Neuropsychology
2001, Vol. 15, No. 2, 230–243

0894-4105/01/$5.00 DOI: 10.1037//0894-4105.15.2.230

Perceptual False Recognition in Alzheimer's Disease

Andrew E. Budson
Brigham and Women's Hospital and
Harvard Medical School

Rahul Desikan
Brigham and Women's Hospital

Kirk R. Daffner
Brigham and Women's Hospital and
Harvard Medical School

Daniel L. Schacter
Harvard University

Previous research has found that patients with probable Alzheimer's disease (AD) show lower levels of false recognition of semantic associates than do healthy older adults. To investigate whether this finding is attributable to semantic impairments in patients with AD, the authors examined false recognition of perceptually related novel objects with little semantic content in patients with AD and healthy older adults. By using corrected recognition scores to control for unrelated false alarms, it was found that patients with AD showed lower levels of both true and false recognition of novel objects than did older adults. These results suggest that the previous difference in false recognition of semantic associates observed between patients with AD and older adults is not entirely attributable to semantic memory deficits in patients with AD but may also involve poorly developed gist information in these patients.

In addition to failing to retrieve desired information, patients with probable Alzheimer's disease (AD) also suffer from distortions of memory (Förstl et al., 1994). These memory distortions may impair the ability of patients with probable AD to live independently (Borson & Raskind, 1997). For example, patients may believe that they turned off the stove or took their medication when they only thought about performing these activities. Memory distortions in AD are thus a clinically important issue; however, the causes of such distortions remain largely unexplored.

Although much of the previous research on memory distortion in AD has focused on the tendency to produce nonstudied items or "intrusions," patients with AD have recently been examined by using a paradigm that allows measurement of a similar type of memory distortion known as false recognition. False recognition occurs when people incorrectly claim to have previously encountered a novel word or event. False recognition has been studied more extensively and analytically than have recall intrusions (see Schacter, Norman, & Koutstaal, 1998) and, therefore, may allow insights into memory distortion in patients with AD that would be difficult to obtain from studies of intrusion errors. Recent experiments using a paradigm originally developed by Deese (1959) and revived and modified by Roediger and McDermott (1995) have demonstrated robust levels of false recognition in healthy adults. After studying lists of semantic associates (e.g., *candy, sour, sugar, bitter, good, taste,* and so forth) that all converge on a nonpresented "theme word" or "related lure" (e.g., *sweet*), participants frequently intruded the related lure on free-recall tests (Deese, 1959) and made very high levels of false alarms to these words on recognition tests (Roediger & McDermott, 1995).

We and others have previously shown that with the Deese/Roediger–McDermott (DRM) paradigm, false recognition of semantically associated words is significantly lower in patients with AD than in healthy older adults (after controlling for false alarms to unrelated words; Balota, Cortese, et al., 1999; Budson, Daffner, Desikan, & Schacter, 2000). (Note that in Balota, Cortese, et al.'s [1999] study, the recognition data were measured only after recall performance and therefore were contaminated by the earlier recall task.) Similar results have been obtained in patients with amnesia with the DRM semantic associates (Schacter, Verfaellie, Anes, & Racine, 1998; Schacter, Verfaellie, & Pradere, 1996), perceptually similar words (Schacter, Verfaellie, & Anes, 1997), and novel abstract objects (Koutstaal, Schacter, Verfaellie, Brenner, & Jackson, 1999). Patients with amnesia exhibit severe difficulties remembering recent experiences as a consequence of damage to the medial

Andrew E. Budson and Kirk R. Daffner, Division of Cognitive and Behavioral Neurology, Brigham and Women's Hospital, Boston, Massachusetts, and Harvard Medical School; Rahul Desikan, Division of Cognitive and Behavioral Neurology, Brigham and Women's Hospital; Daniel L. Schacter, Department of Psychology, Harvard University.

This research was supported by National Institute on Aging Grant AG08441, National Institute of Mental Health Grants F32 MH11767 and K23 MH01870, National Institute of Neurological Disease and Stroke Grant NS26980, and the Human Frontiers Science Program.

We thank Wilma Koutstaal, Carolyn Brenner, Carrie Racine, Dorene Rentz, and Leonard Scinto for their invaluable help and support.

Correspondence concerning this article should be addressed to Andrew E. Budson, Division of Cognitive and Behavioral Neurology, Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115. Electronic mail may be sent to abudson@partners.org.

temporal lobes and related structures in the diencephalon, despite normal perceptual and linguistic functions along with IQ scores in the normal range (e.g., Parkin & Leng, 1993; Squire, 1994). Thus, results from patients with AD may be entirely explained by their poor episodic memory, as is the case for patients with amnesia.

In addition to impairments in episodic memory, however, patients with AD are known to exhibit deficits in performance on some tasks that use semantic memory (van der Hurk & Hodges, 1995). Semantic memory consists of an organized body of knowledge of words, concepts, meanings, and associations (Nebes, 1989). It may be that the lower level of false recognition of semantically associated words that is seen in patients with AD is attributable primarily to their deficit in semantic memory performance. Supporting this idea, Dalla Barba and Wong (1995) demonstrated that patients with AD who showed worse performance on semantic memory tasks produced fewer related intrusions on free-recall tests.

The exact etiology of the semantic memory deficit seen in patients with AD is unclear (for a recent overview, see Ober, 1999). The poor performance of patients with AD on semantic memory tasks may be due to degraded semantic representations (Martin, 1992) or to deficits in retrieval of information, including failure in accessing, evaluating, and applying that information (Nebes, 1992). Supporting the theory of degraded semantic representations, investigators have found that patients with AD show greater difficulty in generating words from semantic (i.e., animals, fruits, vegetables), relative to phonemic (i.e., words beginning with the letters *F*, *A*, and *S*), categories (Monsch et al., 1992; Salmon, Heindel, & Lange, 1999). Supporting the theory of deficits in retrieval of information, investigators have found robust priming in tasks that require intact semantic representations (Balota & Duchek, 1991; Balota, Watson, Duchek, & Ferraro, 1999; Nebes & Halligan, 1996, 1999). Other experimental paradigms have also been used to support the theory of either degraded (e.g., Bayles, Tomoeda, & Cruz, 1999; van der Hurk & Hodges, 1995) or intact (e.g., Ober & Shenaut, 1999) semantic representations.

There are several ways in which the semantic memory dysfunction of patients with AD, regardless of its exact etiology, could contribute to these patients' lower levels of false recognition relative to healthy older adults. One explanation is that patients with AD, with their poor semantic memory, may not generate the implicit associative responses that healthy participants do when DRM study lists are presented. Deese (1959) and others (e.g., Bousfield, Whitmarsh, & Danick, 1958; Underwood, 1965) have suggested that the high levels of recall intrusions in the original paradigm may have been due to participants themselves spontaneously generating nonpresented critical lure words (e.g., *sweet*) during the study phase. That is, even though the critical theme words are not themselves presented, these words may come to mind during the study phase. Source memory confusion between these implicit associative responses and actually presented items could produce both recall intrusions and false recognition in healthy participants. Because patients with AD are less able to recognize the semantic associations between the studied items, they may be less likely to form an implicit associative response from a series of semantically related words. They would, therefore, be at less risk of confusing such an associative response with actually studied items.

Another way that the semantic memory deficits of patients with AD may explain why they exhibit lower levels of false recognition than control groups is that patients with AD may be less able to use semantic associations among target words in a strategic manner. In this way, the semantic memory deficits of patients with AD may work synergistically with their episodic memory deficits to reduce both true and false recognition. Simon, Leach, Winocur, and Moscovitch (1994) found that on the California Verbal Learning Test, patients with AD showed deficits in their ability to cluster words by taxonomic category. If, because of their semantic memory impairments, patients with AD are less able to recognize the semantic relationships between the words and cluster them together, they will then be less able to use a clustering or categorization strategy to remember the words. On recognition testing in the DRM paradigm, this categorization strategy would tend to support true recognition of previously studied words as well as false recognition of related lures.

