ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Brain and Cognition 51 (2003) 251–261

Brain and Cognition

www.elsevier.com/locate/b&c

Semantic versus phonological false recognition in aging and Alzheimer's disease☆

Andrew E. Budson,[a,b,*] Alison L. Sullivan,[a] Kirk R. Daffner,[a,b] and Daniel L. Schacter[c]

[a] *Division of Cognitive and Behavioral Neurology, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115, USA*
[b] *Harvard Medical School, Boston, MA, USA*
[c] *Harvard University, Cambridge, MA, USA*

Accepted 7 October 2002

Abstract

Patients with Alzheimer's disease (AD) have been found to exhibit lower levels of false recognition of semantic associates compared with healthy older adults. Because these patients may show impaired performance of episodic and semantic memory tasks, this finding could be explained by deficits in episodic memory, semantic memory, or both. The authors adapted a paradigm for comparison of semantic versus phonological false recognition. They found that: (a) patients with AD exhibited lower levels of corrected false recognition of semantic, phonological, and hybrid (mixed semantic and phonological) lists than older adults, and (b) patients with AD showed very similar levels of false recognition for all list types. These results suggest that only episodic memory deficits are necessary to explain the lower level of false recognition of semantic associates observed in patients with AD when compared to older adults. Additionally, (c) older adults showed greater levels of semantic, phonological, and hybrid false recognition than younger adults, extending previous false recognition research of semantically related words and categorized colored photographs to phonologically related words.

Keywords: Memory; False memory; False recognition; Memory distortions; Alzheimer's disease; Aging; Older adults; Phonological; Semantic memory; Hybrid

1. Introduction

Patients with probable Alzheimer's disease (AD) not only fail to retrieve desired information but also suffer from distortions of memory (Förstl et al., 1994) that may impair their ability 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 have been explored experimentally using a false memory paradigm originally developed by Deese (1959) and revived and modified by Roediger and McDermott (1995). This Deese/Roediger–McDermott (DRM) paradigm has demonstrated robust levels of false recall and 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 non-presented 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).

Using the DRM paradigm, Balota et al. (1999b) found that, after controlling for false alarms to unrelated items, patients with AD falsely recognized *fewer* related lures than did healthy older adults (note that the recognition data were measured only after recall performance and therefore were contaminated by the

☆ This research was supported by National Institute on Aging AG08441, National Institute of Mental Health K23 MH01870, and National Institute of Neurological Diseases and Stroke NS26980. We are grateful to Jason Watson and David Balota for providing us with the word lists used in this study and for helpful discussions regarding the interpretation of our results. We thank Dorene Rentz and Leonard Scinto for their invaluable help and support.

* Corresponding author. Fax: 1-617-278-6963.
E-mail address: abudson@partners.org (A.E. Budson).

0278-2626/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0278-2626(03)00030-7

earlier recall task). Budson, Daffner, Desikan, and Schacter (2000) also studied patients with AD using the DRM paradigm, and likewise found that patients with AD exhibited lower levels of false recognition compared with healthy older adults.

Similar results have been obtained in patients with amnesia using several types of material, including the DRM semantic associates (Schacter, Verfaellie, & Pradere, 1996; Schacter, Verfaellie, Anes, & Racine, 1998a), categorized pictures (Koutstaal, Verfaellie, & Schacter, 2001), perceptually similar words (Schacter, Verfaellie, & Anes, 1997), and perceptually related novel objects (Koutstaal, Schacter, Verfaellie, Brenner, & Jackson, 1999). Patients with amnesia exhibit severe difficulties remembering recent experiences as a consequence of damage to the medial 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, the results from the patients with AD could be entirely explained by their poor episodic memory, as is the case for patients with amnesia.

In addition to their impairment in episodic memory, however, patients with AD also show impairment on some tasks that draw upon semantic memory. Such tasks include generating words from semantic categories (typically animals, fruits, vegetables; Monsch et al., 1992; Salmon, Heindel, & Lange, 1999) and judging word relatedness (Bayles, Tomoeda, & Cruz, 1999). Patients with AD perform normally, however, on other tasks requiring intact semantic memory, including semantic priming (Balota & Duchek, 1991; Balota, Watson, Duchek, & Ferraro, 1999a) and instantiation of semantic categories in sentence comprehension (Nebes & Halligan, 1999). These seemingly contradictory data have led Balota and colleagues to postulate that the major problem with semantic memory in patients with AD is not in underlying semantic networks, but rather in an attentional control system that provides access to those networks (Balota et al., 1999a, 1999b; Watson, Balota, & Sergent-Marshall, 2001). This theory would explain why patients with AD show normal performance on paradigms such as priming, that use relatively automatic activation processes, whereas they show impairment in paradigms that require cognitive effort, and are heavily dependent upon attentional control systems, such as category word generation.

It may be that the lower level of false recognition of semantically associated words that is seen in patients with AD is attributable in part to a semantic memory deficit (regardless of the precise etiology of that deficit). Supporting this idea, Dalla Barba and Wong (1995) have demonstrated that patients with AD who showed relatively poor performance on semantic memory tasks also produced fewer related intrusions on free-recall tests.

Impairment in semantic memory may contribute to the reduced level of false recognition of semantic associates in patients with AD: if these patients 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 upon 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, 1998; 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 related lures relative to the unrelated lure words, patients with AD would show much less selectivity between these two different types of false positive responses (see Budson, Desikan, Daffner, & Schacter, 2001, for additional theories of how these patients' semantic memory impairment may contribute to their lower level of false recognition of semantic associates).

