COGNITIVE NEUROPSYCHOLOGY, 1999, 16 (3/4/5), 317–341

PERCEPTUALLY BASED FALSE RECOGNITION OF NOVEL OBJECTS IN AMNESIA: EFFECTS OF CATEGORY SIZE AND SIMILARITY TO CATEGORY PROTOTYPES

Wilma Koutstaal and Daniel L. Schacter
Harvard University, Cambridge, USA

Mieke Verfaellie
Veterans Affairs Medical Center and Boston University School of Medicine, Boston, USA

Carolyn Brenner and Eric M. Jackson
Harvard University, Cambridge, USA

Previous research suggests that amnesics may show impaired semantically based false recognition under conditions where control participants show high levels of gist-based errors, but little or no impairment when controls show less robust false recognition. Using abstract novel objects, we examined perceptually based false recognition in amnesics under conditions designed to induce differing levels of false recognition in controls. Whereas amnesics showed significantly impaired false recognition for category prototypes, and numerically impaired false recognition when many perceptually similar exemplars were studied—conditions where controls showed high rates of illusory recognition—amnesics and controls showed lower, and comparable, levels of false recognition when few related exemplars were studied, or lures were at a far transformational distance from the prototype. Although amnesics may be able to extract some information regarding the perceptual "gist" of studied items, they appear to do so less efficiently than controls.

INTRODUCTION

The strategy of using illusory recollection as a means to probe the nature of memory encoding and retrieval processes has a long and diverse history in psychology (for general review and discussion see Estes, 1997; Johnson, Hashtroudi, & Lindsay, 1993; McClelland, 1995; Roediger, 1996; Schacter, 1995; Schacter, Norman, & Koutstaal, 1998a). Although many studies have focused on participants with intact memory, the uncovering of robust forms of false recollection, and the effort to understand the bases of illusory recollection have also prompted investigations of such errors in persons with impaired memory (for review, see Schacter, 1997; Schacter et al., 1998a).

Several studies have focused on amnesic patients: Individuals who, as a consequence of damage to medial temporal and diencephalic brain regions arising from a variety of possible causes (e.g.

Requests for reprints should be addressed to Wilma Koutstaal, Harvard University, Department of Psychology, 33 Kirkland Street, Cambridge, MA 02138, USA (Tel: (617) 495-3856; Fax: (617) 496-3122; E-mail: wk@wjh.harvard.edu).

This research was supported by National Institute on Aging grant AG08441, the Human Frontiers Science Program, and National Institutes of Mental Health grant R01 MH57681.

anoxia, encephalitis, Korsakoff's syndrome), show marked impairments in long-term memory for newly encountered facts and events, yet retain perceptual and linguistic functions within the normal range (Parkin & Leng, 1993; Squire, 1994). Importantly, these studies have revealed that, under some conditions, amnesics may prove to be *less* susceptible to illusory recollection than are individuals with intact memory, with their impairment in veridical memory in some way buffering them against the errors of illusory memory shown by control participants. In other instances, however, the reverse has been found: Under some conditions amnesics have shown *increased* rates of false recognition, proving to be more prone to such errors than are individuals with unimpaired memory.

One significant factor that may mediate which of these two patterns is observed—depressed versus elevated false recognition among amnesics relative to controls—is the extent to which the nonstudied lure items share semantic or conceptual features with previously encountered stimuli. Several recent studies have examined false recognition in amnesia under conditions where multiple items that are semantically or conceptually related to a particular theme word are presented (Schacter, Verfaellie, & Anes, 1997; Schacter, Verfaellie, Anes, & Racine, 1998b; Schacter, Verfaellie, & Pradere, 1996). A key observation from these studies is that under these conditions—where multiple items might serve to reinforce the conceptual or semantic gist of the list—amnesic patients show both impairment in correct recognition and *decreased* false recognition. For example, Schacter et al. (1996) compared the recognition performance of amnesics and controls in the converging verbal associates paradigm developed by Deese (1959) and Roediger and McDermott (1995). In this paradigm, participants are presented several lists of words, where each of the words in a given list is associatively related to a single (non presented) "critical lure" (e.g. participants might be asked to remember the words *bed*, *rest*, *awake*, *tired*, *dream*, etc., without being presented the word *sleep*, the critical lure to which each of these list words is associated). As expected, the average rate of veridical recognition of control participants (83%) substantially exceeded that of amnesics (46%); however, amnesics also showed impaired *false* recognition: Whereas controls falsely recognised 89% of the critical lures, amnesics' false recognition rate was 57%. Largely similar findings were reported by Schacter et al. (1998b), and by Schacter et al. (1997). In the latter study, using different word lists comprised of fewer semantic associates, amnesics showed impaired true recognition in two experiments and impaired false recognition in one experiment (Expt. 1) but not in another (Expt. 2); however, in the second experiment, control participants also showed particularly low levels of false recognition.

These observations of impaired false recognition in amnesics under conditions where multiple semantic associates are presented stand in marked contrast to an early report indicating that, compared to non-memory-impaired controls, amnesics might show *elevated* rates of false alarms under conditions where the lures are related to only one of the studied items (rather than multiple study items). In an early study, Cermak, Butters, and Gerrein (1973) found that, when only a single related item was presented in a continuous recognition paradigm, Korsakoff amnesics were more likely than alcoholic controls to show false recognition of new items that were similar to previously presented items. Compared to controls, Korsakoff amnesics showed significantly greater false recognition of associatively related words (e.g. study *table*, tested with *chair*), and also of homonyms (e.g. *bear–bare*); amnesics also showed a nonsignificant tendency to more often false alarm to synonyms (*robber–thief*) and to unrelated items than did controls. More recently, in the study by Schacter et al. (1996), amnesic patients demonstrated *both* a pattern of "impaired" false recognition of semantically related items for which many study items were presented, and increased baseline rates of false alarms for items that were unrelated to study items.

What might account for these differing patterns? Schacter et al. (1996) suggested that, when multiple related associates are presented, false recognition among the control participants depends on memory for the associative or semantic information that also supports veridical memory. They proposed that, because control participants were able to retain and integrate or "bind together" the studied items, they formed a strong and well-organised representation of the semantic gist of each set of associates—thereby rendering rejection of the semantically similar lures quite difficult, but also, and simultaneously, facilitating rejection of lures that were unrelated to any of the lists. In contrast, amnesic patients may have formed only a weak or degraded representation of the semantic gist of the list items. Such weak and less well-integrated representations would render amnesics less likely to show false recollection of the related lures, but also relatively less likely to reject unrelated new items successfully (which would not "fail to cohere" with their memory of the lists in the same manner that unrelated items for the control participants did). Based on these considerations, Schacter et al. predicted that, by systematically varying the number of related items that were presented, "it should be possible to specify a crossover point where amnesics shift from *enhanced* to *reduced* false recognitions" (Schacter et al. 1996, p. 329, emphasis added).

These outcomes each involved the false recognition of lures that were conceptually or semantically related to the studied items. Importantly, however, in contrast to these converging findings from studies involving similarity of meaning, the findings relating to the effects of *perceptual* similarity on false recognition of amnesics are notably less clear. Whereas some studies have suggested that amnesics may also show impaired false recollection when items share perceptual similarity (Schacter et al., 1997), other studies have not found a deficit (Kolodny, 1994) or—in instances where perceptual features of studied stimuli are recombined to form new items (Kroll, Knight, Metcalfe, Wolf, & Tulving, 1996, Expt. 2; also cf. Reinitz, Verfaellie, & Milberg, 1996)—have even reported significantly higher false recognition among amnesic individuals[1].

Evidence that perceptually based false recognition in amnesics also might be depressed compared to that found for control participants was reported by Schacter et al. (1997). In addition to conditions manipulating the conceptual similarity of studied items, these researchers included a condition in which they manipulated the perceptual similarity of words, exposing amnesic and control participants to some study lists where the words shared a high level of orthographic and phonemic similarity (e.g. for the related lure *fate*, the study list words were *fade*, *fame*, *face*, *fake*, *mate*, *hate*, *late*, *date*, and *rate*, Experiment 1) or were all rhyming words (Experiment 2). In both experiments, amnesics showed impaired true recognition for such perceptually similar items, and also impaired false recognition (after correcting for baseline levels of false alarms). In one experiment, the level of corrected false recognition of amnesics did not exceed chance; in the second, both amnesics and controls showed above-chance levels of corrected false recognition, but perceptual false recognition of amnesics fell significantly behind that of controls.

[1] The studies of Kroll et al. (1996) and Reinitz et al. (1996) are of particular interest in that, in agreement with many researchers, they underscore the notion that the hippocampal system may play a key role in allowing the "cohesion" or "binding" of different elements into a single trace (e.g. Moscovitch, 1994; Rudy & Sutherland, 1994; Schacter, Church, & Bolton, 1995; Shapiro & Olton, 1994; Sutherland & Rudy, 1989). For example, Reinitz et al. (1996) found that both amnesics and controls often falsely recognised compound words that were comprised of a conjunction of words embedded in previously studied items (e.g. study "*tooth*paste" and "heart*ache*;" test with the lure "*toothache*"); however, whereas controls showed substantially greater veridical recognition for actually studied items (that is, items that preserved both the elements and their relation with one another), amnesics showed largely similar levels of false recognition of recombined conjunction lures and veridical recognition of the target words. These researchers proposed that amnesics may have especial difficulty retaining information concerning *global stimulus structure*, so that although they can remember individual features or parts of items, they have marked deficits in tasks that require access to information about how the parts or features are interrelated to one another. Similar findings of especially heightened susceptibility to conjunction errors among amnesics were reported by Kroll et al. (1996), both for words comprised of recombined syllables (e.g. study "val*ley*" and "*bar*ter," test with "*barley*") and for schematic pictures of faces. Nonetheless, it is also important to note that, unlike the experiments of both Kroll et al. (1996) and Reinitz et al. (1996), in the present study, the lures were not created by recombining previously presented features but were comprised of new elements that were perceptually similar—but not identical—to presented items. Because of this difference, we focus on studies where perceptual similarity was manipulated without the recombination of previously studied elements.

