Memory illusions in amnesic patients: Findings and implications

Chapter · January 2002

Mieke Verfaellie
VA Boston Healthcare System/ Boston University School of Medicine

Daniel L Schacter
Harvard University

9

Memory Illusions in Amnesic Patients

Findings and Implications

DANIEL L. SCHACTER, MIEKE VERFAELLIE, and WILMA KOUTSTAAL

Studies of patients with organic amnesia have contributed enormously to understanding the neuropsychology of memory. Scoville and Milner's (1957) pioneering observations concerning the effects of bilateral medial temporal lobe (MTL) removal in patient H. M. convincingly revealed the key role of the MTL in establishing new memories of day-to-day experiences. Later studies of H. M. and many other amnesic patients revealed dramatic dissociations between impaired explicit or declarative memory and preserved implicit or nondeclarative memory (cf. Schacter, 1987; Squire, 1992). These dissociations have played a critical role in shaping the view that memory consists of multiple forms or systems (see, e.g., Cohen & Eichenbaum, 1993; Schacter & Tulving, 1994; Squire & Zola-Morgan, 1991).

Research concerning amnesic patients has been rather less influential in the analysis of constructive aspects of memory—the errors, distortions, and illusions that sometimes characterize memory in the laboratory and in everyday life (Bartlett, 1932; Loftus, 1979; Neisser, 1967; Schacter, 1999b, 2001). Studies of confabulation in amnesic patients have led to ideas and hypotheses about mechanisms involved in memory monitoring and verification processes (e.g., Burgess & Shallice, 1996; Dalla Barba, 1993; Johnson, 1991; Moscovitch, 1989, 1995; Schacter, Norman, & Koutstaal, 1998; Schnider & Ptak, 1999), but apart from a focus on this question, students of amnesia have been largely mute concerning errors, illusions, and related constructive aspects of memory.

During the past few years, the situation has begun to change. A number of research groups have begun to focus on illusory or false memories in amnesic patients, attempting to use these observations to gain insight into the basic memory processes that underlie such errors. Cognitive psychologists generally agree that memory errors and illusions can provide revealing insights into the operation of basic memory processes (e.g., Jacoby, Kelley, & Dywan, 1989; Johnson, Hashtroudi, & Lindsay, 1993; Roediger, 1996; Schacter, 1996), and a similar consensus seems to be emerging in neuropsychology (see, e.g., Parkin, 1997; Rapcsak, Nielsen, Glisky, & Kaszniak, Chapter 10, this volume; and the papers in Schacter, 1999a). The purpose of this chapter is to review some of our own and others' work that has explored illusory memories in amnesic patients, and to consider several key empirical and theoretical issues that these studies have raised. We focus primarily on the phenomenon of "false recognition," which occurs when an individual mistakenly claims that a novel item or event is familiar. We first review recent studies that examine the nature and basis of false recognition in amnesic patients, and then turn to some unresolved questions and issues that require attention in future research.

FALSE RECOGNITION OF WORDS

Within experimental psychology, Underwood (1965) provided an early and influential demonstration of false recognition. He used a continuous-recognition paradigm in which words were presented and subjects made "old–new" decisions about each one. Some of the words were repeated; others were preceded by semantically, associatively, or physically similar words, or by entirely new words. Subjects responded "old" to the related words at a somewhat higher rate than to entirely unrelated new words, thereby documenting the occurrence of false recognition.

The first experimental study of false recognition in amnesic patients, reported by Cermak, Butters, and Gerrein (1973), employed a procedure similar to that introduced by Underwood (1965): a continuous-recognition paradigm in which new and old words were intermixed, and subjects responded "old" or "new" to each test item. New words were preceded by a single homophone, associate, or synonym. Six amnesic patients with Korsakoff's syndrome and six matched controls participated in the study. The amnesic patients showed a higher level of false recognition than did controls to both homophones and associates; neither amnesic patients nor controls in this experiment showed significant false recognition to synonyms.

Amnesic patients also recognized fewer of the words that had been presented previously. Cermak and colleagues (1973) argued that decreased true recognition and increased false recognition on the part of the amnesic patients reflected poor semantic encoding of target words: The patients did not encode deeply enough to remember words that were actually presented, or to reject new words that were related to previously exposed items.

To our knowledge, no studies followed up on the Cermak and colleagues (1973) results; in fact, we are unable to identify any studies reported during the next two decades that specifically focused on false recognition in amnesic patients. We initiated a series of such studies during the mid-1990s. We were inspired in part by striking results from a paradigm initially developed by Deese (1959), and later revived and modified by Roediger and McDermott (1995), which yields surprisingly high levels of false recognition in healthy young participants with intact brain function. In the Deese/Roediger–McDermott or "DRM" paradigm, participants study a list of associated words that all converge on a nonpresented "theme word," and are then tested for studied words, the nonstudied theme word, and other nonstudied words that are unrelated to the target item. For instance, participants might study a list of words such as "candy," "sour," "bitter," "good," "taste," "tooth," "nice," "honey," "soda," "chocolate," "heart," "cake," "eat," and "pie," which all converge on the nonpresented theme word "sweet." Roediger and McDermott (1995) found that participants claimed to recognize nonpresented theme words such as "sweet" about as often as they claimed to recognize words that actually appeared on the study lists (such as "taste"), falsely recognizing approximately 80% of the theme words—much higher than the relatively modest rates of false recognition reported by Underwood (1965). Participants expressed high confidence in their false-recognition responses and often claimed to possess specific recollections of having encountered the word. These results were soon confirmed and extended by others (e.g., Mather, Henkel, & Johnson, 1997; Norman & Schacter, 1997; Payne, Elie, Blackwell, & Neuschatz, 1996; Robinson & Roediger, 1997). Although there has been some debate regarding the nature and implications of this robust false-recognition effect (cf. Miller & Wolford, 1999; Roediger & McDermott, 1999; Wixted & Stretch, 2000), the existence and reliability of the effect have been firmly established.