To begin to explore the contribution of semantic memory impairments to the lower level of false recognition of semantic associates observed in patients with AD, Budson et al. (2001) examined false recognition of perceptually related novel objects with little semantic content in patients with AD and matched older adults. They found that compared to older adults, patients with AD showed lower levels of false recognition of these perceptually related novel objects as well.

Thus it is clear that the lower level of false recognition of semantic associates observed in patients with AD compared with older adults cannot be entirely attributed to their semantic memory deficit, but may also involve a more general deficit in the acquisition, retention, or retrieval of gist information. What remains unclear is the relative contribution of episodic and semantic memory impairments to the observed results. Because of the differences between experiments using perceptual related novel objects (Budson et al., 2001) and DRM semantically related words (Budson et al., 2000), it is not possible to resolve this issue directly by comparing the results of these two studies. However, comparing false recognition for semantically and non-semantically related items within the same experiment should clarify the relative contributions of semantic and episodic memory impairments to the lower level of false recognition observed in the patients with AD compared with older adults.

Recently, Watson et al. (2001) developed a paradigm that allows such a comparison. They studied true and false recall (but not recognition) in patients with AD, older adults, and younger adults with modified DRM lists in which items were either related semantically, phonologically, or a mixture of both ("hybrid lists"). For example, for the related but non-presented lure word *cold*, lists could be semantic (*chill*, *frost*, *warm*, *ice*, etc.), phonological (*code*, *told*, *fold*, *old*, etc.), or hybrid (*chill*, *told*, *warm*, *old*, etc.). Thus, because the related lure words are the same for semantic and phonological lists, it should be possible in this paradigm to directly compare semantic and phonological false recognition. Watson et al. (2001) found similar rates of false recall between patients with AD and older adults across all conditions. Interestingly, false recall varied by list type, with hybrid lists producing the largest false recall. In fact, hybrid lists showed the highest proportion of false recall relative to true recall across all participants; phonological lists showed somewhat lower false relative to true recall, while semantic lists showed the lowest level of this relative measure. Tests of recall however, do not tap the same underlying processes as tests of recognition and can be influenced differentially by aging (Craik & Jennings, 1992). For example, the attentional control system needed for recalling information may be different from or stressed to a different degree than that required for making old/new judgments during a recognition test. Moreover, false recognition has been studied more extensively and analytically than have recall intrusions (see Schacter, Norman, & Koutstaal, 1998b), and therefore may allow insights that would be difficult to obtain from studies of intrusion errors.

Therefore, to investigate the contributions of episodic and semantic memory deficits to the reduced level of false recognition of semantic associates in patients with AD compared to healthy older adults we modified the paradigm reported by Watson et al. (2001) for use with recognition. We tested patients with AD, matched older adults, and, to make our data more comparable to that of Watson et al. (2001), younger adults. We hypothesized that if semantic memory deficits alone are responsible for reduced false recognition, then semantic false recognition of patients with AD should be reduced, but their phonological false recognition should be normal—i.e., at the level of the healthy older adults. If, instead, episodic memory deficits alone are responsible, then semantic and phonological false recognition of patients with AD should be reduced by the same amount compared to older adults. Lastly, if both semantic and episodic memory deficits are responsible, then compared to older adults, phonological false recognition in patients with AD should be reduced, but not to the extent seen with semantic false recognition.

Previous findings in patients with AD and perceptually related novel objects (Budson et al., 2001), suggest that episodic memory deficits are one important component contributing to the lower level of false recognition observed in patients with AD, compared with older adults, in the semantic associates paradigm (Budson et al., 2000). Because Dalla Barba and Wong (1995) found that patients with AD who showed worse performance on semantic memory tasks produced fewer related intrusions, we also presumed that the semantic memory deficits in these patients would play a role in their low level of false recognition of semantically related items. In addition, neuropathological studies of patients with AD not only show plaques and tangles in structures that are critical to episodic memory performance such as medial temporal lobes, but pathology is also present in inferior lateral temporal lobes (Price & Morris, 1999)—regions that are important for normal semantic memory performance (Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; Hodges & Patterson, 1995; Hodges et al., 1999). Thus, we expected that both phonological and semantic false recognition would be reduced in patients with AD compared to healthy older adults, but that semantic false recognition would be reduced to a greater extent than phonological false recognition.

In addition, to our knowledge comparisons between older and younger adults have not been reported for false recognition of phonological words. Because older adults have been shown to exhibit higher levels of false recognition of semantically related words (Balota et al., 1999b; Norman & Schacter, 1997; Schacter, Israel, & Racine, 1999; Tun, Wingfield, Rosen, & Blanchard, 1998) and categorized colored photographs (Koutstaal & Schacter, 1997) when compared to younger adults, we expected that older adults would also show higher levels of phonological false recognition compared to younger adults, consistent with the recall data of Watson et al. (2001).

