphere function. However, Parkin *et al.* (70) have reported a patient with left frontal lobe damage who makes excessive numbers of false alarms and resembles patient BG in a number of respects. Thus, it is also possible that BG's high false alarm rate is related to its prefrontal locus, irrespective of laterality.

Despite these caveats, it is interesting to consider BG's deficits in light of the idea suggested earlier that right prefrontal cortex may be involved in setting up contextual representations that guide retrieval by allowing one to focus in on a target episode and filter extraneous activity. Norman and Schacter (56) have argued that BG's deficits are related to problems generating a specific, focused representation of a study episode and filtering out irrelevant information. We suggested that BG is capable of generating only a vague or blurry representation of characteristic features of the target episode against which test items are matched. The result is that items that are generally similar to previously studied ones will be accepted as old because they match BG's blurry contextual representation. BG's lesion may have rendered him incapable of engaging in the kinds of effortful mental processes of contextual reconstruction that are normally supported by right prefrontal regions.

Illusory Recognition in Elderly and Amnesic Subjects

One striking feature of our experiments with BG is that some normal elderly control subjects occasionally exhibited surprisingly high levels of false recognition. In view of the apparent contribution of frontal lobe dysfunction to false recognition in BG and other patients, and in view of the previously mentioned evidence on age-related changes in frontal lobe function, it seems plausible to suppose that older adults would be more susceptible than younger adults to false recognition and other memory illusions. A number of experiments have provided evidence that supports this suggestion (71–73).

Norman and Schacter (74) have recently examined illusory memories in older adults using a method for inducing false recognition recently described by Roediger and McDermott (63). Their procedure is based on an earlier study by Deese (75) in which people were exposed to lists of semantically associated words, such as bitter, taste, chocolate, cake, candy, eat, pie, and others, and were then asked to recall the presented words. Deese (75) found that subjects frequently intruded a nonpresented associate of the targets, such as sweet. Roediger and McDermott (63) replicated this result and also found that on a remember/know recognition test, subjects made extraordinarily large numbers of false alarms (i.e., 70–80%) to nonstudied associates such as sweet. Moreover, subjects' false alarms were frequently accompanied by "remember" responses. Although false recognition is a well-established phenomenon in experimental psychology (64), the magnitude of the effect is usually rather modest, with the false alarm rate typically less than one-third of that observed by Roediger and McDermott (63). Subjects in Roediger and McDermott's experiments were behaving much like patient BG, claiming to possess specific recollections of words that were never presented.

In our experiments, older and younger adults listened to a series of 15-word lists, each composed of strong associates of a nonpresented "critical lure" such as sweet. After each list, subjects were given 1 min in which they either performed arithmetic problems or recalled words from the just presented list. After all lists had been presented, subjects were given a remember/know recognition test for studied words, nonstudied critical lures (e.g., sweet), and nonstudied words that were unrelated to study list items. In addition, half of the older and younger subjects were instructed to write down a brief explanation of what they recollected about an item when they made a remember response, to provide some clues concerning the kind of information that people access when they make false recognition responses.

The experiment yielded three main results. First, on the free recall test, older adults produced fewer study list words than did younger adults, yet produced more intrusions of critical lures. Second, on the recognition test, elderly adults correctly recognized fewer studied words than did younger adults, yet falsely recognized slightly more critical lures than did younger subjects; that is, there was a significant interaction between subject group (young versus old) and item type (studied word versus critical lure). Both older and younger adults frequently claimed to "remember" critical lures; young adults made somewhat fewer "remember" responses to critical lures than to studied words, whereas elderly adults actually made more "remember" responses to critical lures than to studied words. Third, requiring subjects to provide explanations of their false recognition responses had no effect on the magnitude of the effect. Explanations typically consisted of semantic associations to target words (e.g., for the critical lure needle, one subject said "this word came in the same list as thread").

