CHAPTER 16

The Cognitive Neuropsychology of False Memories: Theory and Data

Chad S. Dodson
University of Virginia, Charlottesville, VA, USA
and
Daniel L. Schacter
Harvard University, Cambridge, MA, USA

Memory for past events can differ, sometimes in striking ways, from how they were initially experienced (Bartlett, 1932; Kopelman, 1999; Loftus et al., 1978; Neisser, 1967; Roediger, 1996; Schacter, 1995, 2001). Even exceptionally vivid and emotional memories, such as "flashbulb memories", are not immune from memory distortion. Schmolck, Buffalo, and Squire (2000) examined individuals' memory for the events surrounding their learning the verdict of the O.J. Simpson trial, such as where they were, who they were with, what they felt, etc. They tested their participants' memory 3 days after the announcement of the verdict and then either 15 months or 32 months later. Surprisingly, nearly 43% of the individuals tested at the 32 month interval reported memories that contained major distortions, as compared to the account they gave just 3 days after learning the verdict. For instance, one individual reported at the 3 day interval that he learned about the verdict while in a lounge at college. When tested 32 months later, this individual remembered being at home with his sister and father when he learned about the verdict. Ironically, many individuals expressed confidence in the accuracy of their recollection, despite its actual inaccuracy (for similar results, see Bohannon & Symons, 1992; Neisser & Harsch, 1992).

How can a remembered episode differ so dramatically from how that episode was initially experienced? As Roediger & McDermott (2000, p. 149) ask in their review of memory distortion, "Where would the recollection come from, if not from stored traces of actual events?" This chapter will review cognitive and neuropsychological research that illustrates mechanisms that promote or minimize the occurrence of distorted memories. First, we will outline a general framework for understanding the processes that contribute to constructive memory phenomena. We then will consider how false memories are influenced by factors operating primarily at the encoding or retrieval stages of memory.

The Handbook of Memory Disorders. Edited by A.D. Baddeley, M.D. Kopelman and B.A. Wilson

A CONSTRUCTIVE MEMORY FRAMEWORK

Our view of constructive memory, which we refer to as the constructive memory framework (CMF; see also Schacter et al. 1998a), draws on the ideas of several investigators, including Johnson et al. (1993), McClelland et al. (1995), Moscovitch (1994), Norman & Schacter (1996), Reyna & Brainerd (1995) and Squire (1992), among others. This framework focuses on key encoding and retrieval processes that contribute to both accurate and inaccurate memories.

Memories of past events consist of a pattern of features that constitute a record of the processes that were active during the encoding of these events (e.g. Johnson et al., 1993; Johnson & Raye, 1981), e.g. some features in the pattern would represent the output of different sensory processes, whereas others would reflect the output of conceptual processes. This pattern of features is widely distributed across different parts of the brain, such that no single location contains a complete record of the trace or engram of a specific experience (Damasio, 1989; Squire, 1992). In short, the memory representation of an event is this distributed pattern of features. Remembering this event, therefore, involves a process of reactivating the features constituting the desired memory representation. In other words, remembering is a process of pattern completion (McClelland et al., 1995) in which a retrieval cue activates a subset of the features comprising a particular past experience, and activation spreads to the rest of the constituent features of that experience.

There are several problems that must be solved in order for the foregoing memory system to represent past events relatively accurately. First, during the encoding of an event, the features making up the corresponding memory representation must be connected or bound together to form a "coherent" representation (i.e. the feature binding process; see Johnson & Chalfonte, 1994; Moscovitch, 1994; Schacter, 1989). If the feature binding process is incomplete because the individual is distracted, for instance, then he/she may subsequently remember fragments or particular features of the episode but not the entire pattern (Johnson et al., 1993; Schacter et al., 1984; Squire, 1995). Inadequate feature binding can contribute to source memory failure, in which an individual fails to remember the conditions under which an event was acquired, such as remembering a fact but failing to remember the journal that contained this bit of information. A second problem that must be solved at encoding is keeping the bound representations or patterns separate from each other. There must be a process of pattern separation (McClelland et al., 1995), e.g. if an individual regularly has breakfast with a friend, then the memory representations for these different episodes will share many characteristics. The patterns comprising the separate memories of these episodes will overlap. If the patterns overlap extensively with one another, then the person may subsequently only recall the general similarities (Hintzman & Curran, 1994) or gist (Reyna & Brainerd, 1995) common to the many episodes. In other words, if there is a failure in pattern separation, then an individual may fail to recollect distinctive, item-specific information that distinguishes one episode from another.

The memory system must solve similar problems at retrieval in order to reconstruct relatively accurate memories of past events. A common assumption of most memory theorists is that the likelihood of retrieving a past event depends on the degree to which features of the retrieval cue match features of the memory representation (e.g. Bower, 1967; Hintzman, 1988; Johnson & Raye, 1981; Tulving, 1983; Underwood, 1969). One problem is that retrieval cues can potentially match representations other than the one that is desired (Dodson &

Shimamura, 2000; Nystrom & McClelland, 1992). For instance, if given the retrieval cue "having breakfast with a friend", there could be countless memories that contain this characteristic and that would potentially be remembered. To solve this problem, there is often an initial stage of retrieval, referred to as focusing (Norman & Schacter, 1996), in which the rememberer forms a more refined description of the characteristics of the episode to be retrieved (Burgess & Shallice, 1996; Norman & Bobrow, 1979). Poor retrieval focus can lead to recollection of information that does not pertain to the target episode. However, the process of focusing means that how people are oriented to assess their memories can influence what is remembered. This is indeed what is found, e.g. the particular kind of memory test, such as an old–new recognition or a source monitoring test, can affect the probability of making a source misattribution (e.g. Dodson & Johnson, 1993; Lindsay & Johnson, 1989; Marsh & Hicks, 1998; Multhaup, 1995; Zaragoza & Lane, 1994).

Once memorial information has been retrieved, the memory system faces an additional problem, referred to by Johnson (1992) as the source monitoring problem. A decision must be made about whether the activated information is a veridical recollection of a previously experienced event, or whether it is a fantasy or a memory of an earlier imagined event (Johnson & Raye, 1981). This phase of retrieval involves a criterion-setting process: the rememberer needs to consider the diagnostic value of perceptual vividness, semantic detail and other kinds of information for determining the origin of the retrieved pattern (Johnson et al., 1993). It is also during this stage that an individual's expectations, metamemorial beliefs and strategies influence how the remembered information is evaluated (e.g. Dodson & Johnson, 1996; Dodson & Schacter, 2001, 2002, and in press).

Although many brain areas are involved in the preceding memory processes, two brain regions are especially relevant to constructive memory: the medial temporal area, including the hippocampal formation, and the prefrontal cortex. Many researchers view the hippocampus as implementing feature binding and pattern separation (cf. McClelland et al., 1995; Squire & Alvarez, 1995; Treves & Rolls, 1994). According to this view, distributed patterns of activity in the neocortex, constituting the memory representations for different episodes, are linked to sparse neuronal representations in region CA3 of the hippocampus, such that each episode is assigned its own hippocampal "index". To the extent that the hippocampus is able to assign nonoverlapping CA3 representations to different episodes, pattern separation is achieved, which will facilitate remembering distinctive characteristics about particular episodes.

The medial temporal region also contributes to pattern completion at retrieval (e.g. Moscovitch, 1994). According to McClelland et al. (1995; see also Squire, 1992), during retrieval of relatively recent episodes (for which there is still a hippocampal index corresponding to the episode), cues activate the episode's index in region CA3 of the hippocampus, and activation spreads from the index to all the features comprising that episode. Once an episode has been consolidated in the neocortex, however, activation can spread directly between the episode's features, and the hippocampus no longer plays an important role in pattern completion (for a contrasting view on consolidation and the medial temporal lobe, see Nadel et al., 2000). Relevant evidence is provided by Teng & Squire (1999) concerning a patient, E.P., with profound amnesia resulting from bilateral damage to the hippocampus and surrounding structures. E.P. is severely impaired at remembering recent events. Nonetheless, he is able to retrieve remote memories about the spatial layout of the city where he grew up as well as controls who also lived in the same city as E.P. during the same time period and for about the same length of time. Thus, although the medial temporal lobe plays a critical role in the formation of episodic memories, it does not appear to be a necessary part of the retrieval of remote episodic memories, at least spatial memories (cf. Rosenbaum et al., 2000, for a patient with more extensive damage who shows dissociations between different types of remote spatial memories; for neuroimaging evidence indicating hippocampal activation during retrieval of remote memories, see Nadel et al., 2000).