2. Methods

2.1. Participants

Eighteen patients with a clinical diagnosis of probable AD (National Institute of Neurological and Communicative Disorders and Stroke—Alzheimer's Disease and Related Disorders Association criteria used, McKhann, Drachman, Folstein, Katzman, & Price, 1984), 18 healthy older adults, and 18 younger 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. Older adults were recruited from participants in a longitudinal study of normal aging at Brigham and Women's Hospital, from spouses and friends (but not blood relatives) of the patients, as well as from flyers and posters placed in senior centers in and around Boston. Younger adults were recruited through fliers posted at Harvard University, and word-of-mouth. Written informed consent was obtained from all participants and their care-givers (where appropriate). The study was approved by the human subjects committee of Brigham and Women's Hospital and Harvard University. Participants were paid $10/h for their participation. Older adults were all community dwelling and were excluded if they scored below 30 on category word fluency (animals, fruits, vegetables; Monsch et al., 1992), or below 28 on the Mini Mental Status Examination (MMSE; Folstein, Folstein, & McHugh, 1975). Most patients showed mild to moderate impairment on the MMSE, all but one scoring above 16 (mean = 22.3, range 11–28). Participants were excluded if they were characterized by clinically significant depression, alcohol or drug use, focal brain damage such as that from strokes or tumors, or if English was not their primary language. All participants had normal or corrected to normal vision and hearing. The patients were matched to the older adults on the basis of sex (7 male and 11 female patients, 8 male and 10 female older adults), age (AD patient mean = 75.5 years, range = 60–85; older adult mean 73.9 years, range 60–81), and education (patient mean = 15.3 years, range = 9–20; older adult mean = 15.3 years, range = 12–19). The younger adults were matched on the basis of sex (9 male and 9 female); their ages ranged from 19 to 27 years, with a mean of 21.7 years.

2.2. Materials and design

The word lists were taken from Watson et al. (2001), and consisted of semantic, phonological, and hybrid lists for the critical lures *cold*, *dog*, *glass*, *gun*, *hard*, and *smoke*. Each list consisted of 12 words. Each participant studied two semantic, two phonological, and two hybrid lists. The words in each list were presented in the same order each time, and the lists were presented so that the critical lure items were in the same order for each participant, regardless of whether a particular lure was related to a semantic, phonological, or hybrid list. The experiment was counterbalanced so that an equal number of participants saw semantic, phonological, and hybrid lists for each lure word.

Each test list was composed of 48 words, in a different random order for each participant. Twenty-four of these items were studied words (from list positions 1, 2, 7, and 8) and 24 were unstudied: 6 related lures (1 from each list) plus 18 unrelated words taken from standard DRM lists (Roediger & McDermott, 1995; Stadler, Roediger, & McDermott, 1999; positions 1, 8, and 10). (Care was taken to assure that the unrelated words shared little or no semantic or phonological associations with the lure items.)

2.3. Procedure

Participants were tested individually, and instructed to read the study words out loud, and to remember them for a test session that would follow immediately. The words were presented on an Apple Macintosh Powerbook 3400c computer, one word at a time for 2600 ms each, in the center of the screen which was placed a comfortable viewing distance from the participant. There was a 400 ms interval between words. The six study lists were presented without interruptions. Words were presented visually in the same font and size at study and test. At test the words remained on the screen until the participant responded verbally with an "old" or "new" judgement. The experimenter then entered the appropriate response on the keyboard.

3. Results

The table shows the proportion of "old" responses to previously studied words (true recognition) and to related lure words (false recognition) as a function of group (younger adults, older adults, and patients with AD) and list type (semantic, phonological, and hybrid). Also shown are the responses to non-studied unrelated words. These unrelated words provide an index of the baseline false alarms. Comparisons between older adults and patients with AD are reported first, followed by comparisons between younger and older adults.

3.1. Unrelated words

Patients with AD made significantly more "old" responses to non-studied unrelated words compared to older adults, and therefore had a higher baseline false alarm rate. A one-way ANOVA demonstrated a significant effect of group ($F(1,34) = 39.8$, MSE = .027, $p < .0005$) (Table 1).

Older and younger adults made similar numbers of "old" responses to non-studied unrelated words, giving them similar baseline false alarm rates ($F(1,34) < 1$) (Table 1).

3.2. True recognition

Compared to healthy older adults, patients with AD made fewer old responses to studied items for all list types. A 2 (Group: older adults vs. patients with AD) × 3 (List Type: semantic, hybrid, and phonological) ANOVA found an effect of Group ($F(1, 34) = 9.77$, MSE = .125, $p = .004$), no effect of List Type ($F(2,68) = 1.12$, MSE = .022, $p = .332$), and a

Table 1
True and false recognition responses by list type

	True recognition			False recognition			Unrelated
	Semantic	Hybrid	Phonological	Semantic	Hybrid	Phonological	
Younger adults							
M	.83	.88	.89	.44	.47	.47	.06
SD	.20	.11	.14	.42	.42	.40	.07
Older adults							
M	.81	.79	.83	.67	.64	.67	.06
SD	.21	.19	.20	.34	.29	.42	.06
AD patients							
M	.65	.62	.52	.69	.67	.61	.40
SD	.22	.25	.33	.39	.40	.44	.22

only trend toward a Group × List Type interaction ($F(2,68) = 2.64$, MSE = .022, $p = .079$). This trend toward an interaction is present because patients with AD, but not older adults, showed some tendency toward making fewer "old" responses to studied words from the phonological lists compared with the other list types. After correction for baseline false alarms by subtracting the proportion of "old" responses to the unrelated words from the proportion "old" responses to the studied words, the effect of Group became even stronger ($F(1, 34) = 108.0$, MSE = .077, $p < .0005$) (Fig. 1).