Perhaps the most straightforward reason why the semantic memory deficits of patients with AD may contribute to their reduced level of false recognition of semantic associates is that if patients with AD do not recognize the semantic associations between the studied items, then they will not develop the general meaning, idea, or semantic gist conveyed by the collection of semantically related items (gist information; e.g., Reyna & Brainerd, 1995). It may be that in healthy older adults, as the study list is presented in the DRM paradigm, a gist representation is developed. This gist representation may result in an experience of recollection or familiarity when either a studied item or a related lure is presented on a later recognition test. In the DRM paradigm, accurate recognition of previously studied items probably depends on both gist information and the specific details of a prior encounter (item-specific recollection), whereas false recognition of related lure words may be related to remembering gist but not item-specific information (cf. Brainerd & Reyna, 1998a; Payne, Elie, Blackwell, & Neuschatz, 1996; Schacter et al., 1996). Whereas older adults recognize the semantic associations between related items, build up gist, and therefore become much more susceptible to responding "old" to the critical lures relative to the unrelated lure words, patients with AD would show much less selectivity between these two different types of false-positive responses.

To investigate whether the differences in false recognition between patients with AD and healthy older adults are entirely attributable to semantic deficits in AD, we examined false recognition of perceptually related novel objects with little semantic content in patients with AD and matched older adults. We used a paradigm recently used by Koutstaal, Schacter, Verfaellie, et al. (1999). The stimuli consisted of complex, perceptually detailed, abstract objects based on a category prototype (cf. Posner & Keele, 1968,

1970). Each object possessed varying levels of perceptual similarity to the prototype, defined as near, middle, or far transformational distance from the prototype. Thus, as well as providing comparisons between overall levels of perceptually based false recognition between the groups, the manipulation of the degree of perceptual similarity allowed for evaluation of the extent to which false recognition of patients with AD and healthy older adults depends on the perceptual similarity of the items to the category prototype. The number of categorically related items that were presented (one, three, six, or nine items) was also manipulated. Higher levels of false recognition have been demonstrated with increased numbers of categorically related items in studies examining semantically related words (e.g., Arndt & Hirshman, 1998; Robinson & Roediger, 1997; Shiffrin, Huber, & Marinelli, 1995), abstract patterns (e.g., Homa, Cross, Cornell, Goldman, & Schwartz, 1973; Omohundro, 1981), and pictures of everyday objects (Koutstaal & Schacter, 1997; Koutstaal, Schacter, Galluccio, & Stofer, 1999). This manipulation allowed us to determine whether a similar effect would be seen in patients with AD.

In addition to being able to eliminate explicit semantic content from the target stimuli, this paradigm provided two other advantages. First, we used a complexity judgment as an incidental encoding task that would tend to reduce the categorization memory strategy: Because participants were not told that they needed to remember the images for a later recognition test, it was less likely that they would develop strategies to do so. Second, using abstract novel objects should have removed the potential source memory confusion produced by implicit associative responses. In this paradigm, the lures were highly distinctive and detailed abstract colored images, and it was exceedingly unlikely that the participants themselves would have generated the specific images during the study phase of the experiment. Rather, false recognition of lures from the studied categories was likely driven by the perceptual correspondence between the lures and common properties of the studied exemplars, that is, perceptual gist information, analogous to the semantic gist discussed above. Thus, this perceptual paradigm should have provided fewer confounded measures of gist memory than was possible with semantically associated words.

Because episodic memory deficits, relative to semantic memory deficits, predominate in patients with mild to moderate AD (Hodges & Patterson, 1995), and because Koutstaal, Schacter, Verfaellie, et al. (1999) found that patients with amnesia showed reduced false recognition relative to controls (after correction for false alarms to unrelated novel items), we predicted that, as with lists of semantic associates, patients with AD would demonstrate lower levels of corrected false recognition of perceptually related novel objects than would healthy older adults. We expected that, because of their episodic memory deficits, patients with AD would exhibit more difficulty than healthy older adults in remembering the presented items; thus, we expected that their gist memory for those items would not be as developed, thereby reducing their corrected false recognition. We further expected that these differences would be larger either when older adults were able to build up more robust gist memory, as when there are large numbers of categorized items presented at study, or when older adults were most susceptible to the effects of gist memory, as when the lure items are closer in transformational distance to the category prototype. Because higher levels of gist memory should also enhance older adults' recognition accuracy, and because older adults are better able to use item-specific recollection (Budson et al., 2000) that would also increase their recognition of studied items, we similarly expected that patients with AD would show reduced levels of true recognition (after correction for false alarms to unrelated novel items) relative to older adults. Finally, we also expected the greatest differences in corrected true recognition between these groups with larger numbers of categorized items presented at study or with items closer to the category prototype, for the same reasons discussed above.

Method

Participants

Twelve patients with a clinical diagnosis of probable AD (according to the criteria of the National Institute of Neurological and Communications Disorders and Stroke and Alzheimer's Disease and Related Disorders Association; McKhann, Drachman, Folstein, Katzman, & Price, 1984) and 15 healthy older adults were recruited for the experiment. Patients with AD were recruited from the clinical population at the Memory Disorders Unit, Brigham and Women's Hospital, Boston, Massachusetts. Healthy older adults were recruited from individuals who were participating in a longitudinal study of normal aging at Brigham and Women's Hospital, as well as spouses and friends (but not blood relatives) of the patients with AD. Written informed consent was obtained from all participants and their caregivers (when appropriate). The study was approved by the human subjects committee of Brigham and Women's Hospital. Participants were paid $10/hr for their participation. Older adults were excluded from the study if they scored below two standard deviations on any element of the Word List Memory Test of the Consortium to Establish a Registry for Alzheimer's Disease (memory, recall, and recognition; Morris et al., 1989; Welsh, Butters, Hughes, Mohs, & Heyman, 1992), below 30 on category word fluency (animals, fruits, and vegetables; Monsch et al., 1992), or in the impaired range (>3) on either subtest of the Blessed Dementia Scale (Activities, Habits, and Personality or Information, Memory, and Concentration; Blessed, Tomlinson, & Roth, 1968). Patients with AD were excluded from the study if they scored outside of the mild to moderate range (4–16) on the Information, Memory, and Concentration subtests (Locascio, Growdon, & Corkin, 1995). Participants were also excluded if they were characterized by clinically significant depression, alcohol or drug use, or brain damage or if English was not their primary language. Three older adults were excluded on the basis of these criteria, resulting in 6 male and 6 female participants in both groups. All participants had normal or corrected-to-normal vision. The patients with AD were matched to the 12 healthy older adults on the basis of age (for the patients with AD, M = 71.6 years, range = 60–85 years; for the older adults, M = 74.3 years, range = 63–90 years), education (for the patients with AD, M = 15.3 years, range = 12–21 years; for the older adults, M = 16.5 years, range = 12–20 years), and estimated verbal IQ as measured by the National Adult Reading Test—American Version (Blair & Spreen, 1989; for the patients with AD, M = 119.0,

range = 107–132; for the older adults, M = 124.1, range = 110–132). (See Budson et al., 2000, for details on participant demographics.) In addition to the screening and matching tests, performance on controlled word fluency to letters (*F, A,* and *S*; Monsch et al., 1992) was also recorded.

Design

The experimental design included a between-subjects variable of group (patients with AD vs. older adults) and two within-subject variables of transformational distance and category size. Transformational distance had three levels for studied items (near, middle, or far) and four levels for nonstudied items (prototype, near, middle, or 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 aforementioned four levels plus "novel" category items for which no related items were present at study; the novel items provided an estimate of baseline levels of false alarms. In addition, noncategorized items (termed *unrelated*; see the *Stimuli* section below) were included as both studied and nonstudied items.