The prefrontal cortex also plays a role in the retrieval of memories (for review, see Shimamura, 2000). Patients with frontal lobe damage have shown difficulty remembering the source of previously learned facts (e.g. Janowsky et al., 1989a), reconstructing the order of a recently studied list of words (e.g. Shimamura et al., 1990), determining the relative recency of items (e.g. Milner, 1971) and so forth. In addition, neuroimaging studies consistently have shown prefrontal activity during episodic retrieval, often in a right anterior frontal region (for reviews, see Buckner, 1996; Nyberg et al., 1996; Tulving et al., 1994). Although the exact nature of the functions indexed by these activations remains open to debate, they appear to tap effortful aspects of retrieval (Schacter et al., 1996a) related to focusing or entering the "retrieval mode" (Lepage et al., 2000; Nyberg et al., 1995), post-retrieval monitoring and criterion setting (Rugg et al., 1996; Schacter et al., 1997) or both (Norman & Schacter, 1996).

In the remainder of the chapter we will consider both cognitive and neuropsychological evidence focusing on false memories, particularly occurrences of false recognition. We will attempt to explain these occurrences of memory in terms of malfunctioning processes operating at encoding and/or retrieval.

MEMORY DISTORTION AND FALSE RECOGNITION

One often investigated kind of memory distortion, known as false recognition, occurs when individuals report incorrectly that they previously encountered a novel event (e.g. Deese, 1959; Underwood, 1965). False recognition responses reflect, to some degree, a tendency to endorse items that are similar to what was studied, e.g. individuals are likely to falsely recognize: (a) objects that resemble many earlier-studied objects (e.g. Franks & Bransford, 1971; Posner & Keele, 1968, 1970); (b) words that are semantically related to earlier presented words (e.g. Underwood, 1965); or (c) sentences that match the implied meaning of studied sentences (e.g. Bransford & Franks, 1971; Johnson et al., 1973). Interestingly, Johnson et al.'s study indicates that individuals occasionally go beyond the information given in the studied sentence (e.g. Bruner, 1986), i.e. one's preexisting knowledge—and hence, expectancies—can introduce distortion. For example, after listening to the sentence, "the spy threw the secret document into the fireplace just in time, since 30 seconds longer would have been too late", participants on a later recognition test were likely to claim that they had heard "the spy had burned the secret document" (Johnson et al., 1973). Furthermore, becoming more knowledgeable about a topic—increasing one's expertise—does not necessarily reduce memory errors. In fact, Arkes & Freedman (1984) show that, compared to novices, experts were more likely to falsely recognize statements that matched the implied meaning of events in a previously read story. These kinds of errors indicate that distorted memories are occasionally a byproduct of comprehension in which a representation is created that is based on both the information provided and general knowledge (however, see Alba & Hasher, 1983; Brewer, 1977).

During the past several years, a paradigm developed by Deese (1959) and modified by Roediger & McDermott (1995; see also Read, 1996) has attracted much attention because it produces very high levels of false memories. The Deese/Roediger–McDermott (DRM) paradigm consists of presenting individuals with a list of words (e.g. one list might consist of the words truck, bus, train, automobile, vehicle, drive, jeep, Ford, race, keys, garage, highway, sedan, van, and taxi) that are all semantically related to a lure word that was not presented in the list (e.g. car). Subsequently, when asked to recall the presented words individuals frequently report having studied the non-presented lure word (e.g. car) (e.g. Deese, 1959; McDermott, 1996; Read, 1996; Roediger & McDermott, 1995; Smith & Hunt, 1998; Toglia et al., 1999). Similarly, if individuals are asked to distinguish between studied words and new words, they often falsely recognize the related lure word. In fact, false recognition rates to the non-presented lure words are so high that they are often indistinguishable from the correct recognition rates to words that were actually studied (Dodson & Schacter, 2001; Mather et al., 1997; Norman & Schacter, 1997; Payne et al., 1996; Roediger & McDermott, 1995; Schacter et al., 1996c; cf. Miller & Wolford, 1999; Roediger & McDermott, 1999; Wixted & Stretch, 2000).

Although there are a variety of factors contributing to this false memory effect, such as how deeply the items were encoded or whether people were distracted during the study phase, one important variable is the number of related items that were studied (e.g. Seamon et al., 1998; Smith & Hunt, 1998; Toglia et al., 1999). Individuals are more likely to falsely recognize a related new word when they have studied many, rather than few, associates of the item (e.g. Arndt & Hirshman, 1998; Robinson & Roediger, 1997; Shiffrin et al., 1995), e.g. Arndt & Hirshman presented participants with lists that contained either 4 or 16 associated words (filler words were included to equate list lengths). On a later memory test, studied words from the small and large lists of associates were recognized at comparable rates (58% for the 4-associate-word lists and 61% for the 16-associate-word lists). By contrast, new words were much more likely to be falsely recognized when they were related to the large list of associates (67%) than to the small list of associates (41%).

There have been three general accounts of this false recognition effect. The first explanation involves the increased activation, or familiarity, of the lure word (e.g. car) that is a consequence of having studying many related words (e.g. truck, bus, train) (see Arndt & Hirshman, 1998, for a fuller discussion). Specifically, the related new word's familiarity is determined by the degree to which it is similar to or matches earlier studied items. Thus, increasing the number of similar words that are studied will increase the activation and the false recognition rate of the related new word. Related to this account, the activation/monitoring framework of Roediger, McDermott and colleagues views the activation of the lure word as occurring during the study phase (e.g. McDermott & Watson, in press; Roediger et al., in press), i.e. studying the related words (e.g. bed, tired, dream, etc.) activates, via a lexical/semantic network, the lure word (e.g. sleep), thus enhancing its likelihood of false recall and false recognition. A second account of this false memory effect, represented by the CMF (and also by Brainerd & Reyna's "fuzzy trace" theory), proposes that it is a byproduct of pattern separation failure (e.g. Schacter et al., 1998a), i.e. when many related words have been studied there may be very high levels of overlap among the corresponding memory representations. This failure to keep representations separate will result in good memory for what the items have in common, but there will be poor memory for the specific details about each item. However, this difficulty in recollecting specific information about the studied items may encourage responding on the basis of the test item's overall familiarity or the degree to which it matches the gist of what was studied (Brainerd & Reyna, 1998; Payne et al., 1996; Schacter et al., 1996c). A third explanation of this false memory effect involves a process of "implicit associative responses" (Underwood, 1965), in which studying many related words (e.g. truck, bus, train, etc.) may lead individuals to produce on their own the new lure word (i.e. car). On a subsequent memory test, individuals may experience source confusion with the lure words and mistakenly believe that they have studied this word, when in fact it was one that they had generated (e.g. Johnson et al., 1993).

These accounts of the false memory effect in the DRM paradigm are not exclusive, of course, and may even interact with each other. For example, according to Roediger, McDermott and colleagues' activation/monitoring framework, studying many associated words activates the related lure word to such a degree that individuals occasionally think of the lure word (Roediger et al., in press). In other words, the activation of the lure word contributes to the occurrence of an implicit associative response. Moreover, if individuals do consciously generate the lure word during the study episode, then as it is rehearsed with the other studied words the lure word may acquire some of the features that are shared by the other studied words. Thus, according to this framework rehearsal processes during the study phase contribute to "remember" responses on the later memory test, whereby people recollect specific features about the lure word, such as the quality of the voice that spoke it (e.g. Payne et al., 1996).

Koutstaal & Schacter (1997) have provided evidence that activation and/or pattern-separation related processes, instead of implicit associative responses, can produce false recognition. Participants in their study were presented with numerous pictures from various categories, such as different kinds of shoes and cars. On a later memory test, participants often falsely recognized new pictures that were similar to studied pictures, such as a new picture of a shoe. Koutstaal & Schacter reasoned that it is unlikely that participants produced the new picture during the study phase as an implicit associative response. Instead, false recognition responses to the new pictures were likely a result of some combination of the first two accounts: (a) associative activation processes, in which related pictures are falsely recognized because they match many studied pictures; and/or (b) pattern separation failure, whereby individuals have difficulty remembering features about specific studied pictures but do remember what the pictures have in common (i.e. the gist) and, therefore, respond accordingly.