Older and younger adults made similar numbers of "old" responses to studied words for all list types (Table 1). An ANOVA demonstrated no effect of Group ($F(1,34) = 1.46$, MSE = .062, $p = .236$), no effect of List Type, and no interaction (Fs$(2,68) < 1$). Because their baseline false alarms were similar, as in the analysis of the uncorrected data, no effect of Group was seen for corrected true recognition ($F(1,34) = 1.42$, MSE = .064, $p = .242$) (Fig. 1).

3.3. False recognition

Patients with AD and healthy older adults made very similar numbers of "old" responses to lure words. An ANOVA demonstrated no effect of Group ($F(1,34) < 1$), no effect of List Type, and no interaction (Fs$(2, 68) < 1$). Importantly for our initial hypothesis, there was not even a trend towards an effect of List Type nor was there a Group × List Type interaction. Because of their higher rate of baseline false alarms, analysis of the corrected data showed a significant effect of group; the

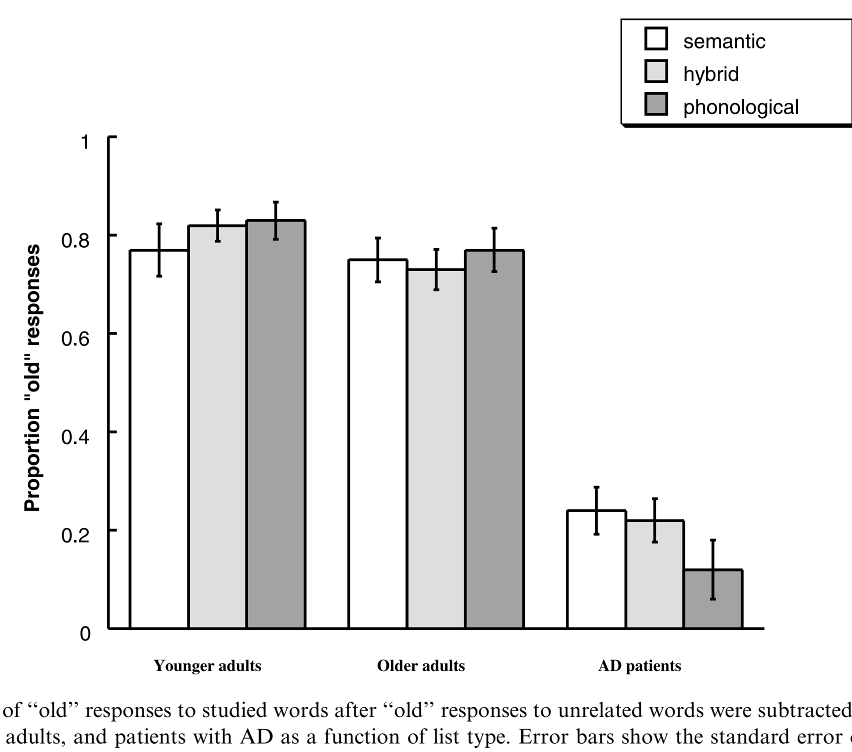

Fig. 1. Mean proportion of "old" responses to studied words after "old" responses to unrelated words were subtracted (corrected true recognition), by younger adults, older adults, and patients with AD as a function of list type. Error bars show the standard error of the mean.

patients with AD demonstrated a significantly lower level of corrected false recognition compared to older adults ($F(1,34) = 17.28$, MSE = .183, $p < .0005$) (Fig. 2). Thus, as with semantically associated words (Balota et al., 1999b; Budson et al., 2000), and perceptually related novel objects (Budson et al., 2001), patients with AD also show lower levels of phonological false recognition after correction for baseline false alarms.

Older adults made significantly more "old" responses to related lure words than younger adults across all list types (Table 1). An ANOVA showed an effect of Group ($F(1,34) = 5.08$, MSE = .201, $p = .031$), no effect of List Type, and no interaction ($Fs(2,68) < 1$). Analysis of the corrected data also produced an effect of Group ($F(1,34) = 5.73$, MSE = .178, $p = .022$) (Fig. 2). Thus, as had been observed in studies of semantically related words (Balota et al., 1999b; Norman & Schacter, 1997; Schacter et al., 1999; Tun et al., 1998) and categorized colored photographs (Koutstaal & Schacter, 1997), older adults in the present study showed higher levels of false recognition of phonologically related words than younger adults.

3.4. Relative false recognition: False recognition divided by true recognition

Because patients with AD have deficits in true recognition and recall in addition to false recognition and recall, meaningful between-group comparisons of false recognition and recall in patients with AD typically involve some type of correction. We have already performed one type of correction, by subtracting baseline false alarms from the uncorrected false recognition data. Balota et al. (1999b) controlled for differing levels of true recall by matching a subset of the participants on their true recall performance. This method has the advantage of allowing between-group false recognition and false recall comparisons directly without mathematical manipulation, but has the disadvantage of selecting a sample of participants who may not be representative of the whole, in addition to requiring relatively large numbers of participants. Using a different method, Watson et al. (2001) controlled for true recall by dividing the participants' false recall performance by their true recall performance—in this way obtaining their relative false recall performance. We also thought this was a good way to explore false memory in patients with AD.

Therefore, we divided the proportion of "old" responses to lure words by the proportion of "old" responses to studied words to obtain a measure of the relative false recognition of the participants. Patients with AD showed greater levels on this relative false recognition measure compared to healthy older adults (Fig. 3). An ANOVA showed an effect of Group ($F(1,34) = 6.25$, MSE = .877, $p = .017$), no effect of List Type and no Group × List Type interaction ($Fs(2,68) < 1$).