Stimuli

The stimuli were color depictions of complex, multifeatured, abstract objects, which were created using a computer graphics program (Aldus Freehand; Aldus Corporation, Seattle, WA). Most of the stimuli were categorized items, which were generated by first creating a novel prototype according to a specific set of construction rules (described below) and then generating additional exemplars that belonged to the same category through manipulations that distorted the initial prototype to a greater or lesser degree. Noncategorized or unrelated items that did not follow the rules of construction for the categorized items were also created (see Koutstaal, Schacter, Verfaellie, et al., 1999, for additional details regarding the stimuli).

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 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. 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, color, outline, and size, to create additional exemplars that possessed varying degrees of similarity to the prototypes. The placement of features on these exemplars was also altered but within a specified range (e.g., a given feature would generally appear more toward the top or the side of an object, but its precise placement varied somewhat). After manipulation of the prototypes in an algorithmic fashion, the stimuli were further altered individually to increase the within-category distinctiveness of the stimuli. The similarity of the items to the prototypes was assessed and confirmed by asking eight raters to place the exemplars in each category on a distance metric corresponding to how similar they were to the prototype (cf. Bahrick, Clark, & Bahrick, 1967). Two exemplars that constituted the clearest representatives of the three distances (near, middle, and far) were chosen as the critical stimulus items (to be counterbalanced across study and test status; see below).

Examples of the categorized items are provided in Figure 1. Shown in Figures 1A and 1B are 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 Figure 1C; unrelated items were never composed 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. To avoid confounding the number of related exemplars that were presented at study with the number of items that were tested, we tested only a subset of all the items in each category: 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 for which only one item was presented, only that 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 for which 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, and far) were balanced; in addition, for the test lists, the number of novel items and prototypes were also balanced across the test thirds. Finally, additional templates were created by using different orderings of the items within each of the test thirds.

For counterbalancing purposes, the 18 categories were assigned to six stimulus subsets made up of 3 categories each; these subsets were used to determine, across participants, whether a given category was shown as studied items or as novel items, with one stimulus subset assigned to each of the large, medium, small, and novel category size conditions (each of which was thus represented by 3 categories) and two stimulus subsets assigned to the single condition (thus represented by 6 categories). Counterbalancing across participants ensured that each category occurred once in each of the six conditions, with each category represented once with the A and B exemplars as the studied and nonstudied exemplars. Thus, counterbalancing required 12 participants.

Procedure

The overall procedure involved three phases: 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 lasting approximately 40 to 60 min. The stimuli at both study and test were presented on an Apple Macintosh Powerbook 5300c computer using PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). The stimuli appeared in the center of the screen, with prompts for responding to the encoding or the recognition test displayed beneath.

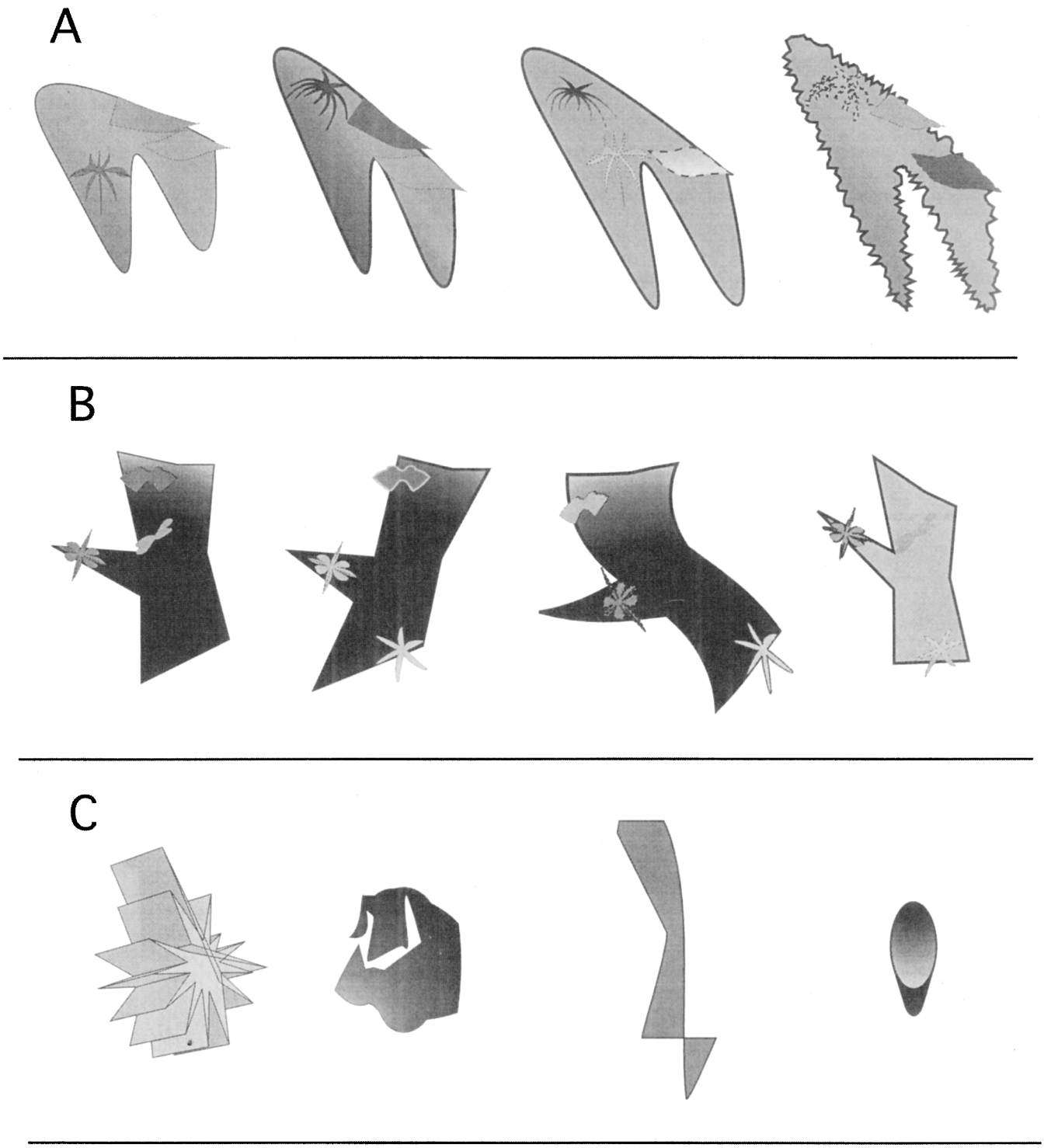


Figure 1. Examples of categorized stimuli (A and B) and unrelated items (C) used in the study. Panels A and B show examples from two different categories, including, from left to right, a category prototype and items from the near, middle, and far transformational distances. Note that, although for illustrative purposes here the stimuli are shown in black and white, the stimuli as shown to the participants were presented in color, with color being an important attribute that was varied both within and across categories. From "Perceptually-Based False Recognition of Novel Objects in Amnesia: Effects of Category Size and Similarity to Category Prototypes," by W. Koutstaal, D. L. Schacter, M. Verfaellie, C. J. Brenner, and E. M. Jackson, 1999, *Cognitive Neuropsychology, 16,* p. 324. Copyright 1999 by Psychology Press Limited. Reprinted by permission of Psychology Press Limited, Hove, UK.

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 s, and participants were asked to rate the "overall complexity" of the stimulus on a 9-point scale, with 1 indicating the stimulus was not complex at all and 9 indicating the stimulus was extremely complex. When making their complexity ratings, participants were instructed to consider all aspects of the stimulus, including both different dimensions of the stimulus (e.g., shape, color, size, and outline) and all of the components of the stimulus. Participants verbally rated the item complexity, and the experimenter entered their response on the computer keyboard using the number keys 1–9. The study phase was followed by a 5-min retention interval.