Our main motivation in extending the DRM paradigm to amnesic patients was to gain insight into the brain mechanisms underlying this memory illusion. We (Schacter, Verfaellie, & Pradere, 1996) exposed 12 amnesic patients (6 patients with Korsakoff's syndrome, 5 patients with MTL damage resulting from anoxia or encephalitis, and 1 patient with thalamic damage) and matched controls to lists of semantic associates (e.g., "candy," "sour," "sugar," "bitter," "good," "taste," etc.). Participants were later tested with previously studied words (e.g., "taste"), theme words that were semantically related to previously presented words (e.g., "sweet"), and new words that were unrelated to previously studied ones (e.g., "point").

As expected from previous studies of recognition memory in amnesia (e.g., Haist, Shimamura, & Squire, 1992), amnesic patients had difficulty distinguishing between previously studied words and unrelated new items: They attained fewer hits to old words, and made more false alarms to new unrelated words, than did matched controls. More importantly, the amnesic patients also showed significantly lower levels of false recognition to the nonpresented theme words than did the controls (who showed the expected high levels of false recognition to the theme words). Reduced or "impaired" false recognition was observed both in the amnesic patients with Korsakoff's syndrome and the mixed-etiology amnesic group that consisted primarily of patients with MTL damage. More recently, Melo, Winocur. and Moscovitch (1999) have replicated these findings in amnesic patients with mixed etiologies involving damage to MTL, diencephalic, or frontal lobe structures.

To assess the generality of our results beyond the DRM paradigm, we (Schacter, Verfaellie, & Anes, 1997) attempted to replicate the finding of reduced false recognition for semantic associates in amnesic patients with a new set of semantically related words. We also sought to determine whether the effect is restricted to the semantic domain, or extends to perceptual false recognition. It has been argued that perceptually based memory processes are preserved in amnesic patients, whereas conceptually or semantically based processes are impaired (e.g., Blaxton, 1989, 1995; see also Wagner, Gabrieli, & Verfaellie, 1997). From this perspective, amnesic patients might show "preserved" false recognition (i.e., false-positive responding rates similar to those shown by controls) after studying perceptually related materials.

To test this hypothesis, we (Schacter, Verfaellie, & Anes, 1997) used materials based on the experiments of Shiffrin, Huber, and Marinelli (1995). Shiffrin and colleagues found that when college students studied lists of words, such as "fade," "fame," "fake," "mate," "late," "date," and "rate," they later showed false recognition to perceptually related words, such as "fate." We adapted their paradigm for use with amnesic patients. Results of two experiments were largely similar to the earlier ones (Schacter, Verfaellie, & Pradere, 1996). Both patients with MTL damage and patients with Korsakoff's syndrome exhibited impaired true recognition of studied words in the semantic and perceptual conditions. More important, they also showed reduced false recognition for new words, whether they were semantically or perceptually related to previously studied words (for a summary of false-recognition results from our two studies, see Table 9.1).

TABLE 9.1. False Recognition of Words by Amnesic Patients and Matched Controls in Experiments by Schacter, Verfaellie, and Pradere (1996) and Schacter, Verfaellie, and Anes (1997)

Type of materials	Amnesic patients	Matched controls
Semantic associates	.16	.57
Conceptually related	.16	.24
Perceptually related	.12	.40

Note. Data concerning semantic associates from Schacter, Verfaellie, and Pradere (1996); data concerning conceptually and perceptually related words from Schacter, Verfaellie, and Anes (1997). The table displays corrected false-recognition scores in which false alarms to unrelated new words are subtracted from false alarms to related new words. Data from Schacter, Verfaellie, and Pradere are averaged across conditions in which either a free-recall test or an arithmetic test preceded the recognition test.

There was one difference worth noting between the results of our two studies. In the Schacter, Verfaellie, and Pradere (1996) study, amnesic patients showed reduced false recognition on two different measures: (1) the absolute proportion of "old" responses to nonpresented theme words; and (2) a corrected false-recognition measure in which the proportion of false alarms to unrelated new words is subtracted from the proportion of false alarms to theme words. The latter measure is an important one, because amnesic patients sometimes exhibit a bias to respond "old" to unrelated new words, reflecting the fact that they have great difficulty distinguishing between old and new items. Because such a tendency could inflate estimates of false recognition to related theme words, it is important to use corrected false-recognition measures (or signal detection analyses) that take into account such a response bias. In the Schacter, Verfaellie, and Anes (1997) study, amnesic patients showed a heightened tendency to respond "old" to new words that were not perceptually or conceptually related to previously studied ones. As a consequence, the overall proportion of "old" responses to semantically or perceptually related new words did not differ significantly between amnesic patients and controls. Instead, reduced false-recognition to semantically or perceptually related new words was evident only in corrected false-recognition measures that took account of this bias. Nonetheless, we (Schacter, Verfaellie, & Anes, 1997) concluded from these data that amnesic patients showed deficient retention of what we termed "conceptual gist" and "perceptual gist"—information about the shared semantic or perceptual features of studied information that underlies false-recognition responses (cf. Brainerd & Reyna, 1998; Reyna & Brainerd, 1995; Schacter, Norman, & Koutstaal, 1998).