Although suggestive, these findings are somewhat difficult to interpret because the manipulation of perceptual similarity involved familiar words, and thus may have invoked both semantic and perceptual processing. This is especially important as these findings conflict with the outcome of a study in which perceptual similarity was manipulated for abstract novel materials that did not possess prior semantic associations. Kolodny (1994) examined recognition performance of amnesics and controls for "random dot patterns"—patterns that were each comprised of nine simple black dots, arrayed in different statistically determined configurations (cf. Posner, Goldsmith, & Welton, 1967; Posner & Keele, 1968). Participants saw multiple exemplars derived from three different "prototype" patterns and then were given a yes/no recognition test for items that had been presented previously, as well as for new items derived from the prototypes and the prototypes themselves. False recognition of amnesics numerically exceeded that of controls, with false recognition of amnesics (69%) actually somewhat greater than true recognition (59%); in contrast, controls showed the reverse pattern (52% false recognition, 67% true recognition). A similar pattern—with amnesics showing numerically greater false recognition (73%) than controls (63%)—was observed for a subset of the new items that were prototypes or low-level distortions of the prototype, that is, items for which, all else being equal, perceptually based similarity- or gist-based responding might be expected to be most pronounced. Moreover, a second experiment reported by Kolodny (1994), involving pictures with representational content (e.g. paintings with religious or biblical themes), but where lures probably shared both conceptual and perceptual similarity with target items, again showed elevated false recognition among amnesics compared to controls.

Does false recognition derived from the perceptual similarity of lures with previously encountered target items differ in some way from that derived from shared semantic similarity? Direct comparisons of these studies are hampered by the many methodological differences across the paradigms. Also, because the Kolodny (1994) dot pattern study did not include an estimation of baseline levels of false recognition, comparisons with the results of Schacter et al. (1997) using perceptually similar words are difficult. Nonetheless, the differing patterns for these studies clearly suggest the need for further evidence regarding the influence of perceptual similarity on illusory recognition in amnesics relative to controls. Evidence is particularly needed under conditions where: (1) the stimuli are novel objects or patterns for which contributions from pre-existing semantic knowledge are minimised, and (2) factors that manipulate the degree to which general similarity might contribute to performance are used. The possibility that there might be a cross-over pattern in the level of false recognition shown by amnesics versus controls as a function of the degree of false recognition induced among controls is especially of interest because documenting such a pattern for novel perceptually related stimuli would provide evidence for the similarity of gist extraction mechanisms across perceptual and conceptual domains.

The experiment reported here was designed to address these issues. We examined false recognition errors arising from perceptual similarity in a paradigm where the stimuli were abstract novel objects: complex, multi-featured shapes constructed from a prototype. We employed two forms of manipulation that—based on findings from other experiments—are known to affect "general similarity" (Curran, Schacter, Norman, & Galluccio, 1997) or "gist based" (Brainerd, Reyna, & Kneer, 1995; Reyna & Brainerd, 1995) false recognition responding in controls. One manipulation was *category size*. Robust effects of category size have been found in several previous studies of false recognition in individuals with intact memory functioning, with higher rates of false recognition found following the presentation of increasing numbers of semantically related words (e.g. Arndt & Hirshman, 1998; Robinson & Roediger, 1997; Shiffrin, Huber, & Marinelli, 1995), categorised pictures of everyday objects (e.g. Koutstaal & Schacter, 1997; Koutstaal, Schacter, Galluccio, & Stofer, in press), and abstract patterns (e.g. Homa, Cross, Cornell, Goldman, & Schwartz, 1973; Omohundro, 1981). The second manipulation involved the degree of perceptual similarity of the lures: *transformational distance* from a category prototype. High rates of false recognition of prototype items have been reported among individuals with unimpaired memory (e.g. Franks & Bransford, 1971; Solso & Raynis, 1979), with false recognition rates decreasing with decreasing similarity to the prototype.

Based on the findings from Kolodny (1994), one might expect no impairment among amnesics compared to controls in false recognition of perceptually related items (at least for the uncorrected data), and equivalent or even enhanced false recognition of the prototypes. However, based on the findings of Schacter et al. (1997) using perceptually similar words, and also on the findings from studies that used multiple semantically related items (Schacter et al., 1996, 1997, 1998b)—all of which pointed to impaired gist-memory among amnesics—it might be expected that amnesics would show a deficit in false recognition relative to controls, particularly under conditions where there was an opportunity for a *strong* build-up of gist (for example, for categories where many related items were presented during the study phase, and for category prototypes). In contrast, under conditions promoting a less strong build-up of gist in control participants (for example, where only one or a few related exemplars had been presented, or for exemplars at a far distance from the category prototypes), one would not necessarily expect false recognition of amnesic individuals to fall behind that of control participants. When only a few categorically related exemplars had been presented, control participants would not have formed a robust gist-like representation of the category, and so should not be strongly inclined to show false recognition of lure items from the category because of their general similarity to studied items (cf. Homa et al., 1973; Koutstaal & Schacter, 1997; Koutstaal et al., in press; Omohundro, 1981; Robinson & Roediger, 1997). Further, when only a few categorically related items had been presented, controls may be more able to retain, and draw upon, more "item-specific" representations of the actually presented items so as to differentiate studied from nonstudied items and thus successfully "oppose" or "suppress" false recognition responses (cf. Schacter et al., 1998a, b). For these items, the combination of a less robust gist representation, and possibly greater retention of item specific information, among the control participants might result in similar levels of false recognition for amnesics and controls; indeed, as in the case of Cermak et al. (1973), for the smaller category sizes, false recognition among amnesics might even exceed that of controls.

METHOD

Participants

Twelve amnesic patients (9 male, 3 female) and 12 individuals with intact memory functioning (8 male, 4 female) took part in the experiment. Both amnesics and their controls were screened at the Memory Disorders Research Center of the Boston Veterans Affairs Medical Center. A subgroup of six amnesic patients were of mixed aetiology (anoxia, encephalitis, and thalamic infarct) and six had a diagnosis of alcoholic Korsakoff's syndrome. All patients were matched to a corresponding control participant on the basis of age, education, and verbal IQ (Wechsler Adult Intelligence Scale-Revised); the Korsakoff patients were matched to individuals with a history of alcoholism (alcoholic controls) whereas mixed aetiology patients were matched to controls with no history of alcoholism (non-alcoholic controls). Table 1 presents the characteristics of the individual amnesic patients, including their performance on the Wechsler Memory Scale-Revised and, where available, anatomical lesion information. Individual lesion data for the Korsakoff patients is not available; however, several of these patients took part in a larger magnetic resonance imaging study (Jernigan, Schafer, Butters, & Cermak, 1991) that used quantitative image-analytic methods to compare Korsakoff patients to age-matched alcoholic controls and age-matched normal controls. In this study, on a group level, and relative to a larger normative sample of 55 normal volunteers, both alcoholic controls and Korsakoff patients showed considerably reduced diencephalic grey matter volumes; however, the volume loss of Korsakoff patients was

Table 1. *Characteristics of Amnesic Patients*

				WMS-R[e]		
Patient: Aetiology	*Age*	*Ed*[c]	*Verbal IQ*[d]	*GM*	*Delay*	*Atn*
CW: Thalamic infarct: bilateral anterior nuclei[a]	56	12	87	79	80	89
AB: Encephalitis: mild diffuse cortical atrophy[b]	58	16	105	76	51	92
PD: Anoxia: enlarged ventricles, diffuse cortical atrophy[b]	60	20	109	65	61	89
JM: Anoxia: bilateral hippocampal complex[b]	48	12	89	70	52	92
PS: Anoxia: bilateral hippocampal complex[b]	39	14	95	90	50	115
RL: Anoxia	68	18	103	68	66	93
Mean	54.8	15.3	98.0	74.7	60.0	95.5
AA: Korsakoff	70	9	93	76	62	109
LB: Korsakoff	63	11	90	99	61	99
PB: Korsakoff	71	14	87	82	60	93
RD: Korsakoff	67	12	83	66	50	99
WR: Korsakoff	69	7	88	76	53	96
RM: Korsakoff	77	14	112	91	68	95
Mean	69.5	11.2	92.2	81.7	59.0	98.5

Anatomical lesion information is provided where available.
[a] Lesions assessed by CT.
[b] Lesions assessed by MRI.
[c] Ed = years of formal education.
[d] Verbal IQ from the Wechsler Adult Intelligence Scale (Revised).
[e] WMS-R = Wechsler Memory Scale-Revised; scores are presented separately for the indices of general memory (GM), delay, and attention (Atn). The WMS-R does not provide scores below 50, and 50 was the lowest score used to compute means.

significantly greater than that of alcoholic controls in anterior regions of the diencephalon; in addition, relative to alcoholic controls, Korsakoff patients showed significant reductions in grey matter in mesial temporal structures and orbitofrontal cortex.

The controls for the mixed amnesic subgroup had a mean age of 50.7 years, with an average of 14.2 years of education and a verbal IQ of 104.2. Alcoholic controls had a mean age of 67.5, an average of 12.8 years of education, and an average verbal IQ of 99.2.

Design

The experimental design included a between-subjects factor of group (amnesic vs. control) and two within-subjects factors: transformational distance and category size. Transformational distance had three levels for studied items (near, middle, far) and four levels for nonstudied items (prototype, near, middle, far). Category size had four levels for studied items: one, three, six, or nine related items presented at study (termed single, small, medium, and large categories respectively) and five levels for nonstudied items: the four levels above, plus "novel" category items for which no related items were presented at study; these latter items provided an estimate of baseline levels of false alarms. In addition, non-categorised items (termed "unrelated," see following) were included both as studied items and as new items.

Stimuli

The stimuli were colour depictions of complex, multi-featured, abstract objects, created using a computer graphics program (Aldus Freehand). Most of the stimuli were "categorised" items, generated by first creating a novel prototype according to a specific set of construction rules (described later), and then generating additional exemplars that belonged to the same category through manipulations that distorted the initial prototype in greater or lesser degree. Also, non-categorised or "unrelated" items were created that did not follow the rules of construction for the categorised items.