The foregoing studies indicate that memory distortion can result from a variety of factors that operate at both encoding and at retrieval. As many have noted, distorted memories are, in part, a by-product of comprehension processes as the mind attempts to integrate what is perceived with what is known (for review and discussion, see Alba & Hasher, 1983; Schacter, 1996). However, as we will discuss later, how events are remembered does not depend solely on the processes occurring during perception and comprehension. For both the source monitoring framework of Johnson and colleagues and the CMF of Schacter and colleagues, retrieval strategies and how individuals are oriented to examine their memories can influence what is remembered about past events (e.g. Dodson & Johnson, 1993, 1996; Hasher & Griffin, 1978; Johnson et al., 1993; Lindsay & Johnson, 1989; Schacter et al., 1998a).

FALSE RECOGNITION AND AMNESIA

Neuropsychologists have recently begun to intensify their efforts to examine memory distortion and the underlying brain systems that give rise to false memories (for reviews, see

Kopelman, 1999; Schacter et al., 1998). In this section we focus on patients with amnesia and the contributions to our understanding of constructive memory that have been provided by studying amnesic patients. Patients with damage to the temporal lobes and related structures in the diencephalon usually exhibit problems in remembering recent events, despite showing relatively intact perceptual, linguistic and general intellectual abilities (e.g. Parkin & Leng, 1993; Squire, 1992).

Schacter et al. (1996c) used the DRM paradigm to investigate memory distortion in amnesics. Amnesics (of varying etiologies) and control subjects studied lists of semantically related words (e.g. bed, tired, dream) and then completed a recognition test containing studied words (e.g. bed), new related words (e.g. sleep), and new unrelated words (e.g. point). As expected, amnesics recognized fewer studied items than did the matched controls. However, they also falsely recognized fewer related lure words than did the controls (for an extension, see Melo et al., 1999). In a follow-up experiment, Schacter et al. (1997) showed that amnesics' reduced false recognition of related lure items extends to perceptual materials. After studying perceptually related words (e.g. fade, fame, face, fake, mate, hate, late, date, and rate) amnesics were less likely than controls both to correctly recognize studied words and to falsely recognize perceptually related lure words (e.g. fate). Thus, for amnesics the same processes that support accurate recognition of studied words also contribute to the false recognition of critical lures that are semantically or perceptually related to earlier studied items (see Koutstaal et al., 1999b, in press, for similar patterns of reduced false recognition in amnesics with different types of pictorial stimuli).

Subsequent experiments using the DRM paradigm have identified conditions in which amnesics show comparable or higher than normal false recognition rates to related lure words. In a study by Schacter et al. (1998b), both amnesic patients (both Korsakoff and amnesics of mixed etiology) and matched controls studied DRM lists of semantically related words and then made old–new recognition judgments about studied words, new related words, and new unrelated words. The key feature of this paradigm was that this study and test procedure was repeated five times, with the same study and test items in all trials. As seen in panels A and C of Figure 16.1, both amnesics and controls correctly recognized increasingly more studied words with repeated study and testing. Panels B and D of Figure 16.1 show the false recognition rates to the related lure words. Across the five study-test trials, control subjects falsely recognized fewer related lure words (see also McDermott, 1996; McKone & Murphy, 2000), whereas the Korsakoff amnesics falsely recognized increasingly more related lure words (Panel B) and the mixed amnesics showed fluctuating levels of false recognition (Panel D). Control subjects presumably encoded more distinct features of the studied words with repeated study and testing, causing their representations to overlap less and less. Moreover, their increasingly better memory for studied words—and the greater pattern separation—heightened the probability that control subjects would notice a difference between the studied words and the related lure word, thus diminishing the false recognition rate across study-test trials.

The Korsakoff amnesics, by contrast, were not able to suppress false recognition responses to the related lures with repeated study and testing. In terms of the CMF, the Korsakoff amnesics formed additional memory representations, and enriched existing ones, with repeated study and testing. However, these representations were not enriched to such an extent that there was a substantial increase in pattern separation. Consequently, there was an increase in false recognition responses across the study-test trials because these amnesics were unable to remember specific information about the studied items to oppose falsely recognizing the

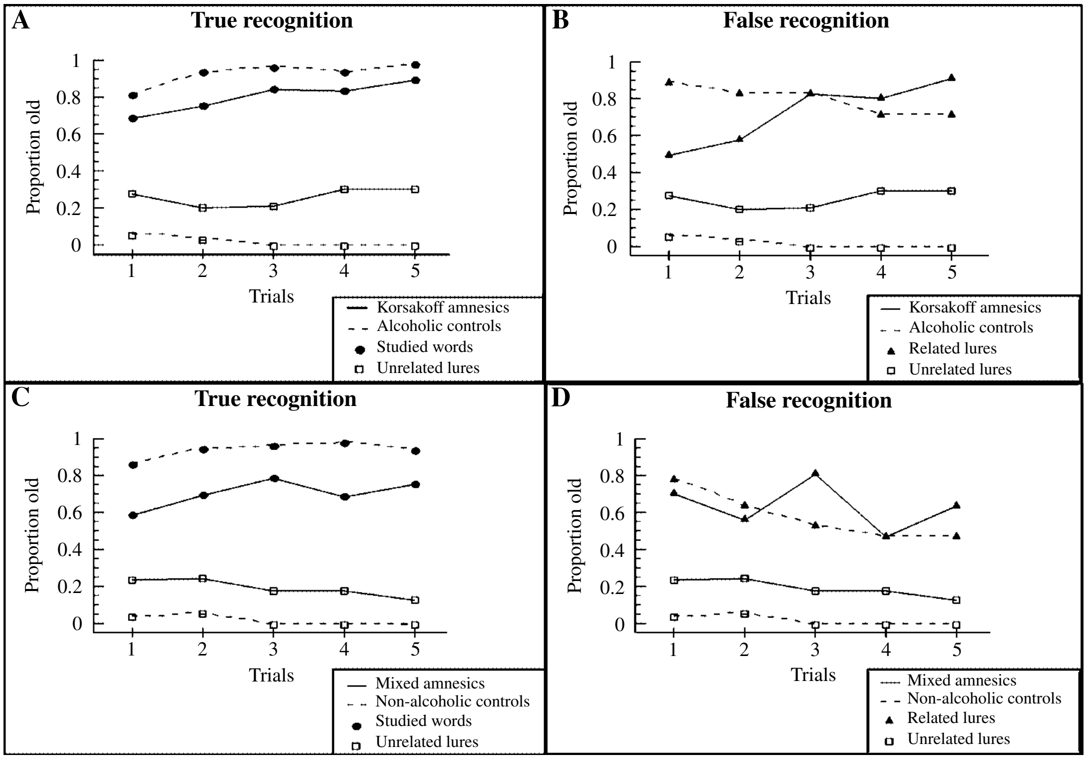

Figure 16.1 Proportions of old responses to studied words (A, C), related lures (B, D) and unrelated lures (A–D) in the two subgroups of amnesics (Korsakoff and mixed) and their respective control groups (alcoholic and nonalcoholic) as a function of study-test trial. Korsakoff and mixed amnesics showed similarly impaired true recognition. However, Korsakoff patients showed increasing false recognition across trials, whereas mixed amnesics showed a fluctuating pattern across trials

lures (for a similar pattern in patients with memory disorders attributable to Alzheimer's disease, see Budson et al., 2000). In sum, these data highlight the importance in healthy individuals of recollecting detailed information for resisting the tendency to falsely recognize similar items.

In contrast to the preceding findings of reduced false recognition by amnesics in the DRM paradigm, when there is a single study and test trial, an earlier study by Cermak et al. (1973) used a different paradigm and showed that amnesics produced *increased* levels of false recognition as compared to control subjects. Cermak et al. presented a series of words to Korsakoff amnesics and alcoholic controls and instructed them to indicate for each word whether or not they had seen it before in the list (i.e. a continuous recognition task). Some words in the series were lures that were semantically or acoustically related to earlier seen words, such as initially seeing "hare" and then later seeing in the list the new lure word "hair". Cermak et al. found that amnesics were more likely than controls to falsely recognize the lure words (see Kroll et al., 1996, who also found that amnesics with either left or right hippocampal damage had higher than normal false recognition rates).