The findings from the foregoing studies suggest that the MTL/diencephalic structures that are damaged in amnesic patients play a role in storing and/or retrieving the semantic or perceptual information that supports false recognition in normal controls. However, these findings contrast sharply with the results discussed earlier from Cermak and colleagues (1973), where amnesic patients showed increased, rather than decreased, false recognition. What accounts for the contrasting pattern of results? In both of our studies, we suggested that when numerous associates are presented for study, as in the DRM paradigm or the related procedures used in the Schacter, Verfaellie, and Anes (1997) study, normal controls establish a well-organized representation of the general semantic or perceptual features of the list—that is, the conceptual or perceptual gist. When this gist representation is matched by a new theme word, normal controls experience a strong sense of familiarity or recollection that results in a powerful false-recognition effect. Amnesic patients, by contrast, encode or retain less gist information and thus show reduced levels of false recognition. Likewise, control subjects can use the potent gist representation to reject new words that are unrelated to, or clearly incongruent with, the perceptual or conceptual gist of studied items, whereas amnesic patients cannot.

Different factors may be operating when only a single related item precedes a new word, as in the study by Cermak and colleagues (1973). Under these conditions, controls should establish a much weaker gist representation than when numerous associates are studied (for particularly clear evidence relating increased category size to increased gist memory in normal controls, see Brainerd & Reyna, 1998, Exp. 3). An equally important factor is that healthy controls can use their intact explicit memory abilities to counter or suppress any familiarity that may be elicited by a related new word. For instance, a healthy control who encounters the nonstudied word "table," and can recollect having previously studied the associate "chair," can use this information to avoid making a false-recognition response. In contrast, amnesic patients are less able to use explicit recollection to suppress whatever weak sense of familiarity might be elicited by a new related word. In this scenario, amnesic patients can exhibit increased levels of false recognition compared with control subjects.

Based on this analysis, we (Schacter, Verfaellie, & Pradere, 1996) suggested that it should be possible to show higher or lower levels of false recognition in amnesic patients compared with controls by creating conditions that are more or less conducive to using explicit recollection to suppress false recognition. We (Schacter, Verfaellie, Anes, & Racine, 1998) tested this idea by repeatedly presenting amnesic patients and matched controls with the same study lists of DRM semantic associates, and testing them with lists composed of previously studied words, semantically related lures, and unrelated lures. We hypothesized that with repeated study and testing of the same lists, normal controls would show increasing explicit recollection of previously presented words, and would use this memory to reduce false alarms to related words that were not presented (see, e.g., McDermott, 1996). By contrast, we hypothesized a contrasting pattern of results in amnesic patients. Based on the data discussed so far, we expected that amnesic patients would continue to show reduced levels of both true and false recognition compared to controls on the first study–test trial. We also expected that with repetition of study–test lists, amnesic patients would continue to show impaired levels of veridical recognition memory compared to controls, although they should show some increases in true recognition across trials.

The key prediction involved levels of false recognition to nonstudied theme words as a function of repetition. We hypothesized that amnesic patients, in contrast to controls, would be unable to use explicit recollection to reduce false recognition across trials. On the contrary, their false recognition might even increase as a function of repetition. Consider that with repetition of study list associates, the theme or gist of the study list should become increasingly accessible. If, as we suggested, the recognition performance of amnesic patients in the DRM paradigm is based largely on a degraded representation of semantic or conceptual gist, and if that representation benefits from repetition, then amnesic patients might show enhanced false recognition across trials. Thus we predicted an interaction between groups and trials for false recognition of related lures: Reduced false recognition in amnesic patients on the first trials should be greatly attenuated, and possibly eliminated or even reversed over the course of multiple study–test trials.

Consider also that with repetition of study list associates, the theme or gist of the study lists should become increasingly accessible. If, as we suggested, the recognition performance of amnesic patients in the DRM paradigm is based largely on a degraded representation of semantic or conceptual gist, and if that representation benefits from repetition, then amnesic patients might even show increased false recognition across trials.

Results were generally consistent with these predictions. Amnesic patients showed reduced levels of corrected false recognition on the first test trial, thus replicating previous results. With repetition, false recognition of theme words in healthy controls declined significantly. In contrast, amnesic patients with MTL damage did not reduce their false-recognition responses with repetition; they showed a flat or fluctuating pattern across trials. Patients with Korsakoff's syndrome actually showed a significant *increase* in false recognition across trials. Although the reasons for the differing patterns in the amnesic group are not entirely clear, it is possible that problems resulting from frontal lobe dysfunction, often observed in Korsakoff patients, made it especially difficult for them to suppress the growing influence of semantic gist across trials (see Schacter, Verfaellie, Anes, & Racine, 1998, for a more detailed discussion). Consistent with this view, in more recent work using the same paradigm in patients with Alzheimer's disease—who are often characterized by both MTL and frontal dysfunction—Budson, Daffner, Desikan, and Schacter (2000) have documented a similar pattern of increasing false recognition across trials.