In a second experiment, we followed up on this finding by probing more formally the characteristics of older and younger adults' accurate and illusory recognition responses. To do so, we used a technique developed by Johnson and colleagues (10) in which people are asked to rate various characteristics of their memories on a 7-point scale, ranging from "no memory" of a particular characteristic (a score of 1) to a "vivid memory" (a score of 7). Thus, when subjects indicated that a word had appeared on the list they were asked to rate their memory for the sound of the word, its serial position in the list, associations they made to the word at the time of study, and so forth. The main data concerning recall and recognition replicated the outcome of the previous experiment: subjects frequently falsely recalled and recognized critical lures, and older adults were relatively more susceptible to these memory illusions than were younger adults. Analysis of subjects' ratings showed that false recognition, like veridical recognition, was based largely on retrieval of associative and semantic information. However, whereas accurate and illusory recognition did not differ in terms of rated vividness of associative and semantic information, veridical recognition of studied words was accompanied by higher ratings of memory for the sound of the word and other contextual information than was false recognition. These patterns were observed in both older and younger adults, although they were somewhat less pronounced in the elderly subjects.

Why are elderly adults more susceptible to illusory recognition than younger subjects? In view of previous demonstrations of impaired frontal lobe functioning in older adults, and older adults' documented problems remembering specific information about particular presented items, such as their source, we think it is likely that their false recognition is based on a similar, though much less severe, impairment as that observed in BG—overreliance on general similarity between test items and general characteristics of the study episode. Critical lures in the Roediger and McDermott (63) paradigm are semantically similar to numerous previously studied words. Thus, preserved memory for general or gist information, together with impaired memory for specific information, would generate the pattern of results observed in the elderly. Likewise, if older adults generated a critical lure as an associative response to target words during list presentation, on the subsequent recognition test they may have been less able than younger people to carry out the kinds of source monitoring activities that are necessary to determine whether they had actually heard the word or only thought of it (e.g., refs. 49, 76, and 77). Older adults appear to be unable to engage in the kinds of effortful retrieval or monitoring processes needed to oppose the strong feeling of recollection or familiarity associated with general similarity between critical lures and previously studied words.

We have also examined false recognition of critical lures in patients with organic amnesia (78). The amnesic syndrome typically results from damage to the medial temporal lobes, including the hippocampal formation, or to diencephalic structures such as the mammillary bodies (e.g., refs. 36 and 79). Amnesic patients have great difficulty remembering recent events and new information, despite preserved intelligence, perception, and language. Our experiment included eight amnesic patients with damage primarily in the medial temporal lobes as a result of anoxia or encephalitis; four patients had sustained damage to the diencephalic region as a result of chronic alcohol abuse and associated thiamine deficiencies (Korsakoff syndrome; see ref. 79). Amnesics and 12 matched control subjects heard lists of associates of critical lures; they were later tested for studied words, critical lures, and unrelated words that had not been studied. Not surprisingly, amnesic patients showed much less accurate recognition of previously presented words than did control subjects. More interestingly, amnesic patients were also much less susceptible to false recognition of critical lures than were control subjects; they made significantly fewer "old" responses to critical lures than did controls. These results suggest that false recognition of critical lures depends on some of the same underlying neuroanatomical substrate as does veridical recognition of words that were actually studied—the medial temporal/diencephalic regions that are damaged in amnesia. Subjects must retain the semantic features of the presented words to exhibit false recognition of the critical lure, and amnesic patients are apparently unable to do so (for further discussion of theoretical issues, see ref. 78).

Our previously mentioned PET study of older and younger adults (48) is consistent with this observation. As noted earlier, older adults showed normal patterns of hippocampal activation in the high recall condition despite abnormal patterns of frontal lobe activation in the low recall condition. Thus, if false recognition of critical lures depends in some way on medial temporal/hippocampal activation during retrieval, and is further enhanced by impaired frontal lobe functions, it would make sense that older adults are especially susceptible to such effects.

Neuroanatomical Correlates of Accurate and Veridical Recognition: A PET Study

To examine further the neural substrates of illusory recognition, we (80) conducted a PET study using an adaptation of the Roediger and McDermott (63) paradigm. Before scanning, subjects listened to a long list of words that were grouped into semantic categories that each included 20 associates of a nonpresented critical lure. We then administered yes/no recognition tests in separate 60-sec scans for old words that had appeared on the study list ("true targets"), critical lures that had not appeared but were related to previously presented words ("false targets"), and new words that had not appeared and were not systematically related to previously presented words ("true target controls" and "false target controls," respectively). In a separate passive fixation scan, subjects simply looked at a crosshair for 60 sec.