Different processes likely contribute to amnesics' higher than normal false recognition rates in the Cermak et al. (1973) paradigm, but lower than normal false recognition rate in the DRM paradigm. In the Cermak et al. study, the number of items separating a word from its lure (i.e. the lag) may have been sufficiently small that control subjects recollected the initial study word when seeing the lure. For instance, when confronted with the new lure word "hair", the control subjects may have remembered that they had earlier seen "hare" and thus concluded that although "hair" is familiar it does not match the initially studied word. Amnesics, by contrast, have difficulty reactivating earlier studied words (e.g. Johnson & Chalfonte, 1994). This deficiency would have hindered their use of recollection to counter the familiarity of the new lure word, and consequently contributed to their higher than normal false recognition rate in the Cermak et al. paradigm.

In the DRM paradigm, however, the amnesics' lower than normal false recognition rate may have had less to do with recollective deficiencies and more to do with the processes that contribute to a related lure item's feeling of familiarity. That is, the amnesics' low false (and true) recognition rate is likely attributable to the reduced amount of activation (and therefore, familiarity) that is generated by studying the related words. Amnesics' inability to associate the related studied items, and construct an organized representation of the list, may have diminished the activation of the lure items and therefore produced the low false recognition rate.

The preceding studies illustrate two processes that contribute to reports of false memories. First, false recognition responses are influenced by variables, such as how many related items have been studied, that contribute to the familiarity of related lure items. Second, recollecting specific information about past events can counteract false recognition responses. As suggested by the Cermak et al. study, the amnesics' deficiency in remembering item-specific information may have contributed to their higher than normal false recognition rate. Thus, healthy individuals may become more vulnerable to distorted memories as they fail to recollect the discriminative features of past events (e.g. Riccio et al., 1994). In short, memory distortion depends upon a dynamic interaction between processes contributing to familiarity and memory for item-specific information.

THE DISTINCTIVENESS HEURISTIC AND FALSE RECOGNITION

Remembering past events is an active process that is influenced by a variety of factors, such as an individual's expectations, metamemorial beliefs, and even the particular way that memory is queried (e.g. Dodson & Johnson, 1993, 1996; Dodson & Schacter, 2001, 2002; Johnson et al., 1993; Koutstaal et al., 1999a; Lindsay & Johnson, 1989; Marsh & Hicks, 1998; Multhaup, 1995; Schacter et al., 1999; Strack & Bless, 1994). According to both the CMF and the SMF of Johnson and colleagues, individuals can recruit a variety of different decision strategies when making memory judgments. In this section, we will focus on the influence of a retrieval strategy, known as the distinctiveness heuristic, that has been effective at reducing false recognition errors (Dodson & Schacter, 2001, submitted; Israel & Schacter, 1997; Schacter et al., 1999, 2001). The distinctiveness heuristic refers to a mode of responding in which people expect to remember vivid details of an experience and make recognition decisions based on this metacognitive expectation. When a novel event lacks the expected distinctive information, individuals can use this absence of critical evidence to reject the item.

Recently, we have shown that people can use the distinctiveness heuristic to suppress the large false recognition rate of lure words in the DRM false memory paradigm (Dodson &

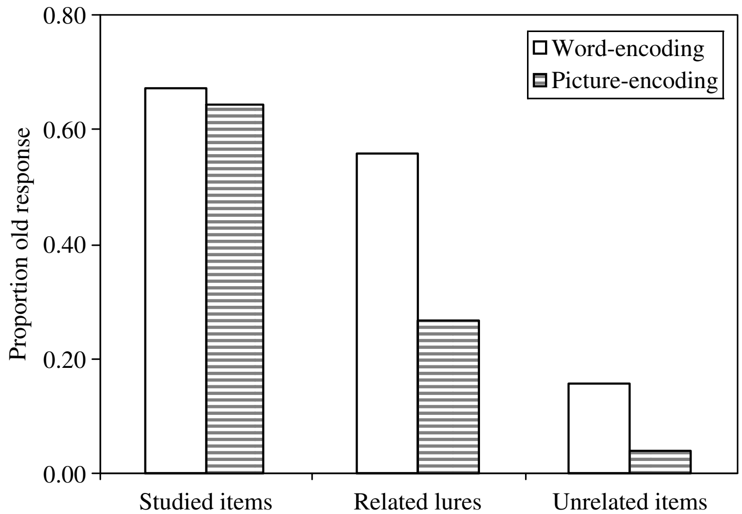

Figure 16.2 Proportions of old responses to studied words, related lures and unrelated lures in the two groups of participants who encoded the words by hearing them either with or without a picture

Schacter, 2001; Israel & Schacter, 1997; Schacter et al., 1999, 2001). Recall that in the DRM paradigm participants initially study lists of words that are associatively related to a nonpresented lure word. On a subsequent old–new recognition test, participants frequently judge the related lure words as having been studied before. However, as seen in Figure 16.2, studying the same items with an accompanying picture, instead of words alone, dramatically reduces false recognition rates to the related lure words (Israel & Schacter, 1997; Schacter et al., 1999, submitted). Schacter et al. argued that the success in rejecting related lure words after picture-plus-word encoding, as compared to word-only encoding, stems from participants' metamemorial belief that they ought to remember the distinctive pictorial information. Participants invoke a distinctiveness heuristic whereby they demand access to pictorial information as a basis for judging items as previously studied; the *absence* of memory for this distinctive information indicates that the test item is new. By contrast, participants who studied only words would not expect detailed recollections about studied items and, thus, would not base recognition decisions on the presence or absence of memory for distinctive information. We (Dodson & Schacter, 2001) reported a similar reduction in false recognition responses to related lures in a condition where participants themselves actually said aloud the target words on the study lists, compared to when they simply heard the target words (participants also saw the studied words in both conditions). Paralleling the reasoning regarding distinctive pictorial information, we suggested that participants who actually said words at study employed a distinctiveness heuristic during the recognition test, demanding access to the distinctive "say" information in order to judge an item as "old". Although the related lure words may feel familiar, the lure words should not evoke the distinctive information associated with actually having said a word out loud, and consequently, can be identified as "new".

Thus far, we have examined the distinctiveness heuristic in a situation in which the studied items share considerable conceptual similarity with one another. However, an important question is whether this retrieval strategy is effective in reducing false recognition rates in situations where the items are unrelated. We (Dodson & Schacter, 2002, in press) examined this question by modifying a "repetition lag" paradigm initially reported by Jennings & Jacoby (1997). In this paradigm, participants study either a list of unrelated words or a list of unrelated pictures, and then make old–new recognition judgments about previously studied items and new words. On the test, the studied items appear once, whereas the new words appear twice, with a varying lag (i.e. number of intervening items) between its first and second occurrence. The new words repeat at lags of either 4, 12, 24, or 48. Participants are instructed to say "old" to studied words only, and to say "new" to non-studied words, even when they repeat. Although participants are explicitly told that if a word occurs twice on the test they can safely conclude that it is a new word, participants in the word-encoding condition nonetheless incorrectly respond "old" to the repeated new words, especially when they repeat at the longer lags. Presumably, individuals mistake the familiarity of the repeated new words—derived from their earlier exposure on the test—for prior presentation in the *study* phase. By contrast, we found that encoding pictures during the study phase, compared to encoding words, reduced the false recognition rate to new words that were repeated during the recognition test. Apparently, participants in the picture-encoding condition, as in the DRM paradigm, used a distinctiveness heuristic during the test, responding "old" only when they could recollect pictorial information about the test items.