Prototypes for 18 different categories were first created according to a set of construction rules. Each prototype consisted of a large central form (the "main" component) together with three smaller features. All of the prototypes were two-dimensional and were created so as to form a single, unified object with multiple parts. The prototypes were created so as to be as dissimilar from one another as possible. Each category had a set of four unique features associated with it; these features were assigned in all possible groupings of three to create four "prototypes" per category, with the only difference between the prototypes for a given category being the features they contained. Specifically, using the letters *a*, *b*, *c*, and *d* to designate the four possible features of a category, the prototypes for a given category consisted of one—invariant—main component, with superimposed features of either *abc*, *abd*, *acd*, or *bcd*. The features associated with a given category were selected to possess some "pairwise" similarity (i.e. the features within a category comprised two pairs of two features, with the features in each pair being relatively similar to one another). The placement of features on the main unit was constant for a given feature, such that if the different prototypes of a category shared a given feature, it appeared in a similar place.

The initial prototypes were then manipulated in various prescribed ways, including alterations of shape, colour, outline, and size, to create additional exemplars that possessed varying degrees of similarity to the prototypes. These manipulations specified particular types and ranges of manipulations and were applied in an algorithmic fashion to items within each category. The placement of features on these exemplars was also altered, but within a specified range (e.g. a given feature would generally appear more towards the top or side of an object, but its precise placement varied somewhat).

Pilot research that examined yes/no recognition of the exemplars derived from these algorithmically applied alterations suggested that the items were too similar to one another: Participants were largely unable to discriminate studied from nonstudied exemplars. Accordingly, the stimuli were further altered to increase the within-category distinctiveness of the stimuli. This involved several further steps: first, each item in a category was further manipulated in a specified manner (e.g. one exemplar's shape was "bent" a certain amount, another exemplar's outline was increased in thickness, etc.); then, following these manipulations, three items within each category that were still relatively similar to one another were selected for additional alterations that rendered the three items less similar to one another and to other exemplars within the category. These latter alterations employed overall manipulations (e.g. rendering the entire shape into a somewhat "swirled" or "warped" shape) and, if necessary, were followed by yet a further manipulation in which particular portions of the shape were altered. Finally, because these further alteration procedures resulted in non systematic variations in the initial "distance metric" for the exemplars, the similarity of the new items to the prototypes was assessed by asking six raters to place the exemplars in each category on a distance metric corresponding to how similar they were to the prototype. The transformational distance of a given item was then determined by the average of the six raters' judgements, in conjunction with the further ratings of two of the experimenters (EJ and WK) in cases where the judges' ratings did not point to a clear outcome. In addition, the two exemplars that comprised the clearest representatives of the three distances (near, middle, and far) were chosen to comprise the critical stimulus items (to be counterbalanced across study and test status; see following).

Examples of the categorised items are provided in Fig 1. Shown in the upper two panels are

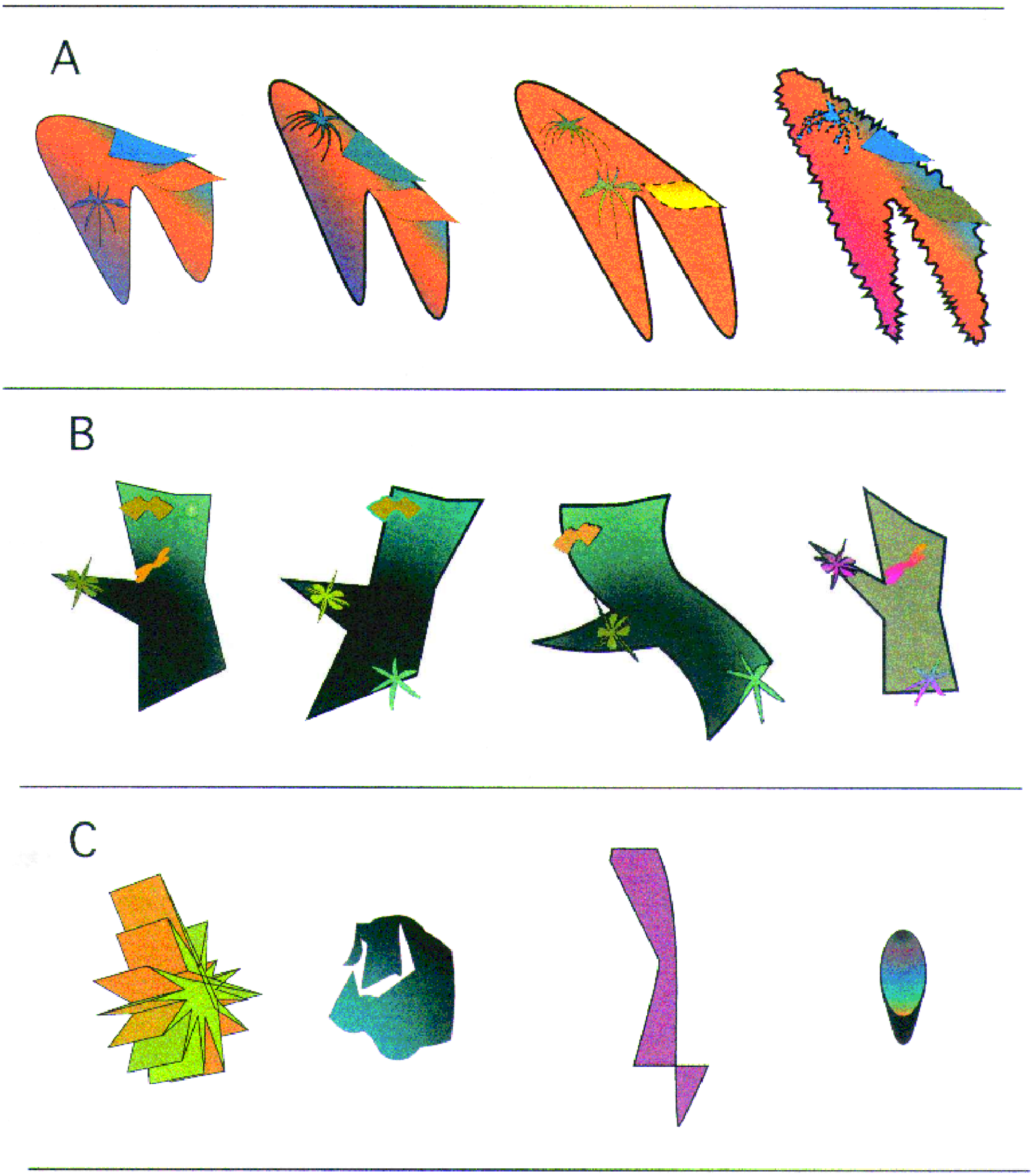

Fig. 1. *Examples of categorised stimuli (A, B) and unrelated items (C). The upper two panels show examples from two different categories including, from left to right, a category prototype, and items from the near, middle, and far transformational distances respectively. Note that, although, for illustrative purposes here, the stimuli are shown in black and white, the stimuli as shown to participants were presented in colour, with colour comprising an important attribute that was varied both within and across categories.*

exemplars from two categories including (from left to right) a prototype, and items from the near, middle, and far distances. In addition, examples of unrelated items are shown in the bottom panel; unrelated items were never comprised of four components and could be either two- or three-dimensional figures.

For each category, the critical exemplars of each distance were randomly assigned to one of two sets (A or B). These sets were subsequently used for counterbalancing the critical exemplars across studied and nonstudied status. In order to avoid confounding the number of related exemplars that were presented at *study* with the number of items that were *tested*, only a subset of all the items in each category was tested: for the three-, six-, and nine-item categories, three old items (one from each distance) and four new items (one from each distance, plus the prototype) were tested; for categories where only one item was presented, only this single target item (always a middle distance item), and two new items (one from the middle distance, plus the prototype) were tested. Category size at study was manipulated by systematically excluding some of the noncritical items depending on the category size to be used: no noncritical items were excluded for nine-item categories, but a given set of three noncritical items (one at each distance) was excluded for six-item categories, and six noncritical items (two at each distance) were excluded for three-item categories. For categories where only one item was presented at study, only one middle distance item from the critical items was presented (from Set A or Set B, depending on the counterbalancing condition).

Templates for the study and test lists were created such that, within each third of the list, the number of items from each category type, category size (including singles, and unrelated items), and transformational distance (near, middle, far) were balanced; in addition, for the test lists, the number of novel items and prototypes was also balanced across the test thirds. Finally, additional templates were created employing different orderings of the items within each of the test thirds.

For counterbalancing purposes, the 18 categories were assigned to 6 stimulus subsets comprised of 3 categories each; these subsets were used to determine, across subjects, whether a given category was shown as studied items or as novel items, with 1 stimulus subset assigned to each of the large, medium, small, and novel category-size conditions (each of which were thus represented by 3 categories), and 2 stimulus subsets assigned to the single condition (thus represented by 6 categories). Counterbalancing across subjects ensured that each category occurred in each of the 6 conditions once, with each category represented once with the "A" and "B" exemplars as the studied, and nonstudied, exemplars. Thus a counterbalancing required 12 subjects. Amnesics and their matched controls were assigned the same counterbalancing lists.

Procedure

The overall procedure involved three phases, including a study phase, in which participants were exposed to the stimuli under an incidental encoding task, a brief retention interval, and the test phase. All participants were tested individually, in a single session of approximately 40 to 50 minutes. The stimuli at both study and test were presented via a powerbook computer with a colour monitor, using the PsyScope experimental presentation program. The stimuli appeared in the centre of the screen, with prompts for responding to the encoding task or the recognition test displayed beneath.

In the study phase, participants were shown a total of 78 items (72 critical items, preceded and followed by 3 buffer items; critical items = 27 large, 18 medium, 9 small, 6 single, and 12 unrelated items). Each item was presented for 6 seconds, and participants were asked to rate the "overall complexity" of the stimulus on a 9-point scale, where "1" indicated the stimulus was not at all complex and "9" indicated it was extremely complex. In making their complexity ratings, participants were instructed to consider all aspects of the stimulus, including both different dimensions of the stimulus (e.g. shape, colour, size, and outline) and all of the components of the stimulus. Participants entered their complexity ratings on the computer keyboard, using the number keys 1–9. The study items were preceded by a brief practice session that included items from two further categories not employed in the experiment and additional unrelated items.