These results support the analysis we put forward in an attempt to resolve the apparent discrepancy between Cermak and colleagues' (1973) finding of increased false recognition in amnesic patients for lures that were related to only a single study item and our findings (Schacter, Verfaellie, & Anes, 1997; Schacter, Verfaellie, & Pradere, 1996) of decreased false recognition in amnesic patients for lures that were related to multiple study items. When experimental conditions allow or encourage the use of explicit recollection to counter or suppress false-recognition responses—as in Cermak and colleagues' experiment, or in the five-repetition condition of the Schacter, Verfaellie, and colleagues (1998) study—healthy controls can do so to a much greater extent than can amnesic patients. By contrast, when experimental conditions promote the development of powerful gist representations in normal controls, but work against the use of explicit recollection to counter them—as in the Schacter, Verfaellie, and Pradere (1996) and Schacter, Verfaellie, and Anes (1997) experiments—amnesic patients show reduced false recognition. These observations lead us to conceptualize false recognition as the outcome of a kind of competition between two opposing forces: conceptual or perceptual gist, which fuels the development of illusory memories, and explicit recollection of specific information, which can help to keep in check the undue influence of gist representations.

FALSE RECOGNITION OF PROTOTYPES AND PICTURES

Though the foregoing analysis seems to capture well the main features of extant data concerning false recognition in amnesic patients, it is not entirely straightforward. One problem concerns our suggestion that gist information (see, e.g., Brainerd & Reyna, 1998; Payne et al., 1996), or overreliance on shared semantic or perceptual features (Schacter, Norman, & Koutstaal, 1998), is primarily responsible for false recognition in the DRM and related paradigms. Though this is a plausible suggestion, other interpretations are also possible. For instance, false recognition in the DRM procedure may be a consequence of "implicit associative responses" (Underwood, 1965) that occur when participants are exposed to lists of semantic associates during the study phase of the experiment. When studying a list of semantic associates, participants may activate or even consciously generate the nonpresented lure word (i.e., "sweet"). On a later memory test, false recognition may occur because participants experience a type of source confusion, mistakenly remembering that they heard or saw the theme word that they themselves have generated (Roediger, Balota, & Watson, 2001; Roediger & McDermott, 1995). From this perspective, amnesic patients may be less likely to generate the theme word at the time of study, and therefore show reduced false recognition. Alternatively, amnesic patients may generate the theme word as frequently as control subjects, but later forget this generated item, just as they forget other individual items that actually appeared on the list. Therefore, reduced false recognition may occur because amnesic patients are not as susceptible as controls to source confusions from internally generated items. In short, reduced false recognition in amnesic patients may not be attributable to degraded gist representations, as we have suggested, but rather to impaired memory for specific, individual items.

To address this issue, and also to expand the generality of our research, we have conducted two sets of experiments involving false recognition of pictorial materials. We made use of paradigms in which false recognition is unlikely to result from generating specific items at study, and hence can be more confidently attributed to the influence of gist representations than in the DRM procedure.

In the first such experiment, we (Koutstaal, Schacter, Verfaellie, Brenner, & Jackson, 1999) examined true and false recognition of novel abstract patterns. Each pattern was a complex, multifeatured shape constructed from a particular prototype. Eighteen different prototypes were used. Individual shapes were manipulated so as to generate a metric of "transformational distance"—the degree of perceptual similarity between a particular item and its prototype, as determined by the number and magnitude of perceptual changes made to particular attributes of each prototype. Thus there were 18 different categories of abstract shapes, each defined relative to a particular prototype. Within each category, 3 different transformational distances were used: near, middle, and far distortions of the prototype. This manipulation was crossed with a manipulation of category size: Participants studied either 1, 3, 6, or 9 shapes from a particular category.

During the study phase, participants rated the complexity of each shape. On a subsequent recognition test, participants made "old–new" decisions about studied and nonstudied shapes from each transformational distance (near, middle, and far), about the prototype from each category (which had never been presented) and also about novel shapes that were unrelated to the previously studied categories.

We reasoned that during the study phase of the experiment, it was highly unlikely that patients or controls would generate a nonpresented prototype or distortion, in the same sense that they might generate a word such as "sweet" in the DRM paradigm. However, previous studies had already shown that normal subjects exhibit high levels of false recognition to nonpresented prototypes of dot patterns (e.g., Posner & Keele, 1968), and we expected to observe similar effects in our controls. The key question was whether amnesic patients would exhibit reduced false recognition of prototypes.

The general answer to this question was "yes." Reduced or "impaired" false recognition of prototypes was more evident in the mixed-etiology amnesic group consisting primarily of patients with MTL damage than in the Korsakoff group (mainly because the alcoholic control patients to whom we compared the patients with Korsakoff's syndrome showed unusually high levels of false recognition even to unrelated new items). Amnesic patients also tended to show reduced false recognition of nonprototypical items in the conditions that encouraged the development of robust gist representations in healthy controls: with near and middle distortions of the prototypes, and following the study of large categories (see Figure 9.1). To the extent that participants are unlikely to generate these novel items during the study phase, these results support our claim that reduced false recognition in amnesic patients is attributable to the influence of a degraded gist representation, rather than to poor memory for a specific item generated during the study phase.

We also found that under conditions that did not promote the development of powerful gist representations in control subjects, such as the far transformational distance or small category size, there was little or no false-recognition "impairment" in the amnesic group (Figure 9.1). This latter finding fits well with our preceding analysis of differences between the early results of Cermak and colleagues (1973) and our later results with the DRM semantic-associates approach and related procedures (Schacter, Verfaellie, & Anes, 1997; Schacter, Verfaellie, & Pradere, 1996; Schacter, Verfaellie, et al., 1998). When experimental conditions favor the development of robust gist representations that cannot be easily countered by explicit recollection, amnesic patients are likely to show reduced false recognition. But when experimental conditions result in weaker gist representations that can be countered with explicit recollection, controls can suppress their false-recognition responses, but amnesic patients cannot. These contrasting processes can reduce, eliminate, or even reverse the differences between amnesic patients and controls that are observed in "strong-gist" conditions.