Compared to the passive fixation condition, both accurate recognition of true targets and illusory recognition of false targets were accompanied by significant blood flow increases in many of the same brain regions, including bilateral anterior/dorsolateral frontal cortex, precuneus, bilateral cerebellum, and the left medial temporal lobe, in the vicinity of the parahippocampal gyrus. This latter observation is intriguing in light of previously mentioned findings that amnesic patients with medial temporal damage exhibited little false recognition of critical lures (78), and that medial temporal blood flow increases are associated with successful recollection (23, 51, 52). However, the evidence linking parahippocampal gyrus activation with episodic recognition (as opposed to visual or lexical processing) was equivocal, so this finding must be viewed cautiously pending further research.

Direct comparison between veridical and illusory recognition yielded virtually no significant findings, suggesting that brain activity during the two forms of recognition is quite similar. Nonetheless, we did observe some suggestive differences. Veridical recognition was accompanied by significantly greater blood flow than illusory recognition in the vicinity of the left supramarginal gyrus and superior temporal gyrus. Previous PET studies have implicated these regions in the processing and storage of phonological information (cf., refs. 81–83). Moreover, studies of brain-damaged patients have linked the supramarginal gyrus with disruptions of phonological analysis and retrieval (84). In light of these observations, we hypothesized that temporoparietal increases associated with veridical recognition may reflect subjects' recognition of having heard or rehearsed the target words at the time of study; no such auditory/phonological information was available for critical lures. Although alternative interpretations are possible, this idea fits with the previously mentioned finding reported by Norman and Schacter (74) that people report more extensive memories of having heard or thought about presented words than critical lures.

One further suggestive finding from the PET study relates to the previously mentioned observations of a link between frontal lobe impairments and heightened susceptibility to false recognition. We found that false recognition was associated with trends for greater blood flow increases in inferior frontal regions (orbitofrontal cortex) bilaterally and right anterior prefrontal cortex than was veridical recognition. One possible interpretation of this finding is that subjects were trying to oppose or inhibit the sense of familiarity or recollection associated with the critical lures. That is, when subjects were deciding whether a critical lure had appeared previously, they likely experienced a strong feeling that it did. At the same time, knowing that many associatively related items were on the list, they may have engaged in effortful retrieval processes as they tried to remember specific information about the test item's appearance in the study list.

Although these findings provide initial clues concerning similarities and differences between veridical and illusory recognition, further studies will be needed to determine the generalizability of the results. It is possible, for instance, that differences in test format, word sets, and other experimental details may produce different patterns of brain activity. At the present, it seems reasonable to conclude that brain activity is largely similar during veridical and illusory recognition, though some evidence for differences can be detected. This description fits reasonably well with findings from purely behavioral experiments (63, 74), indicating many similarities between the properties of veridical and illusory recognition memory, together with a few differences.

Concluding Comments

Our knowledge of the neural systems and processes involved in illusory memories is still meager, and relevant empirical studies are just beginning. Nonetheless, the preliminary findings described in this article suggest that a cognitive neuroscience analysis is likely to provide important new insights into errors and distortions of remembering, which in turn can serve as a useful window on the nature of constructive processes in human memory. By comparing and contrasting different kinds of memory illusions, it should be possible to delineate the component processes involved in each of them. For example, the study by Schacter *et al.* (78) discussed earlier shows that amnesic patients are less susceptible to false recognition of critical lures than are normal controls. By contrast, other recent evidence indicates that amnesics are more susceptible than controls to false recognitions based on illusory memory conjunctions, where people claim to have encountered a stimulus previously when in fact they only saw its component features separately (55, 85). Illusory memory conjunctions appear to be attributable to inadequate binding of features at the time of encoding, a process that likely depends on the hippocampal formation (85). Taken together, these studies highlight the different ways in which medial temporal brain regions may contribute to constructive aspects of remembering. Although concern with constructive processes has long been restricted to investigators taking a purely psychological approach, it seems likely that the analytic tools of modern cognitive neuroscience will enable us to penetrate more deeply into some of the most enigmatic yet revealing aspects of human memory.

I thank Mara Gross for help with preparation of the manuscript. This article was supported by National Institute on Aging Grant AG08441-06 and National Institute of Neurological Disorders and Stroke Grant PO1 NS27950.