The study phase was followed by a 5-minute retention interval, during which participants performed an unrelated number search task. In the test phase, participants were shown a subset of the items shown in the study phase, together with new items, and were asked to designate each item as "old" (previously presented) or "new" (never previously presented). The test list included 117 items, of which 45 were "old" and 72 were "new." The old items consisted of 3 items from each of the studied categories, with the exception of the single categories, where only the single studied item was presented, plus the 12 unrelated items (i.e. old items = 3 × 3 large category items, 3 × 3 medium category items, 3 × 3 small category items, 6 × 1 single category items, and 12 unrelated items). The new items were comprised of 3 related lure items from each of the studied categories (or, for the single categories, 1 new item), together with one prototype from each category; in addition there were 3 items plus the prototype from each of 3 novel categories, and 12 new unrelated items.

RESULTS

All analyses are first presented combining across the two amnesic subgroups (i.e. combining across the mixed aetiology and Korsakoff subgroups), with results first considered as a function of transformational distance, and then as a function of category size. Thereafter, the two subgroups are briefly considered separately. The Appendix presents the central true and false recognition outcomes for individual patients in both the mixed aetiology and Korsakoff groups. The results for the unrelated items are presented first. To increase the normality of the distributions, and to make "equal differences equally detectable" (Rosenthal & Rosnow, 1991; Snedecor & Cochran, 1989), all analyses of proportions were performed after arcsine transformation.

Combined Groups

Unrelated Items

We first considered performance for the unrelated items, that is, items that were markedly different from the categorised items and that did not follow the basic rules of construction for category items (see Table 2). For these items, amnesics showed substantially impaired veridical recognition: means of 65% vs. 43% for controls and amnesics respectively, $F(1,22) = 5.02$, $MSe = 0.09$, $P = .04$. False alarms to unrelated items were numerically, but not significantly, more frequent for amnesics (.24) than controls (.18), $F < 1$. After correcting for these (nonsignificant) baseline differences in false alarms to unrelated lures, veridical recognition of control participants also significantly exceeded that of amnesics, $F(1,22) = 10.52$, $MSe = 0.05$, $P = .004$.

Analyses by Distance

Table 3 presents the proportion of "old" responses given to studied items (true recognition) and nonstudied items (false recognition) for amnesics and controls, with responses shown separately as a function of distance from the prototype (near, middle, far). Also shown are the proportion of "old" responses to the prototypes (averaged for study categories comprised of three, six, and nine items), and the baseline rate of false alarms to novel items. (Note that the results for "single" items, for which only one categorised item was presented at study and which was always from the "middle" distance, are not considered here but are considered in the subsequent section considering the effects of category size.)

Table 2. *Recognition Responses for Unrelated (Noncategorised) Items*

	Hits	*False Alarms*	*Hits − False Alarms*
Amnesic	.43	.24	.20
Korsakoff	.40	.14	.26
Mixed	.46	.33	.13
Control	.65	.18	.47
Alcoholic	.56	.18	.38
Non-alcoholic	.74	.18	.56

Table 3. *True and False Recognition Responses by Transformational Distance*

	True Recognition				*False Recognition*					
	Near	*Middle*	*Far*	*Mean*	*Near*	*Middle*	*Far*	*Mean*	*Prototype*	*Novel*
Amnesic	.60	.59	.56	.58	.58	.50	.39	.49	.52	.35
Korsakoff	.65	.69	.56	.63	.56	.54	.50	.53	.61	.35
Mixed	.56	.50	.56	.54	.61	.46	.28	.45	.43	.35
Control	.80	.80	.62	.74	.69	.62	.41	.57	.72	.38
Alcoholic	.85	.82	.69	.78	.72	.65	.46	.61	.72	.47
Non-alcoholic	.74	.78	.56	.69	.65	.59	.35	.53	.72	.28

Items for categories where only one item was shown at study are omitted because for these categories all exemplars were from the middle distance. The prototype means shown are also based only on categories where three, six, or nine exemplars were shown. Novel indicates the baseline level of false alarms to "novel" category exemplars (0 exemplars shown at study).

From Table 3 it can be seen that, whereas true recognition of amnesics was largely unaffected by the distance of studied exemplars from the prototype (means of .60, .59, and .56 for near, middle, and far respectively), true recognition of the controls was more strongly affected by distance. Control participants showed a high level of correct recognition for near and middle distance items (.80) but considerably reduced recognition for items that were far from the prototype (.62). Table 3 also shows that false recognition for both amnesics and controls declines with increasing distance, but with controls showing higher false recognition for near and middle distance items (means for near, middle, and far respectively, for amnesics of .58, .50, and .39, compared to .69, .62, and .41 for controls). Importantly, from Table 3 it can also be seen that the group differences in false recognition are most pronounced for the prototypes. Relative to the controls, amnesics show a false recognition impairment of 20% for the prototype items compared to an impairment of 11%, 12%, and 2% for lures from the near, middle, and far distances respectively. Amnesics and controls showed moderately high and very similar baseline rates of false alarms: "old" responses to novel category items of .35 and .38 respectively. False alarms to these novel category items—that is, items that followed the same general rules as all of the categorised items but were not from the categories that were presented—provide a measure of participants' general willingness to respond "old" under conditions of uncertainty, but where the stimuli share broad similarity with the types of items that were studied.

The true and false recognition scores following correction for differences in baseline levels of false alarms (Hits − Novel Category False Alarms, and Related False Alarms − Novel Category False Alarms, or "novel-corrected" scores) are shown in Table 4. Focusing first on the true recognition results, an initial 2 (Group: amnesic or control) × 3 (Distance: near, middle, far) analysis of variance (ANOVA) on the novel-corrected true recognition scores showed a significant effect of Distance, $F(2,44) = 4.36$, $MSe = 0.03$, $P = .02$, but no effect of Group, $F = 2.1$, and no Group × Distance interaction, $F < 1.6$. Pairwise comparisons showed that, averaging across groups, novel-corrected true recognition for near and middle distance items exceeded that of far distance items (smallest $F = 6.66$), but near and middle distance items did not differ from one another ($F < 1$). Combining the near and middle distance items and comparing novel-corrected true recognition for these items with that for the far distance items, and also including Group as a factor, again showed a main effect of Distance, $F(1,22) = 10.49$, $MSe = 0.02$, $P = .004$; however, in this more focused analysis, the differentially enhanced performance of the controls for the near and middle distance items was more strongly apparent: $F(1,22) = 4.08$, $MSe = 0.02$, $P = .06$ for the Group × Distance interaction.

Next considering false recognition, a 2 (Group) × 4 (Distance: prototype, near, middle,

Table 4. *Novel-corrected True and False Recognition Responses by Transformational Distance*

	True Recognition				*False Recognition*				
	Near	*Middle*	*Far*	*Mean*	*Near*	*Middle*	*Far*	*Mean*	*Prototype*
Amnesic	.26*	.25*	.21*	.24*	.24*	.15*	.04	.14*	.17*
Korsakoff	.30*	.34*	.21*	.28*	.21*	.19*	.15	.18*	.26*
Mixed	.21*	.15	.21[a]	.19[a]	.26*	.12	−.07	.10	.08
Control	.42*	.42*	.25*	.36*	.31*	.25*	.03	.20*	.35*
Alcoholic	.38*	.34*	.21[a]	.31*	.25*	.18[a]	−.01	.14[a]	.25*
Non-alcoholic	.46*	.50*	.28*	.41*	.37*	.32*	.07	.25*	.45*

Items for categories where only one item was shown at study are omitted because for these categories all exemplars were from the middle distance. The prototype means shown are also based only on categories where three, six, or nine exemplars were shown.

* $P < .05$ for comparisons against zero.

[a] Marginal effects: $.05 < P < .10$ for comparisons against zero.

far) ANOVA performed on the novel-corrected false recognition scores revealed a main effect of Distance, $F(3,66) = 15.34$, $MSe = 0.02$, $P < .0001$, no effect of group, $F < 1.5$, and a slight trend toward a Group × Distance interaction, $F(3,66) = 2.04$, $MSe = 0.02$, $P = .12$. Pairwise comparisons showed that, except for the prototype and near items, novel-corrected false recognition decreased regularly with increasing distance from the prototype (prototype = near > middle > far; for the near, middle, and far comparisons, smallest $F = 5.77$). Critically, however, although there was no overall effect of group, an analysis restricted to the novel-corrected false recognition scores for the *prototypes* alone revealed significantly impaired false recognition among amnesics relative to controls, $F(1,22) = 4.36$, $MSe = 0.05$, $P = .05$.

Analyses by Category Size

Table 5 presents the proportion of "old" responses for the amnesic and control groups separately as a function of category size (one, three, six, or nine related exemplars shown at study) for both studied items ("true recognition") and nonstudied items ("false recognition"). Also shown are the proportion of old responses to novel items (i.e. the baseline measure of false alarms). In order to reduce the number of comparisons, and to increase the power of individual comparisons, we combined the two relatively smaller category sizes (one and three related items presented) and the two relatively larger categories (six and nine related items presented), contrasting the two larger and two smaller categories.

We first considered veridical recognition as a function of category size after correcting for novel false alarms (see Table 6). For novel-corrected true recognition, a 2 (Group) × 2 (Category Size: larger vs. smaller) ANOVA revealed no overall effect of Group, $F = 1.9$, an effect of Category Size, $F(1,22) = 12.52$, $MSe = 0.01$, $P = .002$, and a Group × Category Size interaction, $F(1,22) = 4.71$, $MSe = 0.01$, $P = .04$. Overall, novel-corrected veridical recognition was greater for larger (.32) than smaller (.23) categories, but this difference was largely carried by control participants (.24 vs. .20 for amnesics compared to .41 vs. .25 for controls).

We next conducted a similar 2 (Group) × 2 (Category Size: larger vs. smaller) analysis on the novel-corrected false recognition scores. This analysis showed a main effect of Category Size, $F(1,22) = 14.51$, $MSe = 0.02$, $P = .001$, with no effect of Group, $F < 1$, and no Category Size × Group interaction, $F = 2.0$. False recognition was greater for lures from many-exemplar categories (.21) than for lures from few-exemplar categories (.07), with both amnesics and controls showing this pattern (for amnesics, .17 vs. .08, difference of 9%; for controls, .26 vs. .06, difference of 20%).