Note, however, that these findings were obtained in an experimental paradigm that relies almost exclusively on perceptual memory processes. It is thus unclear whether the principal conclusions we have drawn from this experiment also apply to conditions in which semantic or conceptual factors play a role—as in our previous studies with the DRM and related procedures. To examine the issue, we (Koutstaal, Verfaellie, & Schacter, 2001) used a para-

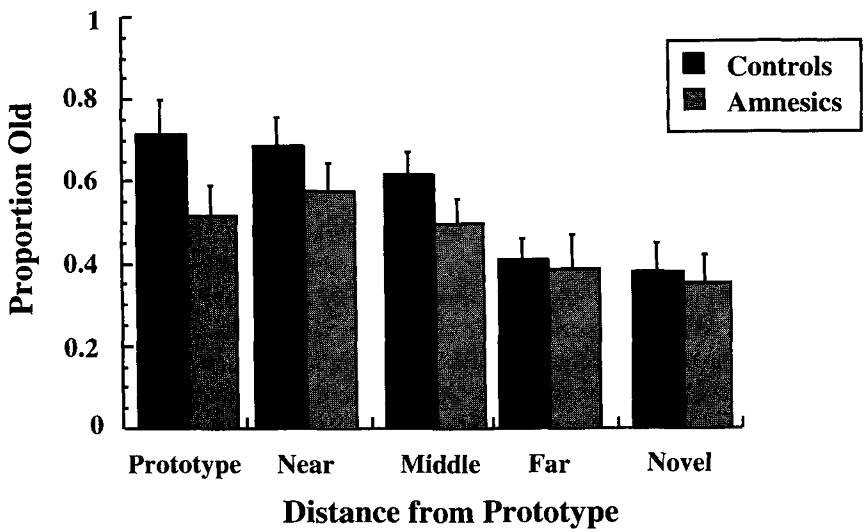

FIGURE 9.1. False recognition of visual patterns by amnesic patients and matched controls in an experiment by our group (Koutstaal, Schacter, Verfaellie, Brenner, & Jackson, 1999). In this experiment, participants studied complex, multifeatured shapes that were each constructed from a particular prototype. There were 18 different categories of shapes, each defined relative to a particular prototype. Within each category, three different transformational distances were used: near, middle, and far distortions of the prototype. The data in the figure show the proportions of "old responses" (false alarms) to the prototype, each of the three types of distortions, and novel items that were not related to a prototype. See text for further explanation.

digm involving categorized pictures that draws on both perceptual and conceptual memory processes. Initially developed (Koutstaal & Schacter, 1997) to examine age-related differences in false recognition, the procedure involves presenting participants with detailed colored pictures of common objects, such as cats or teapots, and varying the size of each category (ranging from 18 different exemplars to only a single exemplar). Participants are later given an "old–new" recognition test consisting of studied pictures, related lure items consisting of nonstudied pictures from previously studied categories, and nonstudied pictures that are unrelated to studied categories. Because the related lures are detailed pictures that participants have not seen prior to the experiment, it is highly unlikely that participants generate specific representations of those pictures at the time of study (just as we have argued for the abstract-shapes paradigm discussed above). Thus false recognition in the paradigm we used probably reflects the influence of general or gist-based similarity—both perceptual and conceptual—between related lure pictures and previously studied pictures.

Based on our previous results, we predicted reduced or impaired false recognition of related lures in amnesic patients following the study of a large number of categorically related pictures. An initial experiment provided evidence of trends in this direction, but the tendency toward reduced false recognition in the amnesic group did not attain statistical significance. We suspected that the relatively weak effects might have been attributable to the use of item-specific recollection by control subjects to suppress false-recognition responses. Accordingly, we carried out an additional experiment designed to minimize such influences. This experiment yielded strong evidence for reduced false recognition of related lures from large categories in amnesic patients (for further details and qualifications, see Koutstaal et al., 2001). These results thus extend the findings of the Koutstaal and colleagues (1999) study beyond the domain of purely perceptual memory processes into the domain of conceptual memory. They also further confirm that amnesic patients show reduced false recognition under conditions where memory for gist, rather than source confusions regarding individual items, is the primary determinant of illusory memory.

One final observation from the latter experiments should be noted. Although our emphasis here has been on establishing the conditions under which amnesic patients show impaired memory for the perceptual or conceptual gist of studied items, a further question one might ask concerns *the relative magnitude* of their impairment in gist-based memory versus their impairment in the explicit recollection of item-specific information. To address this question, one might contrast in both amnesic patients and controls the level of false recognition of related lures with the level of veridical recognition for related targets. However, the usefulness of this comparison is to some extent vitiated by the possibility that the same general similarity or gist information that contributes to false-recognition responses might *also* help to support veridical recognition for related list items. The inclusion of a manipulation of category size in the experiments using categorized pictures (Koutstaal et al., 2001) enabled us to examine the extent of impairment of gist-based versus item-specific memory in amnesic patients while circumventing this difficulty. Rates of false recognition for category lures that were related to a large number of study items provided an index of gist memory, whereas rates of veridical recognition for targets that were unrelated to any of the other list items (one-of-a-kind or single-item categories) provided an index of item-specific memory. A comparison of these two measures showed that although (as noted above) gist memory in amnesic patients was significantly impaired relative to gist memory in controls, the impairment in item-specific memory was even greater. Stated differently, although amnesic patients showed reduced "memory" for items that were similar to but not actually identical to items that they had studied, this deficit in gist memory was not as pronounced as the deficit they demonstrated for actually studied items when those items were unrelated to any of the other items they had encountered.