In the test phase, participants were shown a subset of the items shown earlier 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, for which 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 consisted 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. Following the test phase, which was self-paced, participants were debriefed.

Results

Standard Neuropsychological Tests

The results of the standard neuropsychological tests revealed, not surprisingly, that patients with AD performed significantly worse than healthy older adults on word fluency to letters (33.3 vs. 45.5, respectively), $F(1, 22) = 7.60$, $p = .011$, and categories (24.7 vs. 52.2, respectively), $F(1, 22) = 56.64$, $p < .001$, as well as on both subtests of the Blessed Dementia Scale (Activities, Habits, and Personality: 4.4 vs. 0.4), $F(1, 22) = 34.27$, $p < .001$, and (Information, Memory, and Concentration: 9.5 vs. 0.1), $F(1, 22) = 55.23$, $p < .001$. (See Budson et al., 2000, for details concerning participants' performance on these measures.) Note that the expected reversal in performance on word fluency to letters versus categories in patients with AD was found (Monsch et al., 1992): Whereas the healthy older adults were able to generate numerically more words to semantic categories than letters, the patients with AD generated fewer words to semantic categories than letters, highlighting their semantic memory deficit.

True and False Recognition

Data on the differences and similarities in true and false recognition between patients with AD and healthy older adults are presented first as a function of transformational distance and then as a function of category size.

Analyses by Transformational Distance

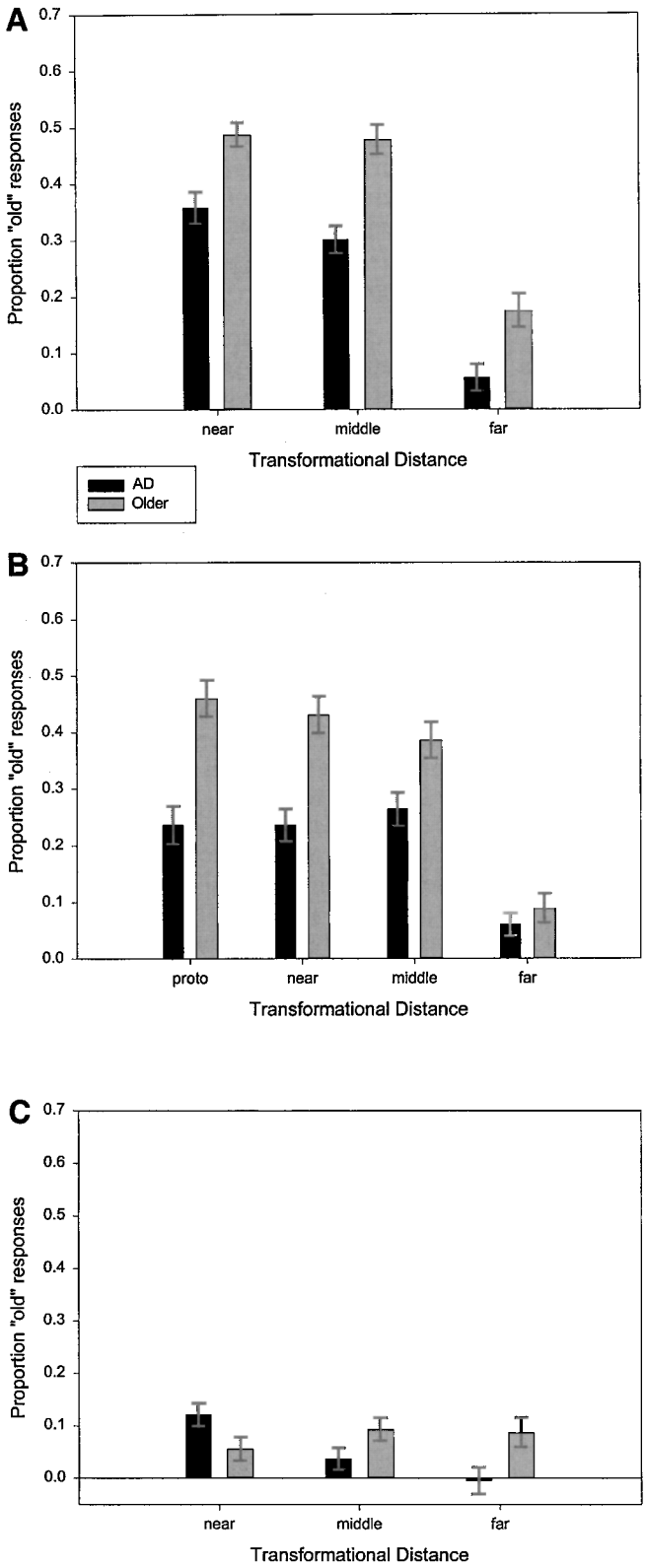
Table 1 presents the proportion of "old" responses to studied items (true recognition) and to nonstudied items (false recognition), separately as a function of group (AD or older) and transformational distance (near, middle, far, or prototype). Note that the results for single items, for which only one categorized item was presented at study and which was always from the middle distance, are not considered here but in the subsequent section concerning the effects of category size. Also shown in Table 1 is the proportion of "old" responses to unrelated items (i.e., the items that did not follow the rules of construction of the categorized items) and the proportion of "old" responses to nonstudied items that were from novel categories, for which no categorically related items were presented at study. The unrelated items provided one measure of true recognition and false alarms, whereas the novel items provided an index of the baseline level of false alarms to items that followed the same constructional rules as the studied categorized items.

Unrelated items. Compared with healthy older adults, patients with AD made numerically fewer "old" responses to the unrelated items (items that did not follow the rules of construction of the categorized items) that they had previously studied, although this difference was not significant, $F(1, 22) = 1.60$. Patients with AD did make significantly

Table 1
True and False Recognition Responses by Transformational Distance

	True recognition				False recognition					
Group	Near	Middle	Far	Unrelated	Prototype	Near	Middle	Far	Novel	Unrelated
AD										
M	.73	.68	.43	.47	.61	.61	.64	.44	.38	.24
SD	.22	.31	.22	.19	.25	.27	.30	.24	.22	.22
Older										
M	.70	.69	.38	.57	.67	.64	.59	.30	.21	.08
SD	.23	.23	.20	.18	.28	.25	.28	.22	.17	.11

Note. False recognition responses for the prototypes were based on category sizes of three, six, and nine items presented at study (i.e., excluding singles). Unrelated = unrelated items, that is, those that did not adhere to the rules used in generating the categorized items; Novel = categorized items not presented at study (the baseline level of false alarms); AD = patients with Alzheimer's disease; Older = healthy older adults.



more false alarms to nonstudied unrelated items, $F(1, 22) = 5.23$, $MSE = 0.153$, $p = .032$. After correction for these false alarms, patients with AD showed a significantly lower corrected hit rate for unrelated items as compared with older adults (.23 vs. .49, respectively), $F(1, 22) = 14.00$, $MSE = 0.396$, $p = .001$.

Novel items (baseline false alarms). Compared with healthy older adults, patients with AD made significantly more false alarms to nonstudied novel items that followed the same constructional rules as the studied items from perceptually related categories, $F(1, 22) = 4.33$, $MSE = 0.167$, $p = .049$.

True recognition. As one can see in Table 1, patients with AD and healthy older adults were affected by transformational distance, and both groups made similar numbers of "old" responses to studied items at each distance. A 2 (group: AD vs. older) × 3 (distance: near, middle, or far) analysis of variance (ANOVA) showed a significant effect of distance, $F(2, 44) = 52.37$, $MSE = 0.694$, $p < .001$, but no effect of group, $F(1, 22) < 1$, and no Group × Distance interaction, $F(2, 44) < 1$. After correction for baseline false alarms by subtracting the proportion of "old" responses to the novel items from the proportion of "old" responses to the studied items, an ANOVA demonstrated a significant effect of group, showing an overall lower level of corrected true recognition in patients with AD as compared with older adults, $F(1, 22) = 4.75$, $MSE = 0.359$, $p = .040$ (see Figure 2A).