Table 5. *True and False Recognition Responses by Category Size*

	True Recognition					*False Recognition*									
	Category Items					*Category Items*					*Prototype*				
	9	*6*	*3*	*1*	*Mean*	*9*	*6*	*3*	*1*	*Mean*	*9*	*6*	*3*	*1*	*Novel*
Amnesic	.62	.56	.57	.53	.57	.55	.48	.44	.40	.47	.56	.44	.56	.29	.35
Korsakoff	.76	.57	.56	.56	.61	.67	.50	.43	.42	.50	.72	.61	.50	.25	.35
Mixed	.48	.54	.59	.50	.53	.43	.46	.46	.39	.44	.39	.28	.61	.33	.35
Control	.81	.76	.65	.60	.70	.68	.58	.45	.42	.53	.86	.70	.61	.32	.38
Alcoholic	.83	.82	.70	.64	.75	.65	.63	.56	.50	.58	.89	.61	.67	.36	.47
Non-alcoholic	.78	.70	.59	.56	.66	.70	.54	.35	.33	.48	.83	.78	.56	.28	.28

False recognition responses to the prototype are also shown separately as a function of category size (nine, six, three, or one item shown at study). Also shown are false alarms to "novel" items.

We also considered novel-corrected false recognition of the prototypes separately as a function of the number of exemplars that were studied. A 2 (Group) × 2 (Category Size: larger vs. smaller) ANOVA revealed a marginal effect of Group, with controls tending to show greater overall false recognition of the prototypes on this measure than amnesics, $F(1,22) = 3.34$, $MSe = 0.08$, $P = .08$. (Note that, because this analysis includes single categories, it differs somewhat from the analyses reported earlier for transformational distance, which excluded the single item categories, and where amnesics were significantly impaired on this measure.) There was also a main effect of Category Size, $F(1,22) = 10.51$, $MSe = 0.05$, $P = .004$, and a trend toward a Group × Category Size interaction, $F(1,22) = 4.10$, $MSe = 0.05$, $P = .06$. Comparing the level of novel-corrected false recognition of the prototypes for the amnesics versus controls within each category size separately showed that amnesics were significantly impaired in false recognition of prototypes from large categories, $F(1,22) = 4.92$, $MSe = 0.12$, $P = .04$, marginally impaired in false recognition of prototypes from medium categories, $F(1,22) = 2.75$, $MSe = 0.15$, $P = .11$, and not impaired for prototypes from small or single-item categories, Fs < 1.

Table 6. *Novel-corrected True and False Recognition Responses by Category Size*

	True Recognition					*False Recognition*									
	Category Items					*Category Items*					*Prototype*				
	9	*6*	*3*	*1*	*Mean*	*9*	*6*	*3*	*1*	*Mean*	*9*	*6*	*3*	*1*	*Mean*
Amnesic	.27*	.21*	.23*	.18*	.22*	.20*	.13*	.10	.06	.12*	.21[a]	.10	.21*	−.06	.12*
Korsakoff	.41*	.23[a]	.21[a]	.21[a]	.26*	.32*	.15	.08	.07	.16*	.38*	.26[a]	.15	−.10	.17*
Mixed	.13	.19	.25*	.15	.18[a]	.08	.12	.12	.04	.09	.04	−.07	.26*	−.01	.06
Control	.43*	.38*	.27*	.22*	.33*	.30*	.21*	.08	.04	.16*	.49*	.32*	.24*	−.06	.25*
Alcoholic	.36*	.34*	.23	.17	.28*	.18*	.16	.08	.03	.11	.42*	.14	.20*	−.11	.16[a]
Non-alcoholic	.50*	.43*	.32*	.28*	.38*	.43*	.26[a]	.07	.06	.20*	.56*	.50*	.28*	1.3E-4	.33*

Results are shown for true and false recognition following correction for baseline levels of false alarms to novel category items.
* $P < .05$ for comparisons against zero.
[a] Marginal effects: .05 < P < 10 for comparisons against zero.

Subgroup Analyses

From the means shown in Table 3 and Table 5 it can be seen that, in absolute terms, the Korsakoff patients generally showed higher levels of both true and false recognition than did mixed amnesics (average true recognition for all four category sizes for Korsakoff vs. mixed amnesics respectively of .61 vs. .53, average false recognition of .50 vs. .44). Korsakoff patients also showed greater false recognition of the prototypes than did mixed amnesics, with this effect especially pronounced for items for which nine or six related exemplars had been studied (Korsakoff patients, means of .72 and .61; mixed amnesics, means of .39 and .28 respectively). However, a largely similar pattern was found for the two subgroups' respective *control groups*, with the alcoholic controls on average showing numerically higher true recognition than did the mixed aetiology controls (means for all four category sizes of .75 vs. .66 respectively) and also higher false recognition (means of .58 vs. .48). An exception, however, concerned unrelated items (see Table 2), where alcoholic controls showed relatively less true recognition than did mixed controls (.56 and.74 respectively).

Because the alcoholic controls also showed an especially high rate of false alarms to novel items (.47 compared to .35 for the Korsakoff amnesics), after correction for baseline levels of false alarms, there were relatively small and unsystematic differences in corrected true and false recognition for the Korsakoff group compared to their controls (e.g. for all four category sizes, overall novel-corrected true recognition of .26 vs. .28 for Korsakoff patients and alcoholic controls respectively; overall novel-corrected false recognition of .16 vs. .11; likewise, for false recognition of the prototypes, novel-corrected false recognition averages of .26 vs. .25 respectively). In contrast, mixed amnesics and their controls showed a less pronounced difference in baseline rates of false alarms to novel categorised items (.35 and .28 respectively).

Focusing on the performance of the mixed amnesia subgroup, Table 3 and Table 5 show that the recognition performance of the mixed amnesic patients was quite consistent with the expectation that, under conditions of relatively little or weak gist, amnesics might show levels of false recognition similar to, or exceeding, that shown by controls. False recognition of the mixed amnesics numerically exceeded that shown by controls for *novel items* (.35 vs. .28), for *single-item category lures* (.39 vs. .33) and for *three-item category lures* (.46 vs. .35). However, also as expected, under conditions of relatively greater build-up of gist, this pattern reversed, with the control participants showing higher levels of false recognition than the mixed amnesics for larger category sizes, where either six or nine items had been presented (for six-item categories, false recognition of .46 vs. .54 for mixed amnesics versus their controls respectively; for nine-item categories, .43 vs. .70). This difference was also observed for the prototypes, where the false recognition rate of mixed amnesics fell nearly 30% behind their controls (combining across categories of three, six, and nine items, means of .43 vs. .72). Indeed, within the prototypes themselves, a similar "crossover" in performance as a function of category size (and presumably "gist") was observed, with mixed amnesics falling well behind their controls in false recognition of the prototypes for the six- and nine-item categories (.39 vs. .83, and .28 vs. .78 respectively) but with the pattern reversed for the three-item and single-item categories (.61 vs. .56, and .33 vs. .28, respectively).

For the novel-corrected data, subgroup analyses revealed no significant differences for the Korsakoff amnesics compared to their controls. For the mixed amnesics, significant or nearly significant differences in novel-corrected true recognition were found for near-distance items [means of .21 vs. .46, $F(1,10) = 4.98$, $MSe = 0.05$, $P = .05$]; for middle-distance items [means of .15 vs. .50, $F(1,10) = 4.64$, $MSe = 0.09$, $P = .06$]; items from large categories [means of .13 vs. .50, $F(1,10) = 9.64$, $MSe = 0.05$, $P = .01$]; and for unrelated items [means of .13 vs. .56, $F(1,10) = 13.01$, $MSe = 0.05$, $P = .005$]. For novel-corrected false recognition, significant or nearly significant differences were found for middle-distance items [means of .12 vs. .32, $F(1, 10) = 3.66$, $MSe = 0.04$, $P = .08$]; for items from large categories [means of .08 vs. .43, $F(1,10) = 5.27$, $MSe = 0.08$, $P = .04$]; prototypes for large categories [means of .04 vs. .56, $F(1,10) = 6.55$, $MSe = 0.16$, $P = .03$]; prototypes for medium categories [means of − .07 vs. .50, $F(1,10) = 10.09$, $MSe = 0.12$, $P = .01$]; and for the prototypes from large, medium, and small categories combined [means of .08 vs. .45, $F(1,10) = 7.31$, $MSe = 0.06$, $P = .02$]. The pattern of depressed false recognition for the larger categories and for prototypes was quite consistent across patients (see individual patient data in the Appendix): for the larger category sizes, and for the near/middle transformational distances, five of the six mixed aetiology patients showed false recognition scores that fell below the mean of the controls; for the prototypes, all of the mixed aetiology patients showed lower false recognition rates than the mean for their controls.

GENERAL DISCUSSION

A central finding of this experiment is that amnesics may show *either* substantial impairment of perceptually based false recognition, or virtually no impairment, depending on the nature of the experimental conditions and the extent to which they allow the extraction of gist information among control participants. Under some conditions—particularly those that allow maximal strengthening of gist-based representations in controls—amnesics showed a substantial impairment of perceptually based false recognition for pictures of novel abstract objects for which they had no pre-existing semantic or conceptual knowledge. However, under conditions where control participants showed relatively less robust false recognition, false recognition of amnesics was also less impaired. The deficit in amnesics' false recognition was especially apparent for the prototypes. For the control participants, the prototypes elicited the highest levels of false recognition that were observed for any of the experimental conditions, with rates of false recognition that matched, or for some conditions exceeded, those seen for true recognition. Yet for these items, amnesic patients showed a substantial and reliable deficit in false recognition.

Although this pattern was observed both in the overall group analysis (associated with a false recognition deficit of 18% after taking baseline false alarms into account), and in the analysis for the mixed aetiology subgroup alone (associated with a novel-corrected false recognition deficit of 37%), the unusually high baseline level of false alarms among the alcoholic control group largely eliminated the pattern observed for the uncorrected or absolute scores in the subgroup of Korsakoff patients. Thus the conclusions for this subgroup are less clear. Ignoring the baseline differences, Korsakoff and mixed aetiology patients generally showed not dissimilar patterns for transformational distance. Korsakoff patients showed comparatively greater false recognition "deficits" for the prototype and near and middle distance items (deficits of .11, .16, and .11) than for far distance items (deficit of − .04); mixed aetiology patients did not show a very marked deficit for near items (.04) but showed strong deficits for middle distance items (.13) and an extremely marked deficit for the prototypes (.29) with little deficit for far distance items (.07). For the factor of category size, however, mixed aetiology amnesics showed more systematic "gist-related" deficits than did Korsakoff patients.