This finding underscores an important point regarding the nature of the impairment of gist memory in amnesia. Under conditions that foster the formation and retention of robust gist representations among control participants, patients with amnesia show reduced false recognition. Nonetheless, the rates of false recognition shown by amnesic patients under these same conditions are also usually elevated above the level found for entirely unrelated "new" items. Relative to controls with intact memory and intact brain function, amnesic patients show a partial—and less efficient—ability to extract, retain, and retrieve information concerning the conceptual or perceptual gist of studied items, rather than an *absence* of an ability to benefit from such information.

Several factors may contribute to the differential levels of impairment shown by amnesic patients for gist information relative to item-specific information. Memory for gist and item-specific information may depend on qualitatively different underlying processes (see, e.g., Reyna & Brainerd, 1995), with the latter more impaired than the former in amnesic patients. Alternatively, remembering item-specific information may be in some sense more difficult than remembering gist information. Given their generally impaired explicit or declarative memory capacities, the greater difficulty associated with remembering item-specific information may have a disproportionately large impact on the performance of amnesic patients compared with controls.

In either scenario, two possibly related findings merit consideration. First, research by Verfaellie and Cermak (1994) using a rather different approach (Watkins & Kerkar, 1985) suggested that whereas amnesic patients may show especially impaired memory for specific occurrences, memory for repeated occurrences may be "superadditive": Recall of twice-presented items was higher than would be expected on the basis of recall of one of two (different) once-presented items. However, a particular feature concerning these items (color) was less likely to be recalled than was the color of the once-presented items. Second, recent work examining the relative impairment of recollection versus familiarity in amnesia has suggested that although both components are impaired (Knowlton & Squire, 1995), recollection is particularly adversely affected (Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998). The comparison of veridical recognition for one-of-a-kind target items versus false recognition of many-exemplar lure items may partially map onto a similar distinction, with successful recognition of one-of-a-kind targets especially (albeit not exclusively or entirely) drawing on recollection, and incorrect identification of many-of-a-kind lures drawing especially on familiarity. Again, questions concerning whether these differences reflect qualitatively distinct underlying processes or quantitative variations in difficulty remain to be addressed in future studies.

UNRESOLVED ISSUES AND FUTURE DIRECTIONS

The main lesson from our studies of false recognition in amnesic patients is relatively clear and consistent: Under conditions that favor the development of strong gist representations, and that work against using explicit recollection to suppress such representations, amnesic patients show reduced false recognition compared to healthy controls. Nonetheless, there remain a number of issues and puzzles that require examination and clarification.

First, though we have focused on false recognition, data from the DRM procedure also exist concerning free recall. As shown initially by Deese (1959), and later by Roediger and McDermott (1995) and others (e.g., Read, 1996), participants frequently intrude nonpresented theme words such as "sweet" during a free-recall test. In our initial study using the DRM procedure, we found that amnesic patients intruded the theme word in free recall about as often as control subjects did (Schacter, Verfaellie, & Pradere, 1996). Moreover, because amnesic patients recalled fewer studied items than controls, when expressed as a proportion of correctly recalled items, amnesic patients actually showed relatively *greater* false recall than did controls. Although these false-recall data are apparently in conflict with our results on false recognition, one further finding is critical to interpreting these data: Amnesic patients also made many more intrusions of unrelated words than did the controls. When these intrusions were taken into account, amnesic patients actually intruded a smaller proportion of theme words than did controls. Still, control subjects produced more studied words than theme words on the free-recall test, whereas amnesic patients showed the opposite effect.

Taken together, these findings suggest that amnesic patients have little or no access to specific information about studied items; their recall performance appeared to rely nearly exclusively on degraded gist memory. These considerations may help to make some sense of free-recall data reported by Melo and colleagues (1999), who found that amnesic patients with MTL/diencephalic damage ($n = 4$) intruded more theme words on a free-recall test than did controls (these same patients showed reduced false recognition of theme words on a subsequent recognition test). Contrary to the results of the Schacter, Verfaellie, and Pradere (1996) study, these patients did not, as a group, show heightened intrusions of unrelated words (although one of the four patients did show this effect). Though it is unclear why Melo and colleagues found greater intrusion of theme words in amnesic patients than controls, and we did not (differences in the relative severity of amnesia or other characteristics of patients in the two experiments could have played a role), both studies agree on one important finding: Amnesic patients produce a higher proportion of theme words than list words on a free-recall task, whereas normal controls show the opposite pattern. Thus it seems clear that patients tend to rely more on gist information during recall than do controls. Further research with free-recall tests is needed to characterize the extent to which such information is preserved or degraded.

A second issue that requires further research concerns the relation between our finding of reduced false recognition of prototypes and other forms of gist information in amnesia, and previous reports that amnesic patients show normal acquisition of prototypes in category-learning experiments (e.g., Knowlton & Squire, 1993). If amnesic patients can learn new categorical information, as expressed by their ability to correctly classify prototypes on the basis of recent learning, why do they not falsely recognize prototypes to the same extent as control subjects do? Early evidence reported by Kolodny (1994) in the context of a category-learning study suggested that amnesics falsely recognize prototypes even more often than control subjects. However, as we have noted elsewhere (Koutstaal et al., 1999), Kolodny did not report data concerning amnesic patients' false alarms to unrelated lure items. If amnesic patients showed elevated false-alarm rates to unrelated lures, this bias in responding could account for their apparent "preservation" of false recognition to prototypes. Stated slightly differently, amnesic patients in this experiment might well have shown reduced false recognition if Kolodny had used corrected false-recognition measures (on which we have relied) that take account of general tendencies to respond "old" to both unrelated and related new items.