False recognition. Table 1 shows that patients with AD and healthy older adults made large numbers of "old" responses to nonstudied items that were perceptually similar to studied items. A 2 (group: AD vs. older) × 4 (distance: prototype, near, middle, or far) ANOVA showed a significant effect of distance, $F(3, 66) = 21.50$, $MSE = 0.410$, $p < .001$; no effect of group, $F(1, 22) < 0.1$; and a trend toward a Group × Distance interaction, $F(3, 66) = 2.38$, $MSE = 0.045$, $p = .078$. Figure 2B shows that the level of novel-corrected false recognition in patients with AD was lower than that in older adults at all transformational distances. An ANOVA of the novel-corrected false recognition yielded a marginally significant difference: effect of group, $F(1, 22) = 4.06$, $MSE = 0.478$, $p = .056$. Thus, as we expected, patients with AD demonstrated lower levels of perceptually based corrected false recognition than healthy older adults. To further explore this difference, we compared AD and older novel-corrected false recognition at

Figure 2. Mean proportion of "old" responses to studied items (A: novel-corrected true recognition by transformational distance) and nonstudied items (B: novel-corrected false recognition by transformational distance), after false alarms to novel category items were subtracted. Also shown is the mean proportion of "old" responses to studied items minus nonstudied items (C: item-specific recollection by transformational distance). Results are shown separately as a function of group (patients with Alzheimer's disease [AD] or healthy older adults) and transformational distance (near, middle, far, and prototype [proto] when present). Error bars represent standard errors of the means.

transformational distances that were either closer to or farther away from the category prototype. To reduce the number of comparisons, we averaged the two close distances (prototype and near) and the two far distances (middle and far), contrasting these two closer and two farther categories. Although, compared with older adults, patients with AD demonstrated numerically lower rates of corrected false recognition at both the closer (.24 vs. .45) and the farther (.16 vs. .24) distances, only the items that were closer in transformational distance to the prototype were statistically different between the two groups: closer, $F(1, 22) = 6.35$, $MSE = 0.260$, $p = .020$; farther, $F(1, 22) = 1.08$. As we expected, this difference between the two groups was greater when older adults were most susceptible to the effects of gist memory, that is, when lure items were closer in transformational distance to the category prototype.

Item-specific recollection. True recognition ("old" responses to studied items) can be thought of as a combination of gist memory plus item-specific recollection. In contrast, false recognition of related lures ("old" responses to nonstudied items from the same categories as studied items) is likely a measure of gist memory minus any item-specific recollection that is available to counteract the effect of the gist. Thus, subtracting false recognition of related lures from true recognition should provide a measure of the item-specific recollection used by the groups. As one can see in Figure 2C, neither patients with AD nor healthy older adults were able to use much item-specific recollection—t tests showed that only the near distance was significantly different from zero for the patients with AD, $t(11) = 2.72$, $MSE = 0.120$, $p = .020$; other distances: $ts(11) < 1$, and the middle distance showed an almost significant difference from zero for the older adults, $t(11) = 2.06$, $MSE = 0.092$, $p = .064$; others: $ts(11) < 1.5$. An ANOVA between the two groups yielded no effects of distance, group, or Group × Distance interactions, nor did pairwise comparisons at near, middle, or far distances yield a significant effect, $Fs(1, 22)$ and $Fs(2, 44) < 1.7$.

Analyses by Category Size

Table 2 presents the proportion of "old" responses as a function of group (patients with AD or healthy older adults) and category size (one, three, six, or nine related exemplars shown at study) for both studied items (true recognition) and all nonstudied items (false recognition). Also shown separately is the proportion of "old" responses to the nonstudied category prototypes (false recognition of prototype) as well as to novel items (nonstudied items that followed the same constructional rules as the studied items from perceptually related categories as a baseline measure of false alarms).

True recognition. As one can see in Table 2, patients with AD and healthy older adults were affected by category size, and both groups made similar numbers of "old" responses to studied items at each category size. A 2 (group: patients with AD vs. older adults) × 4 (category size: nine, six, three, or one related exemplar shown at study) ANOVA showed a significant effect of category size, $F(3, 66) = 3.07$, $MSE = 0.075$, $p = .038$; no effect of group, $F(1, 22) < 1$; and no Group × Category Size interaction, $F(3, 66) < 1$. An ANOVA after correction for baseline false alarms produced similar results: a trend toward an effect of group for the novel-corrected data, $F(1, 22) = 3.39$, $MSE = 0.361$, $p = .079$ (see Figure 3A). To further explore the effect of category size, we performed comparisons of novel-corrected true recognition for larger versus smaller category sizes in the patients with AD and in the healthy older adults. As with the distance analyses, to reduce the number of comparisons, we averaged the two larger categories (six and nine related items presented) and the two smaller categories (one and three items presented), contrasting these combined larger and smaller categories. Patients with AD and older adults showed numerically different but statistically similar levels of novel-corrected true recognition for the smaller category (.24 vs. .34, respectively), $F(1, 22) = 1.7$, but patients with AD demonstrated significantly lower levels of novel-corrected true recognition for the larger category than did the older adults (.31 vs. .45, respectively), $F(1, 22) = 4.43$, $MSE = 0.132$, $p = .047$. As we had predicted, the greater numbers of categorized items presented at study in the larger groups allowed the older adults to raise their level of corrected true recognition above that of the patients with AD.

Table 2
True and False Recognition Responses by Category Size

	True recognition				False recognition					False recognition of prototype			
Group	9	6	3	1	9	6	3	1	Novel	9	6	3	1
AD													
M	.69	.67	.62	.61	.61	.54	.54	.44	.38	.58	.72	.53	.31
SD	.26	.26	.36	.23	.26	.30	.28	.30	.22	.32	.24	.36	.25
Older													
M	.67	.66	.56	.53	.60	.53	.40	.25	.21	.67	.81	.53	.21
SD	.24	.25	.25	.27	.27	.26	.27	.26	.17	.32	.30	.36	.19

Note. The numbers 9 (large), 6 (medium), 3 (small), and 1 (single) indicate the number of categorically related items presented at study. Novel = categorized items not presented at study (the baseline level of false alarms); AD = patients with Alzheimer's disease; Older = healthy older adults.