The distinctiveness heuristic is one mechanism that has proven useful for minimizing false memories. In terms of the CMF, the distinctiveness heuristic operates on how retrieved information is evaluated. This heuristic depends on the metamemorial knowledge about the kind of events that are likely to be remembered, such as having said something earlier. Based on this knowledge, people are able to infer that the absence of memory for this expected characteristic is diagnostic that an event did not actually occur, despite how familiar it may feel. The distinctiveness heuristic is similar to retrieval strategies that operate in three other paradigms. First, Johnson and colleagues observe that participants use the perceived memorability of studied items to attribute a test item to a particular source, such as inferring that something must have been heard earlier based on the absence of expected memorial information about having generated the item (e.g. Anderson, 1984; Foley et al., 1983; Hashtroudi et al., 1989; Hicks & Marsh, 1999; Johnson et al., 1981). Second, Collins and colleagues argue that a "lack of knowledge" inference contributes to how people answer questions, such as inferring that you have not met Bill Clinton because you would expect to remember this event (e.g. Collins et al., 1975; Gentner & Collins, 1981). Third, and similar to the above, others have argued that especially salient new test items, relative to less salient new items, are correctly rejected because of the expectation of remembering these particularly salient items, i.e. a "I-would-have-seen-it-if-it-had-been-there" strategy (e.g. Brewer and Treyens, 1981; Brown et al., 1977; Strack & Bless, 1994: cf. Rotello, 1999; Wixted, 1992). In sum, our results in combination with existing findings suggest that there is a fundamental metacognitive inference process, based on the absence of memory for expected information, that contributes to retrieval in a variety of situations.

THE FRONTAL LOBES AND FALSE RECOGNITION

Patients with frontal lobe damage can experience a wide range of disorders involving inhibitory processes, language, memory, motor control, personality and general executive functions (e.g. Luria, 1966; Schacter, 1987; Shimamura, 1995, 2000). Several studies document that patients with frontal lobe damage exhibit pathologically high false recognition rates (e.g. Curran et al., 1997; Delbecq-Derouesne et al., 1990; Rapcsak et al., 1999; Schacter et al., 1996b; Schnider & Ptak, 1999; Ward et al., 1999).

Recently, Schnider and colleagues (2000) have observed abnormally high false recognition rates by patients with damage to the orbital frontal cortex or to areas connected to it, such as the basal forebrain, capsular genu, amygdala, perirhinal cortex and hypothalamus. Schnider et al. use a paradigm in which patients and control subjects are presented with a series of pictures, such as pictures of an airplane, bicycle, ball, etc. (Schnider & Ptak, 1999; Schnider et al. 1996a, 1996b, 2000). Some of the pictures repeat in the series, whereas others occur only once. Participants are instructed that the goal of this task is to respond "yes" only to pictures that repeat (i.e. hit) and "no" to pictures that are appearing for the first time. The critical feature of this paradigm is that after viewing the series of pictures, individuals are exposed to the *same* set of pictures again, with a different subset of the pictures repeating, e.g. if airplane and ball had repeated during the first run, then a different set of items would repeat during the second run, such as the picture of the bicycle. By the end of the experiment, individuals have seen multiple runs of the same series of pictures with different subsets of the pictures repeating in a given run.

Schnider & Ptak (1999) compared healthy controls with frontal patients and patients with medial temporal lobe damage. Overall, hit rates—correctly identifying repeating pictures—were not significantly different between the two patient groups, although both patient groups made significantly fewer hits than did control subjects. With repeated exposures to the same series of pictures, the control subjects and the patients with medial temporal lobe damage were no different in correctly rejecting pictures that occurred for the first time within a given run. By contrast, the frontal patients were increasingly vulnerable to falsely recognizing pictures with repeated exposures, e.g. on the final run patients with frontal damage falsely recognized nearly 36% of the pictures, whereas the matched controls and the medial temporal lobe patients falsely recognized less than 5% of the pictures.

The abnormal false recognition rate on the part of the frontal patients may reflect a defective process of remembering source information or specific item information (e.g. Johnson et al., 1993). These patients may have trouble distinguishing between memories for events that occurred earlier in the current run from memories for events that occurred in previous runs. As Schnider & Ptak (1999) report, their patients with frontal damage "fail to suppress mental associations that do not pertain to the present; memories thus seem to be as real and pertinent for present behavior as representations of current reality" (p. 679). Failing to distinguish between items repeating in the current run from items that were seen in earlier runs may leave the frontal patients prone to relying on an item's overall familiarity as a basis for a response, thus contributing to their high false recognition rate.

Swick & Knight (1999) have also observed elevated false recognition rates, compared to controls, in a group of patients with focal unilateral lesions of the lateral prefrontal cortex. In their study, both frontal patients and controls completed a continuous recognition test: participants saw a series of words and pronounceable nonwords and pushed one button for items occurring for the first time (new items) and another button for items that repeat in the series (old items). Whereas the frontal patients were no different from controls in their recognition rates to old items, they were much more likely to falsely recognize new items.

Like the findings of Schnider and colleagues and Swick & Knight, there are a growing number of case studies of individuals with frontal lobe damage who exhibit pathologically high false recognition rates (e.g. Delbecq-Derouesne et al., 1990; Rapcsak et al., 1999;

Schacter et al., 1996b; Ward et al., 1999). Schacter et al. (1996b) and Curran et al. (1997) described a patient, B.G., with right frontal damage who is generally alert, attentive and cooperative. B.G. shows no signs of amnesia but he is prone to extremely high rates of false recognition, e.g. after studying a series of unrelated words, B.G. recognized studied words at a rate comparable to normal control subjects, but B.G. was much more likely than controls to respond that new test words had also been studied (Schacter et al., 1996b). Interestingly, a subsequent experiment identified a condition in which B.G. was able to reduce his high false recognition rate. B.G. and control subjects studied pictures of inanimate objects and then completed a memory test in which the distractors were either related or unrelated to studied items. Some distractors were similar to the studied pictures in that both were from the same semantic category, such as pictures of tools or toys. Other distractors were taken from a semantic category that had not been studied, such as a picture of an animal (Schacter et al., 1996b). As in the prior studies, B.G. recognized the studied items at a rate comparable to controls and showed an abnormally high false recognition rate for related distractors. By contrast, B.G. almost never falsely recognized distractors that were from a different semantic category.

Schacter et al. (1996b) proposed that B.G. uses inappropriate decision criteria during the test. In terms of the CMF, his pathological false recognition rate stems from an excessive reliance on information about the general correspondence between a test item and earlier studied words. Responding on the basis of overall similarity led B.G. to exhibit very high false recognition rates to similar distractors, but not to dissimilar distractors. Alternatively, it is possible that B.G. fails to encode specific features of items at study, so that he is forced to rely on overall similarity. That is, the memory representations of the items do not include enough item-specific information so that at test they can be identified on the basis of this specific information. Moreover, encoding items in a vague manner would result in feelings of familiarity for features that are common to many items, including both studied and new lure items that contain these features (see Curran et al., 1997).

Rapcsak et al. (1999) discuss two other frontal lobe patients that exhibit a pattern of false recognition that is similar to that of B.G. J.S. sustained bilateral damage to the basal forebrain/septal area and the ventromedial frontal region; B.W. is characterized by widespread damage to the right frontal lobe. Both patients show high false recognition rates and, like B.G., are able to lessen their false recognition rates when presented with new items that are substantially different from studied items, e.g. both frontal patients and control subjects recognized previously studied faces at roughly similar rates (94% vs. 81%, respectively) and, as expected, the frontal patients were more likely to falsely recognize the distractor faces than were the control subjects (42% vs. 11%, respectively). Consistent with the behavior of B.G., the frontal patients made many more false alarms to the similar distractor faces (i.e. other white males) than to the dissimilar distractor faces (i.e. white females and non-white males). Apparently, the frontal patients, as Rapscak et al. suggest, use overall familiarity as a basis for a response.

Subsequent experiments by Rapcsak et al. (1999) examined whether the frontal patients' reliance on familiarity stems primarily from deficient encoding of specific item information. Alternatively, there may be a bias at retrieval to rely on familiarity as a basis for a response. J.S., B.W. and healthy controls were given a famous faces test: they were presented a series of famous faces (politicians, entertainers, etc.) intermingled with nonfamous faces and were instructed to respond "yes" when the face was that of a famous person, and "no" when it was not. This test probes memory for information that was learned long before the patient's brain damage and, thus, specifically taps processes operating during retrieval. The frontal patients and control subjects identified comparable numbers of faces that were actually famous (94% vs. 85%, respectively). However, J.S. and B.W. showed a much greater tendency to respond that the nonfamous faces were also famous (52%) than did the controls (6%).