1. Loftus, E. F. (1979) *Eyewitness Testimony* (Harvard Univ. Press, Cambridge, MA).
2. Bartlett, F. C. (1932) *Remembering* (Cambridge Univ. Press, Cambridge, U.K.).
3. Roediger, H. L., III (1996) in *Attention and Performance XIV: A Silver Jubilee*, eds. Meyer, D. E. & Kornblum, S. (MIT Press, Cambridge, MA), in press.
4. Schacter, D. L. (1995) in *Memory Distortion: How Minds, Brains and Societies Reconstruct the Past*, eds. Schacter, D. L., Coyle, J. T., Fischbach, G. D., Mesulam, M.-M. & Sullivan, L. E. (Harvard Univ. Press, Cambridge, MA), pp. 1–43.
5. Neisser, U. (1967) *Cognitive Psychology* (Appleton-Century-Crofts, New York).
6. Bransford, J. D. & Franks, J. J. (1971) *Cognit. Psychol.* **2,** 331–350.
7. Loftus, E. F., Miller, D. G. & Burns, H. J. (1978) *J. Exp. Psychol. Hum. Learn. Mem.* **4,** 19–31.
8. Dywan, J. & Bowers, K. S. (1983) *Science* **222,** 1184–1185.
9. Jacoby, L. L., Woloshyn, V. & Kelley, C. M. (1989) *J. Exp. Psychol. Gen.* **118,** 115–125.
10. Johnson, M. K., Foley, M. A., Suengas, A. G. & Raye, C. L. (1988) *J. Exp. Psychol. Gen.* **117,** 371–376.
11. Lindsay, D. S. & Read, J. D. (1994) *Appl. Cognit. Psychol.* **8,** 281–338.
12. Kihlstrom, J. F. (1996) in *Truth and Memory*, eds. Lynn, S. J. & Spanos, N. P. (Guilford, New York), in press.
13. Loftus, E. F. (1993) *Am. Psychol.* **48,** 518–537.
14. Schacter, D. L., Norman, K. A. & Koutstaal, W. (1996) in *False and Recovered Memories*, ed. Conway, M. A. (Oxford Univ. Press, New York), in press.
15. Ceci, S. J. (1995) in *Memory Distortion*, eds. Schacter, D. L., Coyle, J. T., Fischbach, G. D., Mesulam, M.-M. & Sullivan, L. E. (Harvard Univ. Press, Cambridge, MA), pp. 91–128.
16. Dalla Barba, G. (1993) *Cognit. Neuropsychol.* **10,** 1–20.
17. Johnson, M. K. (1991) in *Awareness of Deficit After Brain Injury: Clinical and Theoretical Issues*, eds. Prigatano, G. P. & Schacter, D. L. (Oxford Univ. Press, New York), pp. 176–197.
18. Moscovitch, M. (1995) in *Memory Distortion: How minds, Brains, and Societies Reconstruct the Past*, eds. Schacter, D. L., Coyle, J. T., Fischbach, G. D., Mesulam, M.-M. & Sullivan, L. E. (Harvard Univ. Press, Cambridge, MA), pp. 226–254.
19. Baddeley, A. (1994) in *Memory Systems 1994*, eds. Schacter, D. L. & Tulving, E. (MIT Press, Cambridge, MA), pp. 351–367.
20. Goldman-Rakic, P. S. (1994) in *Motor and Cognitive Functions of the Prefrontal Cortex*, eds. Thierry, A.-M., Glowinski, J., Goldman-Rakic, P. S., Christen, Y. (Springer, Berlin), pp. 112–121.
21. Andreasen, N. C., O'Leary, D. S., Arndt, S., Cizadlo, T., Hurtig, R., Rezai, K., Watkins, G. L., Boles Ponto, L. L. & Hichwa, R. D. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 5111–5115.