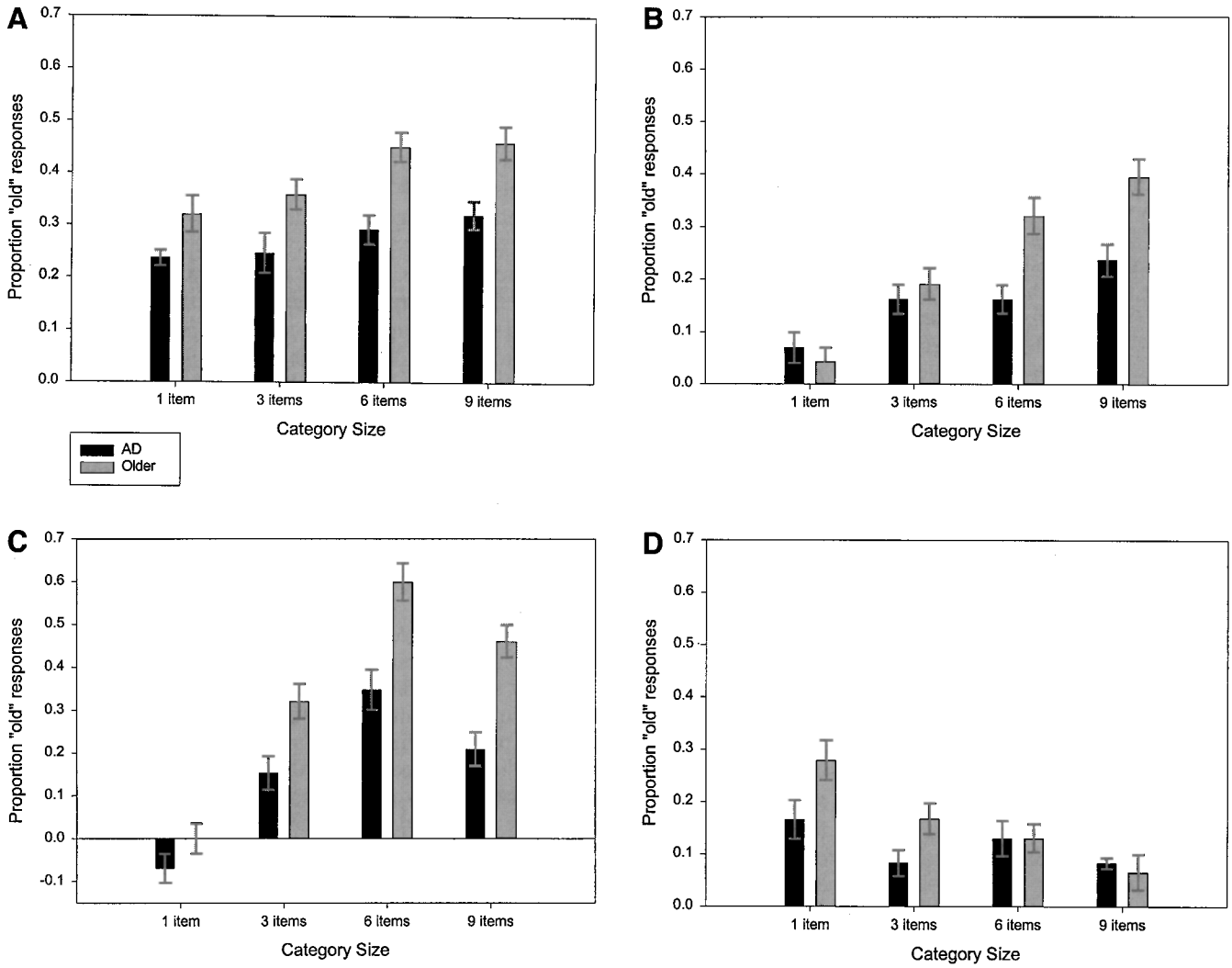


Figure 3. Mean proportion of "old" responses to studied items (A: novel-corrected true recognition by category size) and nonstudied items (B: novel-corrected false recognition by category size), after false alarms to novel category items were subtracted. For false recognition, responses are also shown for the prototypes (C: novel-corrected false recognition of prototype by category size). Also shown is the mean proportion of "old" responses to studied items minus nonstudied items (D: item-specific recollection by category size). Results are shown separately as a function of group (patients with Alzheimer's disease [AD] or healthy older adults) and category size (one, three, six, or nine category exemplars shown at study). Error bars represent standard errors of the means.

False recognition. Table 2 shows that patients with AD and healthy older adults showed similar rates of false recognition at all category sizes. An ANOVA revealed a significant effect of category size, $F(3, 66) = 8.77$, $MSE = 0.331$, $p < .001$; no effect of group, $F(1, 22) < 1$; and no Group × Category Size interaction, $F(3, 66) = 1.6$. After correction for baseline false alarms, an ANOVA produced similar results: no effect of group, $F(1, 22) = 1.8$ (see Figure 3B). However, comparisons between corrected false recognition did show significant differences between patients with AD and older adults for the average of the two larger categories, $F(1, 22) = 4.34$, $MSE = 0.149$, $p = .049$, but not the smaller categories, $F(1, 22) < 0.1$, reflecting the more robust gist memory available to the older adults with the larger but not the smaller categories.

We also considered novel-corrected false recognition of the prototypes separately as a function of the number of exemplars that were studied (Figure 3C). An ANOVA revealed no Group × Category Size interaction, $F(3, 66) < 1$, along with significant effects of category size, $F(3, 66) = 20.17$, $MSE = 1.10$, $p < .001$, and group, $F(1, 22) = 5.43$, $MSE = 0.813$, $p = .029$, again suggesting that effects of transformational distance were relevant: Older adults showed greater effects of gist memory than patients with AD when the items presented were closer in transformational distance (or, in this case, identical) to the category

prototype. Regarding the effects of category size, pairwise comparisons demonstrated significant differences between the patients with AD and the older adults for the average of the larger categories, $F(1, 22) = 6.14$, $MSE = 0.375$, $p = .021$, but not the smaller categories, $F(1, 22) = 2.4$, $p = .135$.

Item-specific recollection. As one can see in Figure 3D, patients with AD and older adults demonstrated similar levels of item-specific recollection: There were no overall group differences or differences in pairwise comparisons for the two larger or two smaller categories between patients with AD and older adults, $Fs(1, 22) < 2.4$, nor was there an effect of category size or a Group × Category Size interaction, $Fs(3, 66) < 2.4$, $ps > .11$. For this analysis, however, with the exception of the three- and six-item sizes for the patients with AD and the nine-item size for the older adults, t tests revealed that all other values were significantly different from zero, $ts(11) > 2.2$, $ps < .05$. Thus, in this analysis of category size, patients with AD and older adults showed similar evidence of low levels of item-specific recollection that allowed them to distinguish between true and false items at test to only a limited extent.

Discussion

Previous research has shown that patients with AD exhibit lower levels of false recognition of semantically related words than do healthy older adults (Balota, Cortese, et al., 1999; Budson et al., 2000). The present experiment has extended this earlier research by demonstrating that, compared with healthy older adults, patients with AD show lower levels of perceptually based false recognition of abstract novel objects. Thus, the observed false recognition differences between patients with AD and healthy older adults are not a limited effect, seen only with semantically related words, but instead are likely to be a more general phenomenon that may be found with many different types of stimuli, as is the case for patients with amnesia.

Overall, comparisons between patients with AD and healthy older adults revealed that patients with AD consistently made higher numbers of "old" responses to nonstudied items that were either unrelated (items unlike any of the categorized objects) or novel (items that followed the same constructional rules as the categorized items), raising their baseline level of false alarms and lowering their corrected true and false recognition. Analysis of the unrelated items showed that the patients with AD made fewer "old" responses to studied items and many more "old" responses to nonstudied unrelated items than did older adults, producing a significantly lower corrected hit rate. In a similar manner, patients with AD made almost twice as many false alarms as older adults to novel items, giving them a much higher rate of baseline false alarms.

After correction for baseline false alarms, in the analyses of transformational distance, patients with AD showed significantly lower levels of true recognition and marginally significant lower levels of false recognition of abstract novel objects than did healthy older adults. Furthermore, both patients with AD and older adults were more likely to respond "old" to an item, whether studied or unstudied, when the item was closer in distance to the category prototype. As we predicted, this effect of transformational distance for false recognition showed a trend toward being stronger for the older adults than the patients with AD. This conclusion is suggested by the trend toward a Group × Distance interaction as well as the fact that the groups were statistically different only when lure items were close in transformational distance to the category prototype. Thus, degraded gist memory in patients with AD was most evident when older adults showed the greatest susceptibility to the effects of gist—when items were most similar to the prototype.

In the overall analyses of category size, patients with AD showed lower levels of novel-corrected true and false recognition only when a relatively large number of category exemplars were presented at study—a condition that allowed the older adults to build up robust gist memory. As we expected from our discussion of transformational distance, novel-corrected false recognition of the category prototypes alone yielded similar but more significant results: an overall effect of group as well as an effect for larger, but not smaller, categories.