In a follow-up experiment, the same famous faces were intermixed with a new set of unfamiliar faces and all subjects were instructed to base fame judgments solely on whether or not they could remember the name and occupation of the person. With these instructions, the false recognition rates of the nonfamous faces dropped down to normal levels for the frontal patients (i.e. 6% false alarm rate for frontal patients and 1% for controls). Hit rates to the famous faces were also no different for the frontal patients and control subjects. The finding that test instructions greatly improved the frontal patients' performance is consistent with the results of group studies of patients with prefrontal cortex damage (Gershberg & Shimamura, 1995; Hirst & Volpe, 1988).

The preceding studies suggest that, in part, the pathological false recognition rate associated with frontal damage stems from a bias at retrieval to rely on overall familiarity or similarity as a basis for a response. Interestingly, performance improves dramatically when the patients are instructed to respond on the basis of more specific memorial information. In terms of the CMF, a malfunctioning focusing mechanism can explain the behavior of B.G. and the frontal patients of Rapcsak et al. (1999; see Curran et al., 1997; Norman & Schacter, 1996; and Schacter et al., 1996b, for further discussion of B.G.'s memory deficit). That is, the frontal patients may generate a search description that is extremely vague, such as whether or not the test item is a member of one of the studied categories of items. This vague description is sufficient for correctly rejecting distractors that are from nonstudied categories, but it does not exclude similar distractors from studied categories. In the Rapcsak et al. study, the instructions to base fame judgments on the retrieval of specific information about the person may have the effect of focusing or refining the search description. Interestingly, when they are instructed to do so, these frontal patients are capable of focusing their search description and relying on more specific information, since they performed as well as the control subjects in this condition.

One reason for frontal patients' bias to rely on overall familiarity as a basis for a response comes from a study by Janowsky et al. (1989b). Janowsky et al. required frontal patients and matched controls to learn a series of sentences, e.g. "Mary's garden was full of marigolds". On a subsequent cued recall test (e.g. "Mary's garden was full of ____"), all participants gave a "feeling-of-knowing" score for each item that they could not recall, e.g. if an individual failed to recall "marigolds" to the above cue then he/she would predict how likely he/she would be able to recognize the answer on a multiple-choice test—the feeling-of-knowing judgment. Recall and recognition performance was comparable between the two groups, but the frontal lobe patients were at chance levels in their feeling-of-knowing judgments, whereas the control subjects achieved much higher scores. Interestingly, the frontal patients attained poor feeling-of-knowing scores because they tended to overestimate their knowledge. Specifically, the frontal patients only recognized 27% of the unrecalled items that they were moderately or highly confident of subsequently recognizing, whereas the control subjects recognized 42% of these items. This tendency on the part of frontal patients to overestimate their knowledge could contribute to abnormally high false recognition rates. Their apparent overconfidence that an item's familiarity means that it must have been studied earlier may have lead them *not* to use a more focused search description (e.g. search memory for more specific item information) and rely on a test item's familiarity.

CONCLUSION

A basic strategy in cognitive neuropsychology and cognitive neuroscience is to examine how mental processes malfunction in order to increase understanding of how they work. This strategy has proved quite useful in uncovering basic memory mechanisms of encoding, storage and retrieval, and more recently has begun to yield insights into various kinds of memory errors (e.g. Schacter, 1999, 2001). We have examined memory distortion in various kinds of brain-damaged patients from the perspective of the CMF. The reviewed studies indicate that healthy individuals are susceptible to falsely recognizing items when they fail to recollect detailed item information. This may occur when there is a pattern separation failure and the representations of similar studied items overlap, such as in the DRM paradigm. The medial temporal lobes appear to be important for storing and/or retaining both familiarity information and more specific information about an item. False recognition also can be a byproduct of the retrieval process—e.g. when patients (or healthy controls) use lax criteria to search memory, such as when they accept memories as true that are only vaguely familiar. Conversely, a retrieval strategy such as the distinctiveness heuristic can be a powerful mechanism for reducing false recognition responses. Some of the evidence we considered indicates that the frontal lobes are implicated in the criteria and strategies people use to search and evaluate their memories. In short, a variety of brain mechanisms are likely responsible for the occurrence of both true and false memories.

ACKNOWLEDGEMENT

We gratefully acknowledge the support provided by grants from the National Institute on Aging (AG08441), National Institute of Mental Health (NS60941) and National Institute of Neurological Disorders and Stroke (NS26985).

REFERENCES

Alba, J.W. & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, **93**, 203–231.

Anderson, R.E. (1984). Did I do it or did I only imagine doing it? *Journal of Experimental Psychology: General*, **113**, 594–613.

Arkes, H.R. & Freedman, M.R. (1984). A demonstration of the costs and benefits of expertise in recognition memory. *Memory & Cognition*, **12**, 84–89.

Arndt, J. & Hirshman, E. (1998). True and false recognition in MINERVA2: explanations from a global matching perspective. *Journal of Memory and Language*, **39**, 371–391.

Bartlett, F.C. (1932). *Remembering*. Cambridge: Cambridge University Press.

Bohannon, J.N. & Symons, L.V. (1992). Flashbulb memories: confidence, consistency, and quantity. In E. Winograd & U. Neisser (eds), *Affect and Accuracy in Recall: Studies of "Flashbulb" Memories* (pp. 65–91). Cambridge: Cambridge University Press.

Bower, G.H. (1967). A multi-component theory of the memory trace. In K.W. Spence & J.T. Spence (eds), *Coding Processes in Human Memory* (pp. 85–123). Washington DC: V.H. Winston.

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

Bransford, J.D. & Franks, J.J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, **2**, 331–350.

Brewer, W.F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition*, **5**, 673–678.

Brewer, W.F. & Treyens, J.C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, **13**, 207–230.

Brown, J., Lewis, V.J. & Monk, A.F. (1977). Memorability, word frequency and negative recognition. *Quarterly Journal of Experimental Psychology*, **29**, 461–473.

Bruner, J. (1986). *Actual Minds, Possible Words*. Cambridge, MA: Harvard University Press.

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

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

Buckner, R.L. (1996). Beyond HERA: contributions of specific prefrontal brain areas to long-term memory retrieval. *Psychonomic Bulletin and Review*, **3**, 149–158.

Cermak, L.S., Butters, N. & Gerrein, J. (1973). The extent of the verbal encoding ability of Korsakoff patients. *Neuropsychologia*, **11**, 85–94.

Collins, A., Warnock, E.H., Aiello, N. & Miller, M.L. (1975). Reasoning from incomplete knowledge. In D. Bobrow & A. Collins (eds), *Representation & understanding: Studies in cognitive science*. New York: Academic Press.

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

Damasio, A.R. (1989). Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. *Cognition*, **33**, 25–62.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58**, 17–22.

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

Dodson, C.S. & Johnson, M.K. (1993). Rate of false source attributions depends on how questions are asked. *American Journal of Psychology*, **106**, 541–557.

Dodson, C.S. & Johnson, M.K. (1996). Some problems with the process dissociation approach to memory. *Journal of Experimental Psychology: General*, **125**, 181–194.

Dodson, C.S. & Schacter, D.L. (2001). "If I had said it I would have remembered it": reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin and Review*, **8**, 155–161.

Dodson, C.S. & Schacter, D.L. (2002). When false recognition meets metacognition: the distinctiveness heuristic. *Journal of Memory and Language*, **46**, 782–803.

Dodson, C.S. & Schacter, D.L. (in press). Aging and strategic retrieval processes: the suppresssion of false memories with a distinctiveness heuristic. *Psychology and Aging*.

Dodson, C.S. & Shimamura, A.P. (2000). Differential effects of cue dependency on item and source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **26**, 1023–1044.

Foley, M.A., Johnson, M.K. & Raye, C.L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Development*, **54**, 51–60.

Franks, J.J. & Bransford, J.D. (1971). Abstraction of visual patterns. *Journal of Experimental Psychology*, **90**, 65–74.

Gentner, D. & Collins, A. (1981). Studies of inference from lack of knowledge. *Memory & Cognition*, **9**, 434–443.

Gershberg, F.B. & Shimamura, A.P. (1996). The role of the frontal lobes in the use of organizational strategies in free recall. *Neuropsychologia*, **13**, 1305–1333.

Hasher, L. & Griffin, M. (1978). Reconstructive and reproductive processes in memory. *Journal of Experimental Psychology: Human Learning and Memory*, **8**, 318–330.