22. Grady, C. L., McIntosh, A. R., Horwitz, B., Maisog, J. M., Ungerleider, L. G., Mentis, M. J., Pietrini, P., Schapiro, M. B. & Haxby, J. V. (1995) *Science* **269,** 218–221.
23. Schacter, D. L., Alpert, N. M., Savage, C. R., Rauch, S. L. & Albert, M. S. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 321–325.
24. Tulving, E., Kapur, S., Craik, F. I. M., Moscovitch, M. & Houle, S. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 2016–2020.
25. Tulving, E., Kapur, S., Markowitsch, H. J., Craik, F. I. M., Habib, R. & Houle, S. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 2012–2015.
26. Roediger, H. L., III & McDermott, K. B. (1993) in *Handbook of Neuropsychology*, eds. Spinnler, H. & Boller, F. (Elsevier, Amsterdam), Vol. 8, pp. 63–131.
27. Schacter, D. L. (1987) *J. Exp. Psychol. Learn. Mem. Cognit.* **13,** 501–518.
28. Schacter, D. L. (1994) in *Memory Systems 1994*, eds. Schacter, D. L. & Tulving, E. (MIT Press, Cambridge, MA), pp. 244–256.
29. Squire, L. R. (1994) in *Memory Systems 1994*, eds. Schacter, D. L. & Tulving, E. (MIT Press, Cambridge, MA), pp. 203–231.
30. Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C. & Damasio, A. R. (1995) *Science* **269,** 1115–1118.
31. LeDoux, J. E. (1995) *Annu. Rev. Psychol.* **46,** 209–235.
32. Cahill, L., Prins, B., Weber, M. & McGaugh, J. L. (1994) *Nature (London)* **371,** 702–704.
33. Damasio, A. R. (1989) *Cognition* **33,** 25–62.
34. McClelland, J. L. (1995) in *Memory Distortion: How Minds, Brains and Societies Reconstruct the Past*, eds. Schacter, D. L., Coyle, J. T., Fischbach, G. D., Mesulam, M.-M. & Sullivan, L. E. (Harvard Univ. Press, Cambridge, MA), pp. 69–90.
35. Schacter, D. L. (1996) *Searching for Memory: The Brain, the Mind, and the Past* (Basic Books, New York).
36. Squire, L. R. (1992) *Psycholog. Rev.* **99,** 195–231.
37. Tulving, E. (1983) *Elements of Episodic Memory* (Clarendon, Oxford).
38. Buckner, R. L. & Tulving, E. (1995) in *Handbook of Neuropsychology*, eds. Boller, F. & Grafman, J. (Elsevier, Amsterdam), Vol. 10, pp. 439–466.
39. Ungerleider, L. G. (1995) *Science* **270,** 760–775.
40. Schacter, D. L. (1987) *Psychobiology* **15,** 21–36.
41. Squire, L. R. (1987) *Memory and Brain* (Oxford Univ. Press, New York).
42. Milner, B., Corkin, S. & Teuber, H. L. (1968) *Neuropsychologia* **6,** 215–234.
43. Squire, L. R., Ojemann, J. G., Miezin, F. M., Petersen, S. E., Videen, T. O. & Raichle, M. E. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 1837–1841.
44. Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Frackowiak, R. S. J. & Dolan, R. J. (1994) *Nature (London)* **368,** 633–635.
45. Kapur, S., Craik, F. I. M., Jones, C., Brown, G. H., Houles, S. & Tulving, E. (1995) *NeuroReport* **6,** 1880–1884.
46. Nyberg, L., Tulving, E., Habib, R., Nilsson, L.-G., Kapur, S., Houle, S., Cabeza, R. & McIntosh, A. R. (1995) *NeuroReport* **6,** 249–252.
47. Craik, F. I. M. & Tulving, E. (1975) *J. Exp. Psychol. Gen.* **104,** 268–294.
48. Schacter, D. L., Savage, C. R., Alpert, N. M., Rauch, S. L. & Albert, M. S. (1996) *NeuroReport* **7,** 1165–1169.
49. Craik, F. I. M., Morris, L. W., Morris, R. G. & Loewen, E. R. (1990) *Psychol. Aging* **5,** 148–151.
50. Glisky, E. L., Polster, M. R. & Routhieaux, B. C. (1995) *Neuropsychology* **9,** 229–235.
51. Nyberg, L., McIntosh, A. R., Houle, S., Nilsson, L.-G. & Tulving, E. (1996) *Nature (London)* **380,** 715–717.