Patients with AD and healthy older adults were able to use item-specific recollection to allow them to distinguish studied from nonstudied but related items to only a limited extent. There were no differences in item-specific recollection between the groups for the analyses of either transformational distance or category size, but this result is largely attributable to a floor effect on item-specific memory. As we expected, both groups demonstrated the numerically greatest item-specific recollection in the condition in which gist influences were least potent: the single categories in which only one exemplar was presented at study.

To assure that the conclusions reached by our analyses were not solely due to the fact that the patients with AD could remember fewer items than healthy older adults, following the analyses of Balota, Cortese, et al. (1999), we matched 6 patients with AD with 6 older adults on the basis of their memory for the unrelated items (for which influences of gist would be relatively small). The resulting analysis showed numerical trends for the majority of key comparisons, similar to the results observed with all participants, although the reduction in power resulted in these trends being nonsignificant (see Table 3). Further studies with larger samples are needed to determine definitively whether the evidence for impaired gist memory in patients with AD is attributable, at least in part, to deficits in item-specific memory or to overall memory differences.

The fact that patients with AD showed reduced false recognition of perceptually related novel objects has important implications for previous findings of reduced false recognition in patients with AD (Balota, Cortese, et al., 1999; Budson et al., 2000). Specifically, our data suggest that the reduced false recognition of semantically related words in patients with AD is not produced entirely by their semantic memory deficits. This result is consistent with previous research indicating that although semantic memory

Table 3
Critical Analyses for All Participants and for a Subset of Participants Matched by Their Unrelated Hit Rate

Selected analyses	*N*	AD	Older	*F*(1, 22) *F*(1, 11)	*p*
Unrelated items					
Correlated unrelated true recognition	12	.23	.49	14.00	.001
	6	.36	.36	<.001	*ns*
Transformational distance					
Corrected mean true recognition	12	.24	.38	4.75	.040
	6	.22	.32	1.30	*ns*
Corrected prototype–near false recognition	12	.24	.44	6.35	.020
	6	.24	.33	<1	*ns*
Corrected middle–far false recognition	12	.16	.24	1.08	*ns*
	6	.14	.13	<0.1	*ns*
Category size					
Corrected larger (9 & 6) true recognition	12	.31	.45	4.43	.047
	6	.29	.38	<1	*ns*
Corrected smaller (1 & 3) true recognition	12	.24	.34	1.69	*ns*
	6	.20	.26	<1	*ns*
Corrected larger (9 & 6) false recognition	12	.20	.36	4.34	.049
	6	.21	.22	<0.1	*ns*
Corrected smaller (1 & 3) false recognition	12	.12	.12	<.001	*ns*
	6	.13	.13	<0.01	*ns*
Corrected larger (9 & 6) prototype false recognition	12	.28	.53	6.14	.021
	6	.26	.35	<1	*ns*
Corrected smaller (1 & 3) prototype false recognition	12	.04	.16	2.41	*ns*
	6	.06	.14	<1	*ns*

Note. AD = patients with Alzheimer's disease; Older = healthy older adults; *ns* = nonsignificant $p > .05$.

deficits are present in AD, they are rarely as severe as episodic memory deficits (Hodges & Patterson, 1995).

Patients with AD showed significantly lower levels of false recognition even though an incidental encoding task was used to reduce the chances that participants would use a categorization memory strategy. In the DRM semantic-associates paradigm, the semantic relationships among studied words are apparent and may be used by healthy participants to help organize and remember target lists. On recognition testing, this strategy would tend to support both true recognition of studied words and false recognition of related lures. Because patients with AD exhibit deficits in their ability to cluster words in taxonomic categories (Simon et al., 1994), they would be less likely than healthy older participants to notice the semantic organization of the study list. Moreover, patients with AD are less likely than older adults to develop strategies (Bondi, Monsch, Butters, Salmon, & Paulson, 1993; Paolo, Axelrod, Troster, Blackwell, & Koller, 1996). The failure of patients with AD to use a semantic organization strategy during encoding of study lists may contribute to their lower levels of true and false recognition as compared with older adults in the standard DRM paradigm. However, in this study, items were grouped together by perceptual, not semantic, similarities, and an incidental encoding task was used, reducing the incentive for older adults to develop semantic strategies to help them remember the items. Thus, it is unlikely that the differences in true and false recognition between patients with AD and older controls documented here can be attributed to older adults' use of semantic encoding strategies.

Although in the present study the within-category lures shared many perceptual features with the studied items, it is highly unlikely that the specific objects used as lures were themselves generated or imagined by the older adults during their initial encounter with the study stimuli. As discussed in Koutstaal and Schacter (1997), although the critical lure might be produced or spring to mind in the converging semantic-associates paradigm used by Deese (1959), Roediger and McDermottt (1995), and others, the use of highly distinctive lures as images in this perceptual paradigm allows us to eliminate source confusions involving implicit associative responses as a potential reason for the lower level of false recognition in patients with AD than in older adults.

One additional factor that may play a role in the false recognition of perceptually similar objects in patients with AD has recently been discussed by Kéri et al. (1999). They found that, like patients with amnesia, patients with AD were impaired in explicit recognition of dot patterns. Unlike patients with amnesia, however, patients with AD were also impaired in their ability to implicitly categorize dot patterns by prototype. Kéri et al. suggested that this difference between patients with AD and patients with amnesia was due to functional disruption of visual association areas in the patients with AD. Thus, impaired higher order visual processing may be an additional factor that contributed to the poor ability to acquire gist information in our patients with AD.

Thus, rather than reflecting the semantic memory deficits of patients with AD, their poor use of a categorization

memory strategy, or the implicit associative responses of the older adults, the lower level of false recognition observed in patients with AD for perceptually related novel objects is likely due to their poor memory for the general perceptual features, or gist, of the studied items. In this paradigm, we found no differences in item-specific recollection between the patients with AD and healthy older adults. Thus, it may be that for patients with AD, true and false recognition in both the semantic and perceptual paradigms are largely or entirely based on a degraded representation of the gist of the study list. This gist-based account of the data is supported by the analyses of both transformational distance and category size. False recognition of patients with AD was most reduced relative to older adults under conditions in which either the older adults were most susceptible to the effects of gist (i.e., with items at close transformational distance to the category prototype) or the older adults were most able to develop robust gist memory (i.e., when category size was largest).

The present results are consistent with our previous findings that patients with AD showed lower levels of false recognition of semantic associates than older adults on the first study–test trial of a paradigm that included repeated study–test trials (Budson et al., 2000). Budson et al. showed that with repeated presentations of the study list, older adults were able to use increasing item-specific recollection to distinguish between studied items and related lures. In contrast, patients with AD showed no evidence that they could acquire item-specific information and were thus driven by their gist memory into making more false alarms to related lures than older adults. In the present experiment, both older adults and patients with AD demonstrated very little use of item-specific recollection; in this setting, patients with AD made fewer "old" responses to nonstudied items than older adults when these false alarms were driven by gist, but patients with AD made more "old" responses to nonstudied items than older adults when gist influences were small (e.g., in the case of the unrelated and novel items). Thus, in a setting in which older adults demonstrate item-specific recollection as well as gist memory (Budson et al., 2000), patients with AD are less able to use item-specific recollection and show greater false recognition because of their inability to counter gist influences. In the present setting, in which neither patients with AD nor older adults were able to use much item-specific recollection, both true and false recognition were predominantly driven by gist. We suggest that patients with AD show lower levels of both types of recognition because their ability to acquire and retain gist information is degraded compared with that of healthy older adults.