The nature of this "gist-related" pattern of deficits in mixed amnesics is perhaps most clearly seen by comparing the differences in false recognition of mixed amnesics versus their controls for the various categories. These differences were positive, showing numerically higher false recognition for amnesics than for controls, for each of five categories involving few items: novel items, single items, single item prototypes, small category items, and small category prototypes (differences of + 7%, + 6%, + 5%, + 11% and + 5% respectively); in marked contrast, for medium and large category items, and medium and large category prototypes, the differences were all negative and, in several instances, of substantial magnitude (− 8%, − 27%, − 50%, and − 44%), now in each case showing higher false recognition for controls than for amnesics or "impaired false recognition" for amnesics. A focused comparison, contrasting the performance of mixed amnesics versus their controls for the former set of items, drawn from "low-gist" conditions, against those for the latter set of items drawn from "high-gist" conditions, yielded no effect of Group, $F < 1$, a main effect of Gist (high gist = .55, low gist = .39), $F(1,10) = 9.15$, $MSe = 0.02$, $P = .01$, and a highly significant Group × Gist interaction, $F(1,10) = 14.40$, $MSe = 0.02$, $P = .004$.

This interaction provides clear evidence for a "crossover" in the false recognition performance of amnesics versus control participants as a function of the number of categorically related items that had been encountered previously. Moreover, consistent with the proposal made earlier by Schacter et al. (1997), the across-group difference tracks with the *magnitude of the false recognition effect in the control group*, with mixed aetiology amnesics showing the greatest deficits under conditions where gist-based responding among their normal controls is most robust, namely for the large categories (.70) and medium categories (.54) and especially for the prototypes from these categories (.83 and .78 respectively). Mixed amnesics showed no such deficit, however, for the small or single item categories, where gist-based responding among controls was also much lower (.35 and .33 respectively).

What might account for these differential patterns, with amnesics showing little or no impairment under some conditions (few or far transformational distance exemplars) and substantial impairment under others (many or near transformational distance exemplars)? A pattern of this form might emerge if amnesics showed uniformly little effect of the manipulations of category size or transformational distance but controls were sensitive to these manipulations—acquiring an increasingly strong sense of familiarity with the categories as more exemplars were encountered, and more often incorrectly succumbing to this sense of familiarity if a lure item was similar to many previously encountered items, or if it was similar to (or actually was) the most representative or "prototypical" instance of the category. On this account, a differential pattern of impairment across the different conditions would emerge because, in the control participants, general familiarity with the categories was relatively weak in some conditions (and thus matched that of amnesics) but continued to accumulate in other conditions (and thus exceeded that of amnesics).

This simple account, however, does not appear to provide a sufficiently modulated portrayal of the data. Although, overall, the performance of amnesic patients was less affected by the factors of transformational distance and category size than was found for control participants, amnesics were not entirely immune to the influence of perceptual gist or general similarity information. Combining across the two amnesic subgroups, amnesics' overall level of false recognition for the category lures (.47) significantly exceeded their baseline level of false alarms to novel category items (.35), $F(1,11) = 7.37$, $MSe = 0.01$, $P = .02$. Further, at least in numerical terms, amnesics showed some "gist-like systematicity" in their false recognition responses to categorised lures, showing numerically greater false recognition of categorised lure items that were closer in similarity to the prototypes (means of .58, .50, and .39 for near, middle, and far distance items respectively) and also for lures for which a larger number of related items had been presented at study (.55, .48, .44, and .40 for study categories comprised of nine, six, three, and one items respectively). This result suggests that amnesics were able to retain at least some "category-specific" information regarding the stimulus items, and that this information—like that for controls—was influenced by the factors of transformational distance and category size, albeit to a lesser degree.

The obtained pattern of results, therefore, can be better understood in terms of a differential build-up of gist information in the amnesics compared to the controls. If category specific information was extracted and/or retained less efficiently by amnesics than by control participants, this difference might be least evident in cases where controls also were able to construct only relatively weak gist representations (as when few category exemplars were presented), but might emerge more strongly as the gist representations of controls became increasingly robust. Given the relatively low levels of within-category discriminability shown by both amnesics (on average, approximately 10%) and controls (less than 20% for the categorised items, excluding the prototypes), it is likely that this category-specific information was of a relatively global and "coarse" form, most often involving general familiarity with the nature of "category-like" items rather than specific or richly detailed recollection.

This account of the differences in false recognition between amnesics and controls is based on the notion that an impairment in gist extraction might selectively affect performance for large category sizes and near transformational distance exemplars. The same pattern of results, however, might emerge if control participants had stronger gist representations for the smaller category sizes or far transformational distance exemplars, but if controls—unlike amnesics—were able to *counteract* false recognition responses successfully for these items (cf. Schacter et al., 1998b). That is, it is possible that individuals in the control group actually had a stronger gist-like representation than did amnesics for the smaller categories (as they apparently did for the larger categories), but they were able to oppose or suppress gist-based responding for these items by relying on more detailed, item-specific memory (which might, for example, suffer from less interference when few rather than many overlapping exemplars had been studied, cf. Schacter et al., 1998a).

The overall pattern of performance by the control participants argues against a strong contribution from this factor. For a "selective suppression" account of this form to be viable, it would be expected that control participants would demonstrate a high level of item-specific memory in those conditions in which suppression was occurring. Such "item-specific" memory might be shown in various ways, but one measure—readily obtained from the present data—is participants' ability to discriminate studied items from similar but not studied lures (i.e. hits minus related false alarms). Overall, controls showed somewhat greater ability than did amnesics to discriminate studied targets from related category lures. Yet, contrary to what would be expected on the basis of a selective suppression account, amnesics and controls showed numerically *less* pronounced differences in such within-category discriminability for the far distance items (True Target − Related False Target means of .17 for amnesics compared to .21 for controls) than for middle distance items (means of .09 vs. .18) or near distance items (means of .02 vs. .11). Similarly, for the factor of category size, the differences in within category discriminability for amnesics versus controls were not noticeably more pronounced for the smaller categories (means of .13 and .19 for amnesics and controls) than for the larger categories (means of .07 vs. .15 for amnesics and controls).

One qualification concerning the differential patterns of false recognition for amnesics versus controls that we observed concerns the levels of false recognition that emerged for lures at a far transformational distance from the prototype, and for lures from smaller categories, where participants were exposed to only one or three target items per category. For these conditions, neither amnesics nor controls showed rates of false recognition significantly above those found for novel category lures (on average, for both combined groups, false recognition of far distance items was approximately 4% greater than for novel category lures and between 6% and 8% for the smaller categories; also see Table 4 and Table 6, where asterisks indicate the conditions where novel-corrected true recognition and novel-corrected false recognition exceeded chance levels). On the one hand, this pattern is of concern because, where false recognition rates are at floor or at near-floor levels, it would be difficult to detect a difference between the amnesic and control groups. On the other hand, and while acknowledging this as a legitimate concern, it might also be noted that the absolute levels of false recognition for the far distance lures and smaller category lures were not inconsiderable, ranging between .28 and .50 for the four subgroups for far distance items, and between .33 and .56 for the smaller category sizes. Further, the near-equivalence of false recognition rates in these cases to the false alarm rate observed for lures from novel categories—that is, items that followed the same general rules of construction as the studied categorised items, false recognition rates of .28 to .47—may reflect a real functional equivalence: The overall levels of general familiarity of novel category lures would not be expected to differ radically from that for lures from categories where only one or three related exemplars had been presented. Still, it is possible that, especially under conditions involving lower baseline levels of false recognition, some differences in the false recognition rates of amnesics versus controls might be observed for these conditions.

These considerations—which have thus far been confined to categorised items—have emphasised the parallels in the memory processes shown by amnesics and controls, with amnesics showing similar but less efficient gist extraction than that shown by individuals with intact memory. However, the emergence of clear group differences for the *non-categorised* or unrelated items, where the memory performance of controls far exceeded that of amnesics, suggests that, in addition to building up coarse or global forms of category information, control participants were also able to encode and retain more detailed, item-specific information. Evidence for this latter form of more detailed or specific recollective information—and the need to differentiate it from other forms of information—especially derives from a comparison of participants' corrected true recognition for the *unrelated* items versus their recognition of items for which a *single* categorised item was presented. These two conditions were similar in that, for both conditions, only one item of a given sort was presented at study; however, whereas the single items followed the rules of construction for the categorised items (and were themselves category items), the unrelated items did not follow those rules and were noticeably different from the categorised items. Whereas amnesic and control participants showed essentially identical novel-corrected veridical recognition for the single categorised items (. 18 and .22 respectively), corrected veridical recognition of the two groups diverged markedly for the unrelated items. For the unrelated items, amnesics achieved a very similar score to that seen for single items ("unrelated-corrected" mean of .20) but control participants—now presumably responding on the basis of detailed item-specific information—showed significantly higher levels of recognition (.47). An analysis contrasting corrected veridical recognition of the single versus unrelated items for the two groups yielded a significant interaction of group with item type, $F(1,22) = 5.62$, $MSe = 0.03$, $P = .03$, indicating reliably differential patterns of responding for these item types for amnesics versus controls.