We also compared older and younger adults by their relative false recognition. Since false recognition was greater in older adults compared to younger adults while their true recognition was similar, it was not surprising

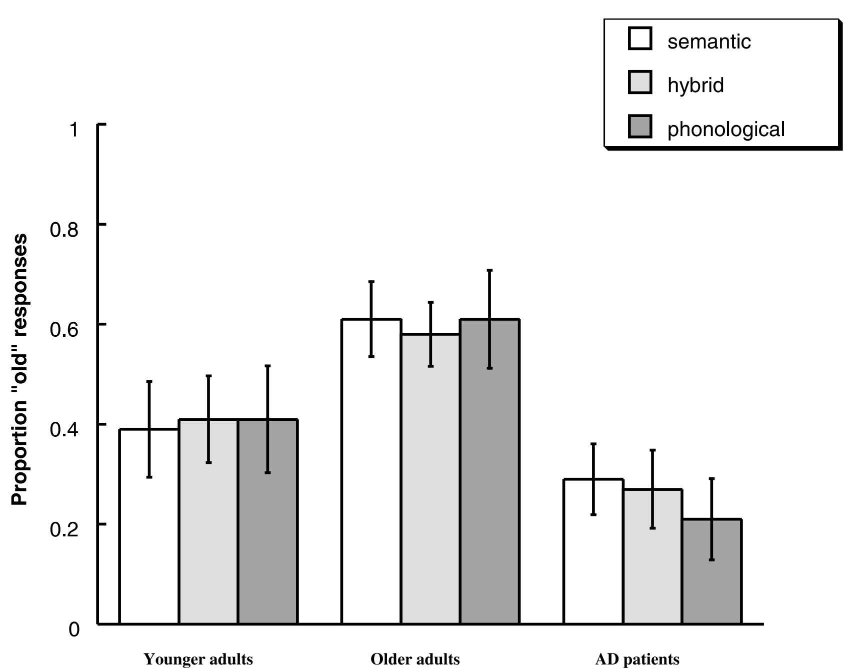

Fig. 2. Mean proportion of "old" responses to related lure words after "old" responses to unrelated words were subtracted (corrected false recognition), by younger adults, older adults, and patients with AD as a function of list type. Error bars show the standard error of the mean.

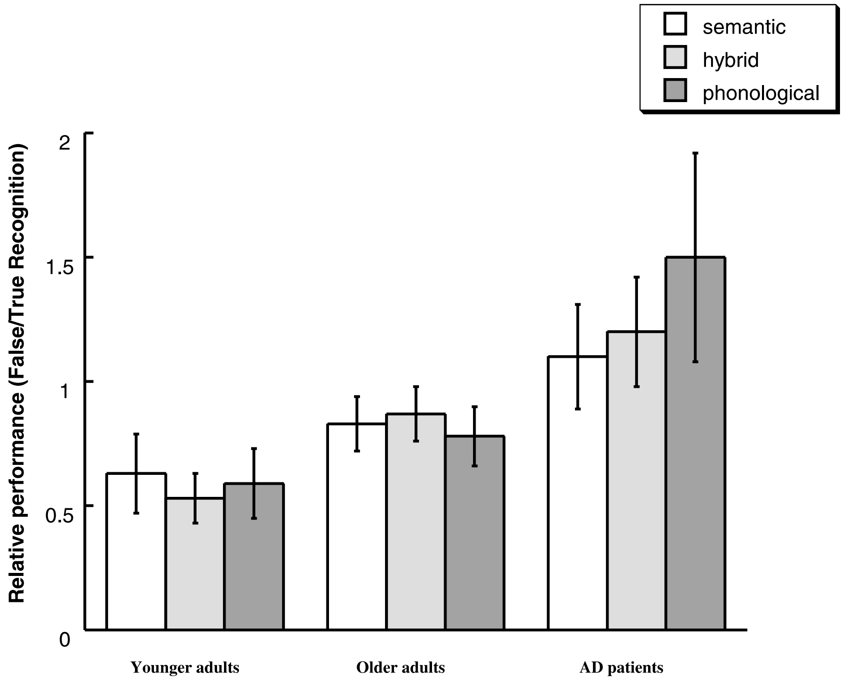

Fig. 3. Mean proportion of "old" responses to related lure words divided by studied words (relative false recognition: false/true), by younger adults, older adults, and patients with AD as a function of list type. Error bars show the standard error of the mean.

that older adults showed higher levels of relative false recognition compared to their younger counterparts (Fig. 3). An ANOVA demonstrated an effect of Group ($F(1,34) = 5.36$, $\text{MSE} = .312$, $p = .027$), no effect of List Type, and no interaction ($Fs(2,68) < 1$). Thus, in contrast to the relative false recall data of Watson et al. (2001), higher relative false recognition of hybrid lists compared with semantic and phonological lists was not observed.