Because both the present and previous (Budson et al., 2000) studies examined only recognition memory, an unanswered question is whether the idea that gist-based memory is degraded in patients with AD can account for the recall data in Balota, Cortese, et al.'s (1999) study. They found that the likelihood of false recall of lures was quite stable across healthy older adults and patients with early stage AD. If one assumes that false recall is based on gist information, this result could indicate relatively intact gist memory in patients with AD. However, if control participants use item-specific recollection on a recall test to counter gist influences, but patients with AD lack the ability to do so, then patients with AD might exhibit comparable levels of false recall because their memory for both gist and item-specific information is impaired. Future studies are necessary to clarify this issue. One possible strategy for addressing the issue would be to try to obtain estimates of gist-based memory without the countering influence of item-specific recollection. For instance, participants could be instructed to respond "old" when the item fits a previously studied semantic or perceptual category, regardless of whether the specific item was actually studied (cf. Brainerd & Reyna, 1998b; Schacter, Cendan, & Dodson, 2000). If patients with AD make fewer "old" responses to lures than do older adults with these instructions, there would be strong evidence for an impairment of gist-based memory. Until such evidence is available, we regard the degraded gist memory account of false recognition in patients with AD as a promising hypothesis that requires further evaluation and testing.

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.

Bahrick, H. P., Clark, S., & Bahrick, P. (1967). Generalization gradients as indicants of learning and retention of a recognition task. *Journal of Experimental Psychology, 75,* 464–471.

Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger, H. L., McDermott, K. B., & Yerys, B. E. (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology, 16,* 361–384.

Balota, D. A., & Duchek, J. M. (1991). Semantic priming effects, lexical repetition effects, and contextual disambiguation effects in healthy aged individuals and individuals with senile dementia of the Alzheimer's type. *Brain and Language, 40,* 181–201.

Balota, D. A., Watson, J. M., Duchek, J. M., & Ferraro, F. R. (1999). Cross-modal semantic and homograph priming in healthy young, healthy old, and in Alzheimer's disease individuals. *Journal of the International Neuropsychological Society, 5,* 626–640.

Bayles, K. A., Tomoeda, C. K., & Cruz, R. F. (1999). Performance of Alzheimer's disease patients in judging word relatedness. *Journal of the International Neuropsychological Society, 5,* 668–675.

Blair, J. R., & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. *The Clinical Neuropsychologist, 3,* 129–136.

Blessed, G., Tomlinson, B. E., & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *British Journal of Psychiatry, 114,* 797–811.

Bondi, M. W., Monsch, A. U., Butters, N., Salmon, D. P., & Paulson, J. S. (1993). Modified version of the Wisconsin Card Sorting Test in the detection of dementia of the Alzheimer type. *The Clinical Neuropsychologist, 7,* 161–170.

Borson, S., & Raskind, M. A. (1997). Clinical features and pharmacologic treatment of behavioral symptoms of Alzheimer's disease. *Neurology, 48,* S17–S24.

Bousfield, W. A., Whitmarsh, G. A., & Danick, J. J. (1958). Partial response identities in verbal generalization. *Psychological Reports, 4,* 703–713.

Brainerd, C. J., & Reyna, V. F. (1998a). Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology, 71,* 81–129.

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

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

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25,* 257–271.

Dalla Barba, G., & Wong, C. (1995). Encoding specificity and intrusion in Alzheimer's disease and amnesia. *Brain & Cognition, 27,* 1–16.

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

Förstl, H., Besthorn, C., Burns, A., Geiger-Kabisch, C., Levy, R., & Sattel, A. (1994). Delusional misidentification in Alzheimer's disease: A summary of clinical and biological aspects. *Psychopathology, 27,* 194–199.

Hodges, J. R., & Patterson, K. (1995). Is semantic memory consistently impaired early in the course of Alzheimer's disease? Neuroanatomical and diagnostic implications. *Neuropsychologia, 33,* 441–459.

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.

Kéri, S., Kálmán, J., Rapcsak, S. Z., Antal, A., Benedek, G., & Janka, Z. (1999). Classification learning in Alzheimer's disease. *Brain, 122,* 1063–1068.

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. (1999). Reducing gist-based false recognition in older adults: Encoding and retrieval manipulations. *Psychology and Aging, 14,* 220–237.

Koutstaal, W., Schacter, D. L., Verfaellie, M., Brenner, C. J., & Jackson, E. M. (1999). Perceptually-based false recognition of novel objects in amnesia: Effects of category size and similarity to category prototypes. *Cognitive Neuropsychology, 16,* 317–341.

Locascio, J. J., Growdon, J. H., & Corkin, S. (1995). Cognitive test performance in detecting, staging, and tracking Alzheimer's disease. *Archives of Neurology, 52,* 1087–1099.

Martin, A. (1992). Degraded knowledge representations in patients with Alzheimer's disease: Implications for models of semantic and repetition priming. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (2nd ed., pp. 220–232). New York: Guilford Press.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., & Price, D. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS–ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology, 34,* 939–944.

Monsch, A. U., Bondi, M. W., Butters, N., Salmon, D. P., Katzman, R., & Thal, L. J. (1992). Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Archives of Neurology, 49,* 1253–1258.

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., Mellits, E. D., & Clark, C. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD): Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology, 39,* 1159–1165.

Nebes, R. D. (1989). Semantic memory in Alzheimer's disease. *Psychological Bulletin, 106,* 377–394.

Nebes, R. D. (1992). Semantic memory dysfunction in Alzheimer's disease: Disruption of semantic knowledge or information-processing limitation? In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (2nd ed., pp. 233–240). New York: Guilford Press.

Nebes, R. D., & Halligan, E. M. (1996). Sentence context influences the interpretation of word meaning by Alzheimer patients. *Brain and Language, 54,* 233–245.

Nebes, R. D., & Halligan, E. M. (1999). Instantiation of semantic categories in sentence comprehension by Alzheimer patients. *Journal of the International Neuropsychological Society, 5,* 685–691.

Ober, B. A. (1999). Semantic memory in Alzheimer's disease: Loss of knowledge or deficits in retrieval? Introduction from the symposium organizer. *Journal of the International Neuropsychological Society, 5,* 623–625.

Ober, B. A., & Shenaut, G. K. (1999). Well-organized conceptual domains in Alzheimer's disease. *Journal of the International Neuropsychological Society, 5,* 676–684.

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

Paolo, A. M., Axelrod, B. N., Troster, A. I., Blackwell, K. T., & Koller, W. C. (1996). Wisconsin Card Sorting Test Short Form in persons with Alzheimer's and Parkinson's diseases. *Journal of Clinical and Experimental Neuropsychology, 18,* 892–897.

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

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

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

Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology, 83,* 304–308.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7,* 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., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803–814.

Salmon, D. P., Heindel, W. C., & Lange, K. L. (1999). Differential decline in word generation from phonemic and semantic categories during the course of Alzheimer's disease: Implications for the integrity of semantic memory. *Journal of the International Neuropsychological Society, 5,* 692–703.

Schacter, D. L., Cendan, D. L., & Dodson, C. S. (2000). *Retrieval conditions and false recognition: Testing the distinctiveness heuristic.* Manuscript submitted for publication.

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

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

Schacter, D. L., Verfaellie, M., Anes, M. D., & Racine, C. (1998). When true recognition suppresses false recognition: Evidence from amnesic patients. *Journal of Cognitive Neuroscience, 10,* 668–679.

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

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

Simon, E., Leach, L., Winocur, G., & Moscovitch, M. (1994). Intact primary memory in mild to moderate Alzheimer disease: Indices from the California Verbal Learning Test. *Journal of Clinical and Experimental Neuropsychology, 16,* 414–422.

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.

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

van der Hurk, P. R., & Hodges, J. R. (1995). Episodic and semantic memory in Alzheimer's disease and progressive supranuclear palsy: A comparative study. *Journal of Clinical and Experimental Neuropsychology, 17,* 459–471.

Welsh, K. A., Butters, N., Hughes, J. P., Mohs, R. C., & Heyman, A. (1992). Detection and staging of dementia in Alzheimer's disease: Use of the neuropsychological measures developed for the Consortium to Establish a Registry for Alzheimer's Disease. *Archives of Neurology, 49,* 448–452.

Received November 12, 1999
Revision received September 22, 2000
Accepted October 27, 2000 ■