Another possible explanation of preserved category learning and reduced false recognition of prototypes in amnesia is that category-learning experiments use indirect or implicit tests to assess classification of prototypes (Knowlton & Squire, 1993), whereas experiments examining false recognition in amnesia have relied exclusively on direct or explicit tests. It is well known that implicit and explicit tests can tap different types of memory processes or representations (cf. Jacoby, 1991; Roediger, 1990; Schacter, Chiu, & Ochsner, 1993; Squire, 1992), and this could be the source of differences between false-recognition and category-learning experiments (Kitchener & Squire, 2000; Koutstaal et al., 1999). One way to examine this issue would be to alter recognition test instructions so that they do not require patients to make "old–new" judgments about specific list items. For example, Brainerd and Reyna (1998; see also Schacter, Cendan, Dodson, & Clifford, in press) used test instructions in a study with the DRM materials that required subjects to respond "old" when a test probe fit a previously studied theme, regardless of whether the specific item had appeared previously. If amnesic patients do indeed possess intact gist representations, but fail to gain access to them on recognition tests, then patients may endorse nonstudied theme words as often as controls do when tested with instructions that require only memory for the theme or gist of previously studied lists. We are currently examining this possibility experimentally.

Note that Kitchener and Squire (2000) have recently reported that amnesic patients showed impaired prototype classification on a category-learning task that used verbal stimuli; prior demonstrations of preserved category-learning had all made use of nonverbal stimuli. If verbal category learning is impaired in amnesic patients, and this deficit reflects an impoverished gist representation, then there may be no conflict at all with our repeated finding of reduced false recognition of verbal materials by amnesic patients in the DRM paradigm (Schacter, Verfaellie, & Anes, 1997; Schacter, Verfaellie, & Pradere, 1996; Schacter, Verfaellie, et al., 1998). Yet this interpretation still leaves open the question of why we found reduced false recognition of prototypes with nonverbal materials (Koutstaal et al., 1999). As Kitchener and Squire (2000) note, in addition to task differences (category classification vs. "old–new" recognition), the contrasting findings could result from another important procedural difference in category-learning and false-recognition studies. Successful category-learning studies typically involve training on only a single category, whereas we have exposed amnesic patients to multiple categories. Further research will be required to examine directly this and other possible sources of the observed differences.

Finally, we have noted earlier that an important goal in studying false recognition in amnesic patients is to gain insight into the brain mechanisms of memory illusions. In addition to studies of amnesic patients, we and others are pursuing complementary approaches to this basic issue. For instance, several neuroimaging studies provide evidence for MTL activation during false recognition (Cabeza, Rao, Wagner, Mayer, & Schacter, 2001; Schacter, Buckner, Koutstaal, Dale, & Rosen, 1997; Schacter, Reiman, et al., 1996). This evidence converges nicely with our findings of reduced false recognition in amnesic patients with MTL damage. Recent electrophysiological evidence has begun to delineate the ways in which true and false recognition differ (Curran, Schacter, Johnson, & Spinks, 2001; Gonsalves & Paller, 2000; Fabiani, Stadler, & Wessels, 2000). Other studies have revealed increased false recognition in patients with damage to specific regions of the frontal lobes (e.g., Parkin, Ward, Bindschaedler, Squires, & Powell, 1999; Rapcsak et al., Chapter 10, this volume; Rapcsak, Reminger, Glisky, Kaszniak, & Comer, 1999; Schacter, Curran, Galluccio, Milberg, & Bates, 1996). When the results from these different approaches are combined, we are optimistic that the study of memory illusions, distortions, and other constructive aspects of remembering that have long intrigued cognitive psychologists will make an increasingly important contribution to the neuropsychology of memory.

ACKNOWLEDGMENTS

Preparation of this chapter was supported by Grant No. NS26985 from the National Institute of Neurological Disorders and Stroke, Grant No. AG08441 from the National Institute on Aging, and Grant No. MH57681 from the National Institute of Mental Health. We thank Steve Prince for his assistance.

REFERENCES

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

Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 657–668.

Blaxton, T. A. (1995). A process-based view of memory. *Journal of the International Neuropsychological Society, 1*, 112–114.

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

Budson, A. E., Daffner, K. R., Desikan, R., & Schacter, F. 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.

Cabeza, R., Rao, S. M., Wagner, A. D., Mayer, A. R., & Schacter, D. L. (2001). Can medial temporal lobe regions distinguish true from false?: An event-related fMRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences USA, 98*, 4805–4810.

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

Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system.* Cambridge, MA: MIT Press.

Curran, T., Schacter, D. L. Johnson, M. K., & Spinks, R. (2001). Brain potentials reflect behavioral differences in true and false recognition. *Journal of Cognitive Neuroscience, 13*, 201–216.

Dalla Barba, G. (1993). Confabulation: Knowledge and recollective experience. *Cognitive Neuropsychology, 10*, 1–20.

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

Fabiani, M., Stadler, M. A., & Wessels, P. M. (2000). True but not false memories produce a sensory signature in human lateralized brain potentials. *Journal of Cognitive Neuroscience, 12*, 941–949.

Gonsalves, B., & Paller, K. A. (2000). Neural events that underlie remembering something that never happened. *Nature Neuroscience, 3*, 1316–1321.

Haist, F., Shimamura, A. P., & Squire, L. R. (1992). On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 691–702.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Erlbaum.