4. Discussion

Previous research has shown that after correcting for unrelated false alarms, patients with AD exhibit lower levels of false recognition of semantically related words (Balota et al., 1999b; Budson et al., 2000) and perceptually related novel objects (Budson et al., 2001) compared to healthy older adults. The present experiment has extended this earlier research by demonstrating that, compared to older adults, patients with AD show lower levels of corrected false recognition of phonologically related words and hybrid lists containing both semantically and phonologically related words. Additionally, we have shown that as with studies of semantically related words (Balota et al., 1999b; Norman & Schacter, 1997; Schacter et al., 1999; Tun et al., 1998) and categorized colored photographs (Koutstaal & Schacter, 1997), older adults exhibit higher levels of false recognition of phonologically related words compared with young adults. Lastly, we have related our research to previous work studying false recall of semantic, phonological, and hybrid words lists (Watson et al., 2001). In contrast to these recall results, no differences were observed in the level of false recognition between semantic, phonological, and hybrid lists in the present experiment.

4.1. Semantic versus phonological false recognition in Alzheimer'sdisease

Patients with AD made more false alarms to nonstudied unrelated words, and fewer "old" responses to studied words, compared with older adults. Importantly for our initial hypothesis, we found that patients with AD exhibited very similar levels of false recognition of semantic, phonological, and hybrid lists. Corrected false recognition of all lists were lower compared to that of the older adults—but there was not even a numerical trend for corrected false recognition of semantic lists to be lower than that of phonological lists. Though a lack of difference must be interpreted cautiously, this finding is contrary to our prediction that patients with AD would show relatively reduced semantic false recognition. Further, the finding raises the possibility that only episodic memory deficits are responsible for the lower level of corrected false recognition of semantic associates observed in patients with AD when compared with older adults. Moreover, patients with AD showed greater levels of relative false recognition (false recognition divided by true recognition) compared to older adults.

There are several possible reasons why the impairment in semantic memory which patients with AD exhibit on some tasks did not reduce the corrected false recognition of semantic associates further than the reduction observed with phonological associates. One explanation relates to the theory that the semantic impairment in patients with AD is attributable to the dysfunctional attentional control system that provides access to those semantic networks (Balota et al., 1999a, 1999b; Watson et al., 2001). This theory would predict that performance on tasks that rely on relatively automatic activation processes would not reveal evidence of special semantic impairments. Because recognition tests may depend in part on such automatic processes (Jacoby, 1991), our results are consistent with this theory.[1]

Another reason why our patients with AD may not have exhibited lower corrected false recognition of semantic lists compared to phonological lists could be related to the specific words used in this paradigm. Poor performance on semantic tasks in patients with AD occurs in a graded fashion (Hodges & Patterson, 1995). For example, Hodges and Patterson (1995) found that while the most demented patients with AD showed impairment on all levels of a picture sorting task, the mild subgroup were impaired on the most difficult sorting task, but not on the easier sorting tasks. Thus, it might be that our patients would have shown difficulty spontaneously associating words in some of the broader categories (e.g., *caring* and *pink* with the category *girl*), while they may not have shown difficulty spontaneously associating words in the more narrow categories used in the present experiment (i.e., *snow* and *frigid* with the category *cold*, *murder* and *weapon* with the category *gun*).

Alternatively, we would argue that patients with AD exhibit similar levels of semantic and phonological corrected false recognition because false recognition for both lists is dependent upon similar kinds of gist memory. Pathology in AD is most prominent in medial temporal lobe structures necessary for episodic memory (Price & Morris, 1999). Patients with AD may simply show impairment in their memory for the general idea or meaning conveyed by the items on the semantic lists—and for the auditory and visual perceptual features of the phonological lists—because of their episodic memory deficits. Thus, corrected false recognition of semantic and phonological lists may be impaired to the same extent because false recognition for both lists may depend upon the same degraded gist memory.[2]

4.2. Interpreting false recognition results in aging and Alzheimer's disease

We have shown that for semantic, phonological, and hybrid lists, patients with AD make similar numbers of "old" responses to related lure words, yet show lower levels of false recognition after correction for false alarms to unrelated words, and higher levels of *relative* false recognition (false relative to true recognition), when compared with healthy older adults. Although these three results may seem contradictory, these different analyses are informative regarding the nature of the neuropsychological impairments in patients with AD, and are similar to the signal detection analyses performed in other studies (Budson et al., 2000; Koutstaal & Schacter, 1997;. Schacter et al., 1998a; Schacter et al., 1999).

Corrected false recognition provides a measure of how likely participants are to respond "old" to a related lure word versus an unrelated word, and thus controls for differences in response bias between the groups. "Old" responses to related lure words represent participants' gist-based false alarms plus their overall tendency to respond "old" to any word, minus any item-specific recollection available to counter the effects of gist. Therefore, subtracting false alarms to unrelated words from false recognition of lure words should provide a measure of participants' tendency to rely on gist information despite any opposing influence of item-specific recollection—but it does not necessarily indicate the amount of gist memory available to participants (unless they show no item-specific recollection). Patients with AD in this study, in fact, showed no ability to distinguish lures from studied words and thus showed no item-specific recollection. Compared to older adults, patients with AD showed lower levels of corrected false recognition for all list types, again suggesting that their gist memory is degraded. Because they made many more

[1] An additional prediction which follows from this line of reasoning is that if the false memory paradigm involved a more cognitively active task, such as recall, a reduction in the number of critical lure intrusions would be seen for semantic lists relative to phonological lists. Although there was no significant difference in the level of false recall of semantic versus phonological lists for patients with AD in the data of Watson et al. (2001), it is interesting to note that the more impaired patients with AD (mild dementia, Washington University Clinical Dementia Rating = 1.0 [Berg et al., 1998]) appeared to intrude critical lures for semantic lists at less than half the rate of critical lure intrusions for either phonological or hybrid lists (Fig. 2, Watson et al., 2001), further bolstering this theory.