The above pattern of findings, with controls showing greater item-specific memory for non categorised items, and also—in the case of the prototypes—greater gist-based false recognition, raises a question regarding the extent to which item-specific memory and gist memory might covary. Is decreased gist-based false recognition an inevitable consequence of decreased item-specific memory? Although generalisations across different paradigms and different populations need to be made very cautiously, it might be noted that, contrary to this suggestion, across a number of studies we have observed several different relations between indices that probably primarily reflect item-specific memory (veridical recognition of single or unrelated "one-of-a-kind" items) and gist-based false memory (false recognition of lures that are perceptually and/or conceptually similar to target items). For example, in other experiments using the same abstract novel objects as in the present experiment but comparing older and younger adults, we observed largely *equivalent* levels of gist-based false recognition in older and younger adults even though, like the control participants in the present study, younger adults demonstrated greater item-specific memory (as indexed by performance for unrelated items) than did older adults. In a paradigm involving categorised pictures of everyday objects, we found that older adults show reliably *greater* gist-based false recognition than younger adults together with *decreased* item-specific memory, as measured by their recognition of unrelated items, or items from categories where only a single related item was presented (Koutstaal & Schacter, 1997; Koutstaal et al., in press). Further, matching older and younger adults on their level of item-specific memory, by requiring younger participants to perform a secondary task during the initial encoding of the pictures (Koutstaal, Schacter, & Brenner, 1998), did not eliminate the age-difference in gist-based responding: Under these conditions, the two age groups showed similar levels of item-specific memory but older adults continued to show greater gist-based false recognition. These varied patterns of gist-based false recognition in relation to item-specific memory for older versus younger adults suggest that the two forms of responding may be determined by multiple factors and, though they may sometimes covary (as they did here), at other times they may not. (For more general discussion and findings concerning the bases underlying true and false recognition, and the conditions under which they may or may not covary, see Arndt & Hirshman, 1998). These multiple patterns also make it hazardous to predict whether, if the amnesic patients and control participants were led to have similar levels of item-specific memory (by, for example, testing control participants after a longer study-to-test interval), they would show false recognition similar to that of the amnesics. Although an increased study-to-test interval might decrease item-specific memory in the control participants, it is possible that false recognition of categorised items, particularly category prototypes, would remain relatively preserved (cf. McDermott, 1996; Payne, Elie, Blackwell, & Neuschatz, 1996; Reyna & Brainerd, 1995), with controls still more often falsely recognising these items than amnesics.

Theoretical Implications

Overall, the outcomes of the present experiment are consistent with the predictions based on prior studies using multiple conceptually or perceptually related words where, under conditions where controls showed high levels of illusory recognition, amnesics also showed substantial deficits in false recognition. The present demonstration of a significant impairment in perceptually based false recognition of prototypes among amnesics using novel abstract visual stimuli extends these earlier findings, showing that impairments in amnesic patients may be observed in cases where the stimuli have no pre-existing semantic associations and where the extraction of similarities depends on representational and associational processes that occur entirely within the experimental session itself.

The parallel patterns found for both perceptually similar items (this study; Schacter et al., 1997) and for semantically or conceptually similar items (Schacter et al., 1996, 1997, 1998b), with amnesics showing "impaired" false recognition under conditions where individuals with intact memory show high levels of false recognition but not under conditions where cognitively intact individuals show less pronounced levels of false recognition, suggest that some of the processes underlying false recognition may be similar across these broad types of similarity. In particular, both forms of false recognition appear, at least in part, to derive from—and also to reflect—limits on the "exact" or "verbatim" encoding, retention, and/or retrieval of information. In both domains, increasing convergence or overlap of the stimulus items, whether in terms of semantic or associative attributes (e.g. Robinson & Roediger, 1997; Vogt & Kimble, 1973) or perceptual features (e.g. Franks & Bransford, 1971; Solso & Raynis, 1979), may act to increase the likelihood of false recognition of nonstudied items that possess similar characteristics.

Nonetheless, these findings do not imply that the processes are entirely parallel across the perceptual and conceptual domains. Although we have underscored the congruencies in the processes relating to these broad types of similarity in regard to certain aspects, particularly in relation to the number of related items that are presented, and their "distance" from a central prototype, other aspects of the processing of perceptually and conceptually similar materials are quite likely to differ. Thus, depending on the particular paradigm used, specific factors may contribute differently to false recognition (or other forms of memory errors, such as false recall) in the conceptual versus perceptual domains. For example, with semantically or associatively related items, such as the converging associates used in the Deese/Roediger-McDermott paradigm where all of the items converge upon a single critical lure, errors may arise from a form of inferential and elaborative processing at the time of encoding and/or retrieval in a way that differs from that for novel perceptually similar items, where pre-existing associative and thematic information is less likely to be strongly evoked. Further, whereas the convergence of items in semantic paradigms (during either study or test) may lead to explicit phenomenological awareness of the nonstudied lure because this item is generated or automatically activated in response to the multiple semantically related items (cf. Gallo, Roberts, & Seamon, 1997; McDermott & Roediger, 1998; Read, 1996; Roediger & McDermott, 1995), the convergence of visually similar items, particularly novel patterns or shapes, is unlikely to result in the "generation" of a "lure" in quite the same way: Although a given exemplar might remind an individual of earlier presented similar target items, participants would appear to be unlikely to generate "novel" variants or exemplars of the targets spontaneously (in part because pre-existing associational support for such generation is lacking). This suggests that direct "source confusions" between items that are presented and items that are implicitly or consciously generated during study (did I hear or read this word versus only think it?) may be more likely to arise for semantically or associatively related words than for perceptually similar items.

Although a number of studies have focused on the effects of *either* perceptual similarity or conceptual similarity on false recognition, the relations between perceptual and conceptual similarity, including the possibility that they may have additive effects, have been much less explored. In the paradigm using categorised pictures of everyday objects that we have used to explore false recognition in older and younger adults (Koutstaal & Schacter, 1997; Koutstaal et al., in press), the objects within a particular category (e.g. cats, chairs, or teddy bears) may share both considerable perceptual similarity (general shape, colour, etc.) and also conceptual (semantic and lexical) information; however, in this paradigm, perceptual and conceptual similarity are confounded and the relative roles of each in false recognition can be assessed only indirectly, by considering the consequences of shifting attention to the different components, or similar manipulations. More direct evidence for the possible additive effects of perceptual and conceptual similarity has been provided by Henkel and Franklin (1998). These researchers found higher rates of source monitoring errors in a judgement task requiring participants to differentiate between visually presented versus imagined items if the items both physically resembled one another and belonged to the same functional category (e.g *bagel–doughnut*) than if items shared primarily only conceptual similarity (e.g. *pants–shirt*) or primarily only physical similarity (e.g. *magnifying glass–lollipop*). The possible effects of combining both perceptual and conceptual similarity on false recognition performance, including false recognition of amnesic patients versus controls, have not been explored.

Although our finding that amnesics showed significantly depressed false recognition for prototypes based on perceptual similarity is consistent with the results from previous studies of false recognition in amnesia using conceptually related words (Schacter et al., 1996, 1997, 1998b), or orthographically and phonologically similar words (Schacter et al., 1997), this outcome clearly appears to be inconsistent with the findings reported by Kolodny (1994), particularly those of the random-dot pattern study, where amnesics showed false recognition rates that exceeded those of controls, even for items that were prototypes or low distortion exemplars of the categories. As noted previously, the absence of baseline rates in that study makes comparisons difficult, and it is possible that—after correction—the results from the two studies would be less discrepant. In addition, the studies differed in many respects. One factor, in particular, that recent research has suggested may have been important is the number of times that the target items were presented: Whereas target items were presented only once in the present study, targets were presented repeatedly in the Kolodny study. This latter factor is important because repeated presentation of the targets may have allowed control participants—at least to some extent—to *suppress* false recognition of the lures. Although, within the current experiment, we found little evidence of "suppression-like" processes, other experiments using the semantic converging associates paradigm have demonstrated suppression among younger persons with intact memory when the target items have been presented repeatedly (Kensinger & Schacter, this issue; McDermott, 1996)—possibly because such repetition facilitates the encoding and retention of greater amounts of item-specific information, or otherwise encourages more careful monitoring of recognition responding. Importantly, in a study comparing the effects of repeated study presentation on the false recognition of amnesics versus controls, Schacter et al. (1998b) found that whereas repeated presentations of semantic associates lead to decreasing false recognition in controls, amnesics showed either increasing rates of false recognition across trials (Korsakoff patients) or fluctuating levels of false recognition (amnesic patients of mixed aetiology). These findings raise the possibility that the repeated target presentations in the study by Kolodny may have acted both to decrease false recognition among controls (albeit not too successfully given that, in absolute terms, the rates of false recognition were not inconsiderable) and either to exacerbate, or at least to maintain, false recognition in amnesic individuals.

From a broader theoretical perspective, our results showing impaired gist-based false recognition of prototypes in amnesics bear upon two more general questions about the nature of preserved versus impaired memory functions in amnesics. One issue concerns the relation between *categorisation or classification tasks*—that is, tasks involving the acquisition and transfer of knowledge concerning the classification of novel stimuli, including the ability to correctly assign novel items to classes based on experience with other category members—and *recognition memory*, involving the explicit recollection of stimuli or events. Previously reported evidence showing a dissociation in the performance of amnesics relative to control participants on these two types of task, with amnesics showing impaired recognition memory but intact or near-normal levels of categorisation learning (Knowlton & Squire, 1993; Kolodny, 1994; Squire & Knowlton, 1995; also cf. Knowlton, Mangels, & Squire, 1996; Knowlton, Ramus, & Squire, 1992; Knowlton, Squire, & Gluck, 1994), has provoked considerable discussion and research concerning the factors underlying this dissociation, especially whether it is necessary to invoke separate memory systems to account for the difference, or if other accounts (e.g. Nosofsky, 1988; Nosofsky & Zaki, 1998) are possible. Our finding of impaired gist-based false recognition of prototypes in amnesics appears to provide additional evidence for the distinction between the processes or structures that support categorisation performance, and those that underlie recognition (in this case, illusory, rather than veridical recognition). Indeed, a growing set of evidence suggests that, in contrast to intact classification performance, including intact classification of perceptually based prototypes (Knowlton & Squire, 1993; Kolodny, 1994; Squire & Knowlton, 1995), amnesic patients show impaired false recognition of "prototypical" but not-presented target items, with such deficits found for semantically similar words (Schacter et al., 1996, 1997, 1998b), perceptually similar words (Schacter et al., 1997), and—in the present study—perceptually similar novel abstract patterns. On the one hand, it is not known if amnesics would show normal or near-normal classification performance for complex abstract stimulus patterns of the form used here, as has been found for the random dot patterns (Knowlton & Squire, 1993; Kolodny, 1994; Squire & Knowlton, 1995) and artificial grammar strings (Knowlton et al., 1992; Knowlton & Squire, 1994) used by previous investigators. On the other hand, the findings of the current study, taken together with those of the several earlier studies employing semantically and perceptually similar words, suggest that different processes, and possibly different representations, may support "gist-based false recognition"—manifested during a task requiring the making of explicit recognition decisions upon the basis of an episodic query of the past—and categorisation performance, requiring the designation of items as belonging or not belonging to a specific category, without any direct or necessary involvement of recollection. Although the factors underlying the dissociation between amnesics' recognition versus classification performance are likely to be a subject of continuing investigation, a clear implication of the present findings is that differences in the performance of amnesics and control participants as manifested in illusory or *false* recognition also comprise one part of the overall data set for which a satisfactory account must be offered.