Johnson, M. K. (1991). Reality m onitoring: Evidence from confabulation in organic brain disease patients. In G. P. Prigatano & D. L. Schacter (Eds.), *Awareness of deficit after brain injury: Clinical and theoretical issues* (pp. 176–197). New York: Oxford University Press.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3–28.
Kitchener, E. G., & Squire, L. R. (2000). Impaired verbal category learning in amnesia. *Behavioral Neuroscience, 114*, 907–911.
Knowlton, B. J., & Squire, L. R. (1993). The learning of categories: Parallel brain systems for item memory and category level knowledge. *Science, 262*, 1747–1749.
Knowlton, B. J., & Squire, L. R. (1995). Remembering and knowing: Two different expressions of declarative memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 699–710.
Kolodny, J. A. (1994). Memory processes in classification learning: An investigation of amnesic performance in categorization of dot patterns and artistic styles. *Psychological Science, 5*, 164–169.
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., Verfaellie, M., Brenner, C. J., & Jackson, E. M. (1999). Percfeptually based false recognition of novel objects in amnesia: Effects of category size and similarity to category prototypes. *Cognitive Neuropsychology, 16*(3, 4, 5), 317–341.
Koutstaal, W., Verfaellie, M., & Schacter, D. L. (2001). Recognizing identical vs. similar categorically related common objects: Further evidence for degraded gist-representations in amnesia. *Neuropsychology*, 268–289.
Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.
Mather, M., Henkel, L. A., & Johnson, M. K. (1997). Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaire compared. *Memory and Cognition, 25*, 826–837.
McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language, 35*, 212–230.
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.
Moscovitch, M. (1989). Confabulation and the frontal systems: Strategic versus associative retrieval in neuropsychological theories of memory. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 133–160). Hillsdale, NJ: Erlbaum.
Moscovitch, M. (1995). Confabulation. In D. L. Schacter, J. T. Coyle, G. D. Fischbach, M.-M. Mesulam, & L. E. Sullivan (Eds.), *Memory distortion: How minds, brains, and societies reconstruct the past* (pp. 226–254). Cambridge, MA: Harvard University Press.
Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
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.
Parkin, A. J. (1997). The neuropsychology of false memory. *Learning and Individual Differences, 9*, 341–357.
Parkin, A. J., Ward, J., Bindschaedler, C., Squires, E. J., & Powell, G. (1999). False recognition following frontal lobe damage: The role of encoding factors. *Cognitive Neuropsychology, 16*, 243–265.
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.
Rapcsak, S. Z., Reminger, S. L., Glisky, E. L., Kazniak, A. W., & Corner, J. F. (1999). Neuropsychological mechanisms of false facial recognition following frontal lobe damage. *Cognitive Neuropsychology, 16*, 267–292.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*, 1–75.

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.

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

Roediger, H. L., III. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45*, 1043–1056.

Roediger, H. L., III. (1996). Memory illusions. *Journal of Memory and Language, 35*, 76–100.

Roediger, H. L., III, Balota, D. A., & Watson, J. M. (2001). Spreading activation and the arousal of false memories. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95–115) Washington, DC: American Psychological Association Press.

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.

Schacter, D. L. (1987). Memory, amnesia, and frontal lobe dysfunction. *Psychobiology, 15*, 21–36.

Schacter, D. L. (1996). *Searching for memory: The brain, the mind, and the past.* New York: Basic Books.

Schacter, D. L. (Ed.). (1999a). *The cognitive neuropsychology of false memories.* Hove, England: Psychology press.

Schacter, D. L. (1999b). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist, 54*(3), 182–203.

Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and remembers.* Boston: Houghton Mifflin.

Schacter, D. L., Buckner, R. L., Koutstaal, W., Dale, A. M., & Rosen, B. R. (1997). Late onset of anterior prefrontal activity during retrieval of veridical and illusory memories: An event-related fMRI study. *NeuroImage, 6*, 259–269.

Schacter, D. L., Cendan, D. L., Dodson, C. S., & Clifford, E. R. (in press). Retrieval conditions and false recognition: Testing the distinctiveness heuristic. *Psychonomic Bulletin and Review.*

Schacter, D. L., Chiu, C. Y. P., & Ochsner, K. N. (1993). Implicit memory: A selective review. *Annual Review of Neuroscience, 16I*, 159–182.

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

Schacter, D. L., Norman, K. A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annual Review of Psychology, 49*, 289–318.

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

Schacter, D. L., & Tulving, E. (Eds.). (1994). *Memory systems 1994.* Cambridge, MA: MIT Press.

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. (1998). When true recognition suppresses false recognition: Evidence from amnesic patients. *Journal of Cognitive Neuroscience, 10*, 668–679.

Schacter, D. L., Verfaellie, M., & Pradere, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language, 35*, 319–334.

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

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery and Psychiatry, 20*, 11–21.

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.

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., & Zola-Morgan, M. (1991). The medial temporal lobe memory system. *Science*, *253*, 1380–1386.

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

Verfaellie, M., & Cermak, L. S. (1994). Acquisition of generic memory in amnesia. *Cortex*, *30*, 293–303.

Wagner, A. D., Gabrieli, J. D. E., & Verfaellie, M. (1997). Dissociations between familiarity processes in explicit–recognition and implicit–perceptual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 305–232.

Watkins, M. J. & Kerkar, S. P. (1985). Recall of a twice-presented item without recall of either presentation: Generic memory for events. *Journal of Memory and Language*, *24*, 666–678.

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

Yonelinas, A. P., Kroll, N. E., Dobbins, I., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: Convergence of remember–know, process dissociation, and receiver operating characteristic data. *Neuropsychology*, *12*, 323–339.