[2] One assumption of this argument is that patients with AD do not exhibit specific deficits in phonological processing until quite late in the disease. If these types of deficits were present in our patients with AD, one could postulate that they showed reduced semantic false recognition due to their semantic memory deficit and reduced phonological false recognition due to their phonological processing deficit. Although some recent studies have suggested that there may be some impairment in phonological processing in selected patients with AD (e.g., Croot, Hodges, Xuereb, & Patterson, 2000), many more studies suggest that phonological processing is spared in patients with mild to moderate AD (e.g., Cohen, Wilcox, & Lerer, 1991; Patel & Satz, 1994; Verfaellie, Keane, & Johnson, 2000).

"old" responses to unrelated words than did older adults, the patients with AD also showed a significantly more liberal response bias than the older adults.

Relative false recognition provides a measure of how likely participants are to respond "old" to a related lure word (due to gist information) versus a studied word (due to either gist or item-specific information), and thus may provide a measure of the tendency of subjects to rely upon gist-based processes versus item-specific recollection, in addition to controlling for differences in response bias between the groups. The older adults showed relative false recognition rates of less than one, suggesting that they were somewhat able to use item-specific recollection to increase their true recognition and/or decrease their false recognition by remembering specific details of particular studied words—enabling them to counter their gist-based memory of the study lists. The younger adults showed even lower levels of relative false recognition than the older adults, suggesting that younger adults are more able to use item-specific recollection than older adults, consistent with previous work (Koutstaal & Schacter, 1997; Norman & Schacter, 1997; Schacter et al., 1999). In contrast, the patients with AD showed rates of relative false recognition greater than one, suggesting that their responses depended primarily upon gist based memory processes, consistent with previous work (Budson et al., 2000; Budson et al., 2001; Budson, Sitarski, Daffner, & Schacter, 2002). That their relative false recognition was greater than—rather than equal to—one is likely attributable to the fact that the lure words were chosen because of their extremely rich semantic and phonological associations. Thus, although both true and false recognition in patients with AD is dependent upon their gist-memory, the lure words—with their rich associations—are even more likely to trigger an "old" response than the studied words (cf. Roediger, Balota, & Watson, 2001).

4.3. Semantic, phonological, and hybrid false recognition in aging

Younger and older adults made similar numbers of "old" responses to studied words; they also made similar numbers of false alarms to unrelated words. However, older adults made more "old" responses to related lure words across all list types, giving them higher rates of uncorrected, corrected, and relative false recognition of semantic, phonological and hybrid lists compared with younger adults. These results extend studies of semantically related words (Balota et al., 1999b; Norman & Schacter, 1997; Schacter et al., 1999; Tun et al., 1998) and categorized colored photographs (Koutstaal & Schacter, 1997) in younger and older adults, and are consistent with the phonological false recognition observed in younger adults (Sommers & Lewis, 1999) as well as the false recall data of Watson et al. (2001). In contrast to the recall data of Watson et al. (2001), however, we did not find an effect of list type. Specifically, we did not find rates of false recognition higher for hybrid lists than for semantic or phonological lists. There are at least two possible reasons for this difference between false recall versus false recognition of hybrid lists.[3]

4.4. False recall versus false recognition of hybrid lists

Participants may be more likely to falsely recall hybrid lure items relative to semantic and phonological lure items because of the greater likelihood of hybrid lure words to be self-generated during free recall compared to semantic or phonological lure words. As discussed by Watson et al. (2001), the hybrid lists provide semantic and phonological boundaries such that the lure words are uniquely constrained. Thus, we would argue that as participants are actively trying to retrieve the study words during the recall test, the specific lure words are more likely to very quickly come to mind in the setting of the hybrid lists compared with the semantic and phonological lists. Participants may thus think of lure words quickly, although at least initially, they may realize that these lures were not presented. As the participants continue to recall items, however, they may forget that they generated these lure words themselves, and thus later intrude these words due to a source monitoring error (cf. Mather, Johnson, & De Leonardis, 1999). Semantic and phonological lists do not produce this unique constraint and may not as strongly converge on a *specific* lure word. Participants may self-generate lure words to semantic and phonological lists later during recall, and, because it would take them longer to forget that they had generated the lure words themselves, they would be intruded later—or not at all.[4]

One test for this theory would be to compare two groups of participants. The conventional group would simply write down on a piece of paper all the study words they could think of after each list. The experimental group would have their paper divided into halves, one half being for the studied words, and the

[3] Recent work by Watson et al. (in press) suggests that the difference between our results and those of Watson et al. (2001) may not only be due to differences between recall and recognition tests.

[4] It is interesting to consider this theory while noting the relative output position of the lure items in the recall data of Watson et al. (2001). Although there was no significant difference in the relative output position of the lures among the different lists, it is interesting to note the numerical trend showing that lures for hybrid lists were output somewhat earlier (.56) than for either semantic (.72) or phonological (.59) lists (Table 2; Watson et al., 2001). This numerical trend was particularly strong for the young adults (.56 vs. .75 and .68, respectively).

other half being for words related to studied words which come to mind but are not actually on the list. For example, if the participant thinks of the word *cold* early in the recall period and knows this word was not on the list, *cold* could be written on the "related but not studied" side of the list. If the participant then thinks of the word *cold* again later in the recall period, the list can be used to reduce source monitoring errors. We would predict that the experimental group would show lower rates of false recall, particularly of hybrid lists, compared with the conventional group.