52. Schacter, D. L., Reiman, E., Uecker, A., Polster, M. R., Yun, L. S. & Cooper, L. A. (1995) *Nature (London)* **376,** 587–590.
53. Buckner, R. L. & Petersen, S. E. (1996) *Semin. Neurosci.*, in press.
54. Tulving, E. & Schacter, D. L. (1990) *Science* **247,** 301–306.
55. Norman, K. A. & Schacter, D. L. (1996) in *Implicit Memory and Metacognition*, ed. Reder, L. M. (Erlbaum, Hillsdale, NJ), in press.
56. McCarthy, G., Blamire, A. M., Puce, A., Nobre, A. C., Bloch, G., Hyder, F., Goldman-Rakic, P. & Shulman, R. G. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 8690–8694.
57. Smith, E. E., Jonides, J., Koeppe, R. A., Awh, E., Schumache, E. H. & Minoshima, S. (1995) *J. Cognit. Neurosci.* **7,** 337–356.
58. Petrides, M., Alivisatos, B., Meyer, E. & Evans, A. C. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 878–882.
59. Schacter, D. L., Curran, T., Galluccio, L., Milberg, W. & Bates, J. (1996) *Neuropsychologia* **34,** 793–808.
60. Gardiner, J. M. & Java, R. I. (1993) in *Theories of Memory*, eds. Collins, A. F., Gathercole, S. E., Conway, M. A. & Morris, P. E. (Erlbaum, Hove, U.K.), pp. 163–188.
61. Tulving, E. (1985) *Can. Psychol.* **26,** 1–12.
62. Roediger, H. L., III & McDermott, K. B. (1995) *J. Exp. Psychol. Learn. Mem. Cognit.* **21,** 803–814.
63. Underwood, B. J. (1965) *J. Exp. Psychol.* **70,** 122–129.
64. Curran, T., Schacter, D. L., Norman, K. A. & Galluccio, L. (1996) *Neuropsychologia*, in press.
65. Janowsky, J. S., Shimamura, A. P. & Squire, L. R. (1989) *Neuropsychologia* **27,** 1043–1056.
66. Schacter, D. L., Harbluk, J. L. & McLachlan, D. R. (1984) *J. Verbal Learn. Verbal Behav.* **23,** 593–611.
67. Metcalfe, J., Funnell, M. & Gazzaniga, M. S. (1995) *Psychol. Sci.* **6,** 157–164.
68. Phelps, E. & Gazzaniga, M. S. (1992) *Neuropsychologia* **30,** 293–297.
69. Parkin, A. J., Bindschaedler, C., Harsent, L. & Metzler, C. (1996) *Brain Cognit.*, in press.
70. Bartlett, J. C., Strater, L. & Fulton, A. (1991) *Mem. Cognit.* **19,** 177–188.
71. Dywan, J. & Jacoby, L. L. (1990) *Psychol. Aging* **3,** 379–387.
72. Rankin, J. S. & Kausler, D. H. (1979) *J. Gerontol.* **34,** 58–65.
73. Norman, K. A. & Schacter, D. L. (1996) *Mem. Cognit.*, in press.
74. Deese, J. (1959) *J. Exp. Psychol.* **58,** 17–22.
75. Johnson, M. K., Hashtroudi, S. & Lindsay, D. S. (1993) *Psychol. Bull.* **114,** 3–28.
76. Schacter, D. L., Osowiecki, D. M., Kaszniak, A. F., Kihlstrom, J. F. & Valdiserri, M. (1994) *Psychol. Aging* **9,** 81–89.
77. Schacter, D. L., Verfaellie, M. & Pradere, D. (1996) *J. Mem. Language* **35,** 319–334.
78. Parkin, A. J. & Leng, N. R. C. (1993) *Neuropsychology of the Amnesic Syndrome* (Erlbaum, Hillsdale, NJ).
79. Schacter, D. L., Reiman, E., Curran, T., Yun, L. S., Bandy, D., McDermott, K. B. & Roediger, H. L., III (1996) *Neuron* **17,** 267–274.
80. Démonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.-L., Wise, R., Rascol, A. & Frackowiak, R. (1992) *Brain* **115,** 1753–1768.
81. Paulesu, E., Frith, C. D. & Frackowiak, R. S. J. (1993) *Nature (London)* **362,** 342–345.
82. Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M. A. & Raichle, M. E. (1989) *J. Cognit. Neurosci.* **1,** 153–170.
83. Roeltgen, D. P. & Heilman, K. M. (1986) *Appl. Psycholinguistics* **6,** 205–230.
84. Kroll, N. E. A., Knight, R. T., Metcalfe, J., Wolf, E. S. & Tulving, E. (1996) *J. Mem. Language* **35**, 176-196.
85. Reinitz, M. T., Verfaellie, M. & Milberg, W. P. (1996) *J. Mem. Language* **35**, 286–299.