Hashtroudi, S., Johnson, M.K. & Chrosniak, L.D. (1989). Aging and source monitoring. *Psychology and Aging*, **4**, 106–112.

Hicks, J.L. & Marsh, R.L. (1999). Attempts to reduce the incidence of false recall with source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **25**, 1195–1209.

Hintzman, D.L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, **95**, 528–551.

Hintzman, D.L. & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: evidence for separate processes of familiarity and recall. *Journal of Memory and Language*, **33**, 1–18.

Hirst, W. & Volpe, B.T. (1988). Memory strategies and brain damage. *Brain and Cognition*, **8**, 379–408.

Israel, L. & Schacter, D.L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychological Bulletin and Review*, **4**, 577–581.

Janowsky, J.S., Shimamura, A.P. & Squire, L.R. (1989a). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia*, **27**, 1043–1056.

Janowsky, J.S., Shimamura, A.P. & Squire, L.R. (1989b). Memory and metamemory: comparisons between patients with frontal lobe lesions and amnesic patients. *Psychobiology*, **17**, 3–11.

Jennings, J.M. & Jacoby, L.L. (1997). An opposition procedure for detecting age-related deficits in recollection: telling effects of repetition. *Psychology and Aging*, **12**, 352–361.

Johnson, M.K. (1992). MEM: mechanisms of recollection. *Journal of Cognitive Neuroscience*, **4**, 268–280.

Johnson, M.K., Bransford, J.D. & Solomon, S.K. (1973). Memory for tacit implications of sentences. *Journal of Experimental Psychology*, **98**, 203–205.

Johnson, M.K. & Chalfonte, B.L. (1994). Binding complex memories: the role of reactivation and the hippocampus. In D.L. Schacter & E. Tulving (eds), *Memory Systems 1994* (pp. 311–350). Cambridge, MA: MIT Press.

Johnson, M.K., Hashtroudi, S. & Lindsay, D.S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3–28.

Johnson, M.K. & Raye, C.L. (1981). Reality Monitoring. *Psychological Review*, **88**, 67–85.

Johnson, M.K., Raye, C.L., Foley, H.J. & Foley, M.A. (1981). Cognitive operations and decision bias in reality monitoring. *American Journal of Psychology*, **94**, 37–64.

Kopelman, M.D. (1999). Varieties of false memory. *Cognitive Neuropsychology*, **16**, 197–214.

Koutstaal, W., & Schacter, D.L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language*, **37**, 555–583.

Koutstaal, W., Schacter, D.L., Galluccio, L. & Stofer, K.A. (1999a). Reducing gist-based false recognition in older adults: encoding and retrieval manipulations. *Psychology and Aging*, **14**, 220–237.

Koutstaal, W., Schacter, D.L., Verfaellie, M., Brenner, C. & Jackson, E.M. (1999b). Perceptually-based false recognition of novel objects in amnesia: effects of category size and similarity to category prototypes. *Cognitive Neuropsychology*, **16**, 317–342.

Koutstaal, W., Verfaellie, M. & Schacter, D.L. (in press). Recognizing identical vs. similar categorically related common objects: further evidence for degraded gist-representations in amnesia. *Neuropsychology*.

Kroll, N.E.A., Knight, R.T., Metcalfe, J. et al. (1996). Cohesion failure as a source of memory illusions. *Journal of Memory and Language*, **35**, 176–196.

Lepage M., Ghaffar, O., Nyberg, L. & Tulving, E. (2000). Prefrontal cortex and episodic memory retrieval mode. *Proceedings of the National Academy of Sciences of the USA*, **97**, 506–511.

Lindsay, D.S. & Johnson, M.K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, **17**, 349–358.

Loftus, E.F., Miller, D.G. & Burns, H.J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, **4**, 19–31.

Luria, A.R. (1966). *Higher Cortical Functions in Man*. New York: Basic Books.

Marsh, R.L. & Hicks, J.L. (1998). Test formats change source-monitoring decision processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24**, 1137–1151.

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

McClelland, J.L., McNaughton, B.L. & O'Reilly, R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, **102**, 419–457.

McDermott, K.B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language*, **35**, 212–230.

McDermott, K.B. & Watson, J.M. (in press). The rise and fall of false recall: the impact of presentation duration. *Journal of Memory and Language*.
McKone, E. & Murphy, B. (2000). Implicit false memory: effects of modality and multiple study presentations on long-lived semantic priming. *Journal of Memory and Language*, **43**, 89–109.
Melo, B., Winocur, G. & 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. *Cognitive Neuropsychology*, **16**, 343–359.
Miller, M.B. & Wolford, G.L. (1999). Theoretical Commentary: the role of criterion shift in false memory. *Psychological Review*, **106**, 398–405.
Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin*, **27**, 272–277.
Moscovitch, M. (1994). Memory and working-with-memory: evaluation of a component process model and comparisons with other models. In D.L. Schacter & E. Tulving (eds), *Memory Systems 1994* (pp. 311–350). Cambridge, MA: MIT Press.
Multhaup, K.S. (1995). Aging, source, and decision criteria: when false fame errors do and do not occur. *Psychology and Aging*, **10**, 492–497.
Nadel, L, Samsonovich, A, Ryan. L. & Moscovitch M. (2000). Multiple trace theory of human memory: Computational, neuroimaging, and neuropsychological results. *Hippocampus*, **10**, 352–368.
Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.
Neisser, U. & Harsch, N. (1992). Phantom flashbulbs: false recollections of hearing the news about Challenger. In E. Winograd & U. Neisser (eds), *Affect and Accuracy in Recall: Studies of "Flashbulb" Memories* (pp. 9–31). Cambridge: Cambridge University Press.
Norman, D.A. & Bobrow, D.G. (1979). Descriptions: an intermediate stage in memory retrieval. *Cognitive Psychology*, **11**, 107–123.
Norman, K.A. & Schacter, D.L. (1996). Implicit memory, explicit memory, and false recollection: a cognitive neuroscience perspective. In L.M. Reder (ed.), *Implicit Memory and Metacognition* (pp. 229–259). Hillsdale, NJ: Erlbaum.
Norman, K.A. & Schacter, D.L. (1997). False recognition in young and older adults: exploring the characteristics of illusory memories. *Memory and Cognition*, **25**, 838–848.
Nyberg, L., Cabeza, R. & Tulving, E. (1996). PET studies of encoding and retrieval: the HERA model. *Psychonomic Bulletin and Review*, **3**, 135–148.
Nyberg, L., Tulving, E., Habib, R. et al. (1995). Functional brain maps of retrieval mode and recovery of episodic information. *NeuroReport*, **6**, 249–252.
Nystrom, L.E. & McClelland, J.L. (1992). Trace synthesis in cued recall. *Journal of Memory and Language*, **31**, 591–614.
Parkin, A.J., Ward, J., Bindschaedler, C. et al. (1999). False recognition following frontal lobe damage: the role of encoding factors. *Cognitive Neuropsychology*, **16**.
Parkin, A.J. & Leng, N.R.C. (1993). *Neuropsychology of the Amnesic Syndrome*. Hillsdale, NJ: Erlbaum.
Payne, D.G., Elie, C.J., Blackwell, J.M. & Neuschatz, J.S. (1996). Memory illusions: recalling, recognizing, and recollecting events that never occurred. *Journal of Memory and Language*, **35**, 261–285.
Posner, M.I. & Keele, S.W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, **77**, 353–363.
Posner, M.I. & Keele, S.W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, **83**, 304–308.
Rapcsak, S.Z., Reminger, S.L., Glisky, E.L. et al. (1999). Neuropsychological mechanisms of false facial recognition following frontal lobe damage. *Cognitive Neuropsychology*, **16**.
Read, J.D. (1996). From a passing thought to a false memory in 2 minutes: confusing real and illusory events. *Psychonomic Bulletin and Review*, **3**, 105–111.
Reyna, V.F. & Brainerd, C.J. (1995). Fuzzy-trace theory: an interim synthesis. *Learning and Individual Differences*, **7**, 1–75.
Riccio, D.C., Rabinowitz, V.C. & Axelrod, S. (1994). *Memory: when less is more. American Psychologist*, **49**, 917–926.