Finally, and also in a broader perspective, it might be noted that our results provide evidence against early proposals suggesting that, whereas conceptual or semantic processing is impaired in amnesia, perceptual processing—not only for implicit tests, but also for explicit tests—might be *unimpaired* in amnesia (Blaxton, 1989, 1995). We found that, under certain conditions, perceptually based false recognition was clearly and significantly impaired in amnesics, thereby demonstrating that the processes involved in the explicit false recollection of abstract perceptual patterns—like that of conceptual information—may also be impaired in amnesics. These findings are thus consistent with our earlier findings, using perceptually related words (Schacter et al., 1997), and with other recent reports indicating that explicit memory of amnesics is impaired, not only for tests that emphasise conceptual processing, such as free or cued recall, but also on tests that draw more extensively upon perceptual processes, such as graphemic cued recall (Cermak, Verfaellie, & Chase, 1995) and word fragment cued recall (Vaidya, Gabrieli, Keane, & Monti, 1995). That, in the present study, this perceptual memory impairment should assume the apparently paradoxical guise of *fewer* false recognition errors in amnesic individuals than in control participants shows, yet again, the versatility and generality of the strategy of using false recollection to uncover the nature of the retrieval and encoding processes operating in veridical memory.

REFERENCES

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.

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, V.F., & Kneer, R. (1995). False recognition reversal: When similarity is distinctive. *Journal of Memory and Language, 34*, 157–185.

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

Cermak, L.S., Verfaellie, M., & Chase, K.A. (1995). Implicit and explicit memory in amnesia: An analysis of data-driven and conceptually driven processes. *Neuropsychology, 9*, 281–290.

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.

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

Estes, W.K. (1997). Processes of memory loss, recovery, and distortion. *Psychological Review, 104*, 148–169.

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

Gallo, D.A., Roberts, M.J., & Seamon, J.G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin and Review, 4*, 271–276.

Henkel, L.A., & Franklin, N. (1998). Reality monitoring of physically similar and conceptually related objects. *Memory and Cognition, 26*, 659–673.

Homa, D., Cross, J., Cornell, D., Goldman, D., & Schwartz, S. (1973). Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. *Journal of Experimental Psychology, 101*, 116–122.

Jernigan, T.L., Schafer, K., Butters, N., & Cermak, L.S. (1991). Magnetic resonance imaging of alcoholic Korsakoff patients. *Neuropsychopharmacology, 4*, 175–186.

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

Kensinger, E.A., & Schacter, D.L. (this issue). When true memories suppress false memories: Effects of ageing. *Cognitive Neuropsychology, 16*, 399–415.

Knowlton, B.J., Mangels, J.A., & Squire, L.R. (1996). A neostriatal habit learning system in humans. *Science, 273*, 1399–1402.

Knowlton, B.J., Ramus, S.J., & Squire, L.R. (1992). Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological Science, 3*, 172–179.

Knowlton, B.J., & Squire, L.R. (1993). The learning of categories: Parallel brain systems for item memory and category knowledge. *Science, 262*, 1747–1749.

Knowlton, B.J., & Squire, L.R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 79–91.

Knowlton, B.J., Squire, L.R., & Gluck, M.A. (1994). Probabilistic classification learning in amnesia. *Learning and Memory, 1*, 106–120.

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., Galluccio, L., & Stofer, K.A. (in press). Reducing gist-based false recognition in older adults: Encoding and retrieval manipulations. *Psychology and Aging*.

Koutstaal, W., Schacter, D.L., & Brenner, C. (1998). *Dual task demands and gist based false recognition of pictures in younger and older adults*. Manuscript submitted for publication.

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

McClelland, J.L. (1995). Constructive memory and memory distortions: A parallel distributed processing approach. 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. 66–90). Cambridge, MA: Harvard University Press.

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

McDermott, K.B., & Roediger, H.L. III (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory and Language, 39*, 508–520.

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. 269–310). Cambridge, MA: MIT Press.

Nosofsky, R.M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 700–708.

Nosofsky, R.M., & Zaki, S.R. (1998). Dissociations between categorisation and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science, 9*, 247–255.

Omohundro, J. (1981). Recognition vs. classification of ill-defined category exemplars. *Memory and Cognition, 9*, 324–331.

Parkin, A.J., & Leng, N.R.C. (1993). *Neuropsychology of the amnesic syndrome*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Payne, D.G., Elie, C.J., Blackwell, J.M., & Neuschatz, J.S. (1996). Memory illusions: Recalling, recognising and recollecting events that never occurred. *Journal of Memory and Language, 35*, 261–285.

Posner, M.I., Goldsmith, R., & Welton, K.E. (1967). Perceived distance and the classification of distorted patterns. *Journal of Experimental Psychology, 73*, 28–38.

Posner, M.I., & Keele, S.W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77*, 353–363.

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.

Reinitz, M.T., Verfaellie, M., & Milberg, W. (1996). Memory conjunction errors in normal and amnesic subjects. *Journal of Memory and Language, 35*, 286–299.

Reyna, V.F., & Brainerd, C.J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 71*, 1–75.

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

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.

Rosenthal, R., & Rosnow, R.L. (1991). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.

Rudy, J.W., & Sutherland, R.J. (1994). The memory-coherence problem, configural associations, and the hippocampal system. In D.L. Schacter & E.

Tulving (Eds.), *Memory systems 1994* (pp. 119–146). Cambridge, MA: MIT Press.

Schacter, D.L. (1995). Memory distortion: History and current status. 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. 1–43). Cambridge, MA: Harvard University Press.

Schacter, D.L. (1997). False recognition and the brain. *Current Directions in Psychological Science, 6*, 65–70.

Schacter, D.L., Church, B., & Bolton, E. (1995). Implicit memory in amnesic patients: Impairment of voice-specific priming. *Psychological Science, 6*, 20–25.

Schacter, D.L., Norman, K.A., & Koutstaal, W. (1998a). The cognitive neuroscience 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. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language, 35,* 319–334.

Shapiro, M.L., & Olton, D.S. (1994). Hippocampal function and interference. In D.L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 87–117). Cambridge, MA: MIT Press.

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*, 261–287.

Snedecor, G.W., & Cochran, W.G. (1989). *Statistical methods* (8th ed.). Ames, IA: Iowa State University Press.

Solso, R.L., & Raynis, S.A. (1979). Prototype formation from imaged, kinesthetically, and visually presented geometric figures. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 701–712.

Squire, L.R. (1994). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. In D.L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 203–231). Cambridge, MA: MIT Press.

Squire, L.R., & Knowlton, B.J. (1995). Learning about categories in the absence of memory. *Proceedings of the National Academy of Sciences. USA, 92*, 12470–12474.

Sutherland, R.J., & Rudy, J.W. (1989). Configural association theory: The role of the hippocampal formation in learning, memory, and amnesia. *Psychobiology, 17*, 129–144.

Vaidya, C.J., Gabrieli, J.D.E., Keane, M.M., & Monti, L.A. (1995). Perceptual and conceptual memory processes in global amnesia. *Neuropsychology, 9*, 580–591.

Vogt, J., & Kimble, G.A. (1973). False recognition as a function of associative proximity. *Journal of Experimental Psychology, 99*, 143–145.

APPENDIX

Novel-Corrected True and False Recognition Data for Individual Patients

	True Recognition				*False Recognition*					
	Category Size		*Distance*		*Category Size*		*Distance*			
Patient: Aetiology	*Larger*	*Smaller*	*Near/ Middle*	*Far*	*Larger*	*Smaller*	*Proto-type*	*Near/ Middle*	*Far*	*Novel*
CW: Thalamic infarct: bilateral anterior nuclei[a]	.42	.39	.47	.25	.14	.11	.25	.25	–.08	.42
AB: Encephalitis: mild diffuse cortical atrophy[b]	.00	.14	.06	.17	–.11	–.08	–.17	–.06	–.28	.50
PD: Anoxia: enlarged ventricles, cortical atrophy[b]	.50	.53	.45	.67	.39	.47	.33	.56	.44	.00
JM: Anoxia: bilateral hippocampal complex[b]	.03	.08	.08	.14	.25	.00	.14	.25	–.08	.42
PS: Anoxia: bilateral hippocampal complex[b]	.03	.00	–.03	.03	–.14	–.03	–.08	.08	–.42	.75
RL: Anoxia	.00	.06	.06	.00	.06	.00	.00	.06	.00	.00
Patient Mean	.16	.20	.18	.21	.10	.08	.08	.19	–.07	.35
Control Mean	.46	.30	.48	.28	.34	.07	.45	.34	.07	.28
AA: Korsakoff	.33	.19	.28	.00	.28	.20	.22	.22	.33	.33
LB: Korsakoff	.39	.45	.39	.39	.22	.33	.39	.17	.28	.50
PB: Korsakoff	.11	.14	.11	.28	.06	–.11	.28	.11	–.17	.17
RD: Korsakoff	.03	.03	.03	–.03	.25	.00	.20	.14	.31	.58
WR: Korsakoff	.61	.28	.72	.33	.33	.11	.22	.39	.11	.00
RM: Korsakoff	.45	.17	.39	.28	.28	–.08	.28	.17	.06	.50
Patient Mean	.32	.21	.32	.21	.24	.07	.26	.20	.15	.35
Control Mean	.35	.20	.36	.21	.17	.06	.25	.21	–.01	.47

Means for prototypes are based on items from category sizes of three, six, and nine items; larger = average of six- and nine-item categories; smaller = average of one- and three-item categories. False alarms for novel category items are shown at the far right; all scores shown here have been corrected for these baseline differences in false alarms. Anatomical lesion information is provided where available.

[a] Lesions assessed by CT.

[b] Lesions assessed by MRI.