Roediger, H.L. III (1996). Memory illusions. *Journal of Memory and Language*, **35**, 76–100.
Roediger, H.L. III & McDermott, K.B. (1995). Creating false memories: remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 803–814.
Roediger, H.L. III & McDermott, K.B. (1999). False alarms about false memories. *Psychological Review*, **106**, 406–410.
Roediger, H.L. III & McDermott, K.B. (2000). Distortions of memory. In E. Tulving & F.I.M. Craik (eds), *The Oxford Handbook of Memory* (pp. 149–162). New York: Oxford University Press.
Roediger, H.L. III, Watson, J.M., McDermott, K.B. & Gallo, D.A. (in press). Factors that create false recall. *Psychonomic Bulletin & Review*.
Robinson, K.J. & Roediger, H.L. III. (1997). Associative processes in false recall and false recognition. *Psychological Science*, **8**, 231–237.
Rosenbaum, R.S., Priselac, S., Kohler, S. et al. (2000). Remote spatial memory in an amnesic person with extensive bilateral hippocampal lesions. *Nature Neuroscience*, **3**, 1044–1048.
Rotello, C.M. (1999). Metacognition and memory for nonoccurrence. *Memory*, **7**, 43–63.
Rugg, M.D., Fletcher, P.C., Frith, C.D. et al. (1996). Differential response of the prefrontal cortex in successful and unsuccessful memory retrieval. *Brain*, **119**, 2073–2083.
Schacter, D.L. (1987). Memory, amnesia and frontal lobe dysfunction. *Psychobiology*, **15**, 21–36.
Schacter, D.L. (1989). On the relation between memory and consciousness. In H.L. Roediger III & F.I.M. Craik (eds), *Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving* (pp. 355–389). Hillsdale, NJ: Erlbaum.
Schacter, D.L. (1995). Memory distortion: history and current status. In D.L. Schacter (ed.), *Memory Distortion: How Minds, Brains and Societies Reconstruct the Past* (pp. 1–43). Cambridge, MA: Harvard University Press.
Schacter, D.L. (1996). *Searching for Memory: The Brain, the Mind, and the Past*. New York: Basic Books.
Schacter, D.L. (1999). The seven sins of memory: insights from psychology and cognitive neuroscience. *American Psychologist*, **54**, 182–203.
Schacter, D.L. (2001). *The Seven Sins of Memory: How the Mind Forgets and Remembers*. Boston, MA: Houghton Mifflin.
Schacter, D.L., Alpert, N.M., Savage, C.R. et al. (1996a). Conscious recollection and the human hippocampal formation: evidence from positron emission tomography. *Proceedings of the National Academy of Sciences of the USA*, **93**, 321–325.
Schacter, D.L., Buckner, R.L., Koutstaal, W. et al. (1997). Late onset of anterior prefrontal activity during true and false recognition: an event-related FMRI study. *NeuroImage*, **6**, 259–269.
Schacter, D.L., Cendan, D.L., Dodson, C.S. & Clifford, E.R. (2001). Retrieval conditions and false recognition: testing the distinctiveness heuristic. *Psychonomic Bulletin & Review*, **8**, 824–833.
Schacter, D.L., Curran, T., Galluccio, L. et al. (1996b). False recognition and the right frontal lobe: a case study. *Neuropsychologia*, **34**, 793–808.
Schacter, D.L., Harbluk, J.L. & McLachlan, D.R. (1984). Retrieval without recollection: an experimental analysis of source amnesia. *Journal of Verbal Learning and Verbal Behavior*, **23**, 593–611.
Schacter, D.L., Israel, L. & Racine, C.A. (1999). Suppressing false recognition in younger and older adults: the distinctiveness heuristic. *Journal of Memory and Language*, **40**, 1–24.
Schacter, D.L., Norman, K.A. & Koutstaal, W. (1998a). The cognitive neurosciences of constructive memory. *Annual Review of Psychology*, **49**, 289–318.
Schacter, D.L., Verfaellie, M. & Anes, M.D. (1997). Illusory memories in amnesic patients: conceptual and perceptual false recognition. *Neuropsychology*, **11**, 331–342.
Schacter, D.L., Verfaellie, M., Anes, M.D. & Racine, C. (1998b). When true recognition suppresses false recognition: evidence from amnesic patients. *Journal of Cognitive Neuroscience*, **10**, 668–679.
Schacter, D.L., Verfaellie, M. & Pradere, D. (1996c). The neuropsychology of memory illusions: false recall and recognition in amnesic patients. *Journal of Memory and Language*, **35**, 319–334.
Schmolck, H., Buffalo, E.A. & Squire, L.R. (2000). Memory distortions develop over time: recollections of the O.J. Simpson trial verdict after 15 and 32 months. *Psychological Science*, **11**, 39–45.

Schnider, A. & Ptak, R. (1999). Spontaneous confabulators fail to suppress currently irrelevant memory traces. *Nature Neuroscience*, **2**, 677–681.

Schnider, A., Ptak, R., von Daniken, C. & Remonda, L. (2000). Recovery from spontaneous confabulations parallels recovery of temporal confusion in memory. *Neurology*, **55**, 74–83.

Schnider, A., von Daniken, C. & Gutbrod, K. (1996a). The mechanisms of spontaneous and provoked confabulations. *Brain*, **119**, 1365–1375.

Schnider, A., von Daniken, C. & Gutbrod, K. (1996b). Disorientation in amnesia: a confusion of memory traces. *Brain*, **119**, 1627–1632.

Seamon, J.G., Luo, C.R. & Gallo, D.A. (1998). Creating false memories with or without recognition of list items: evidence for nonconscious processes. *Psychological Science*, **9**, 20–26.

Shiffrin, R.M., Huber, D.E. & Marinelli, K. (1995). Effects of category length and strength on familiarity in recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **21**, 267–287.

Shimamura, A.P. (1995). Memory and frontal lobe function. In M.S. Gazzaniga (ed.), *The Cognitive Neurosciences* (pp. 803–813). Cambridge, MA: MIT Press.

Shimamura, A.P. (2000). The role of the prefrontal cortex in dynamic filtering. *Psychobiology*, **28**, 207–218.

Shimamura, A.P., Janowsky, J.S. & Squire, L.R. (1990). Memory for the temporal order of events in patients with frontal lobe lesions and amnesic patients. *Neuropsychologia*, **28**, 803–813.

Smith, R.E. & Hunt, R.R. (1998). Presentation modality affects false memory. *Psychonomic Bulletin & Review*, **5**, 710–715.

Squire, L.R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological Review*, **99**, 195–231.

Squire, L.R. (1995). Biological foundations of accuracy and inaccuracy in memory. In D.L. Schacter, J.T. Coyle, G.D. Fischback et al. (eds), *Memory Distortion: How Minds, Brains and Societies Reconstruct the Past*. Cambridge, MA: Harvard University Press.

Squire, L.R. & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: a neurobiological perspective. *Current Opinions in Neurobiology*, **5**, 169–177.

Strack, F. & Bless, H. (1994). Memory for nonoccurrences: metacognitive and presuppositional strategies. *Journal of Memory and Language*, **33**, 203–217.

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

Teng, E. & Squire, L.R. (1999). Memory for places learned long ago is intact after hippocampal damage. *Nature*, **400**, 675–677.

Toglia, M.P., Neuschatz, J.S. & Goodwin, K.A. (1999). Recall accuracy and illusory memories: when more is less. *Memory*, 7, 233–256.

Treves, A. & Rolls, E.T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus*, **4**, 374–391.

Tulving, E. (1983). *Elements of Episodic Memory*. Oxford: Clarendon.

Tulving, E., Markowitsch, H.J., Kapur, S. et al. (1994). Novelty encoding networks in the human brain: positron emission tomography data. *NeuroReport*, **5**, 2525–2528.

Underwood, B.J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, **70**, 122–129.

Underwood, B.J. (1969). Attributes of memory. *Psychological Review*, **76**, 559–573.

Ward, J., Parkin, A.J., Powell, G. et al. (1999). False recognition of unfamiliar people: "seeing film stars everywhere." *Cognitive Neuropsychology*, **16**.

Wixted, J.T. (1992). Subjective memorability and the mirror effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **18**, 681–690.

Wixted, J.T. & Stretch, V. (2000). The case against a criterion-shift account of false memory. *Psychological Review*, **107**, 368–376.

Zaragoza, M.S. & Lane, S.M. (1994). Source misattributions and the suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20**, 934–945.