A second possible explanation is that participants do not, in fact, intrude fewer *related* lure words for semantic and phonological lists compared with hybrid lists, just fewer *critical* lure items—that is, lure items which have been designated by the experimenter as the critical lures. While seeing a semantic list pushes participants toward the critical lure (e.g., *cold*), it also pushes participants toward related but non-critical items (e.g., *sleet*, *glacier*). This phenomenon is likely even stronger for the phonological lists (e.g., *bold* or *sold* in addition to *cold*). Because their lure word is uniquely constrained, hybrid lists are less likely to cause subjects to intrude related but non-critical lures. Thus, if related but non-critical intrusions are included with the critical intrusions in the false recall data, levels of total semantic and phonological false recall may be similar to hybrid false recall.[5]

4.5. Concluding comments

We have shown that patients with AD show lower levels of corrected false recognition of semantic, phonological, and hybrid lists compared to older adults, while older adults themselves show higher levels of false recognition of these lists compared to younger adults. Because patients with AD showed similar levels of semantic and phonological false recognition, we suggest that the previously observed lower level of corrected false recognition of semantic associates in patients with AD compared to older adults (Balota et al., 1999b; Budson et al., 2000) may be attributable only to the poor episodic memory of these patients, and not to their impairment in semantic memory. We have also observed that, in contrast to healthy older adults, patients with AD exhibit greater false compared to true recognition, which we hypothesize is due to their reliance on gist-based memory processes and inability to use item-specific recollection in the present study. Furthermore, we suggest that younger adults' greater ability to use item-specific recollection allowed them to counter gist-based memory processes and reduce their false recognition of all list types, compared to older adults. Lastly, we found that semantic, phonological, and hybrid lists produced very similar levels of false recognition across all participants. Future studies will be needed to better understand the interesting differences observed between false recognition and false recall (Watson et al., 2001) of hybrid lists.

[5] Analysis of the non-critical intrusions of younger and older adults in Watson et al. (2001; Table 3) provides some support for this explanation. Whereas semantic and phonological non-critical intrusions on hybrid lists (mean .043 and .083, respectively) were similar to semantic non-critical intrusions on semantic lists (mean .117), phonological non-critical intrusions on phonological lists (mean .250) were somewhat greater. Thus, if related but non-critical intrusions were included with the critical intrusions in the false recall data, levels of phonological false recall would likely be similar to hybrid false recall, although the semantic false recall would still be lower.

References

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 type. *Brain and Language*, *40,* 181–201.

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

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

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.

Berg, L., McKeel, D. W., Miller, J. P., Storandt, M., Rubin, E. H., Morris, J. C., Baty, J., Coates, M., Norton, J., Goate, A. M., Price, J. L., Gearing, M., Mirra, S. S., & Saunders, A. M. (1998). Clinicopathologic studies in cognitively healthy aging and Alzheimer's disease: Relation of histologic markers to dementia severity. *Archives of Neurology*, *55,* 326–335.

Borson, S., & Raskind, M. A. (1997). Clinical features and pharmacologic treatment of behavioral symptoms of Alzheimer's disease. *Neurology*, *48,* S17–24.

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

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.

Budson, A. E., Desikan, R., Daffner, K. R., & Schacter, D. L. (2001). Perceptual false recognition in Alzheimer's disease. *Neuropsychology*, *15,* 230–243.

Budson, A. E., Sitarski, J., Daffner, K. R., & Schacter, D. L. (2002). False recognition of pictures versus words in Alzheimer's disease: The distinctiveness heuristic. *Neuropsychology*, *16,* 163–173.

Cohen, J. B., Wilcox, C. S., & Lerer, B. E. (1991). Development of an "early" detection battery for dementia of the Alzheimer's type. *Progress in Neuropsychopharmacology and Biological Psychiatry*, *15,* 433–479.

Croot, K., Hodges, J. R., Xuereb, J., & Patterson, K. (2000). Phonological and articulatory impairment in Alzheimer's disease: A case series. *Brain and Language*, *75,* 277–309, doi:10.1006/brln.2000.2357.

Craik, F. I. M., & Jennings, J. M. (1992). Human memory. In F. I. M. Craik, & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51–110). Hillsdale, NJ: Erlbaum.

Dalla Barba, G., & Wong, C. (1995). Encoding specificity and intrusion in Alzheimer's disease and amnesia. *Brain and Cognition*, *27,* 1–16, doi:10.1006/brcg.1995.1001.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, *380,* 499–505.

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

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12,* 189–198.

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.

Hodges, J. R., Patterson, K., Ward, R., Garrard, P., Bak, T., Perry, R., & Gregory, C. (1999). The differentiation of semantic dementia and frontal lobe dementia (temporal and frontal varients of frontotemporal dementia) from early Alzheimer's disease: A comparative neuropsychological study. *Neuropsychology*, *13,* 31–40.

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

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, doi:10.1006/jmla.1997.2529.

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.

Koutstaal, W., Verfaellie, M., & Schacter, D. L. (2001). Recognizing identical vs. similar categorically related common objects: Further evidence for degraded gist-represenations in amnesia. *Neuropsychology*, *15,* 268–289.

Mather, M., Johnson, M. K., & De Leonardis, D. M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology*, *16,* 437–458.

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.

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.

Norman, K. A., & Schacter, D. L. (1997). False recognition in young and older adults: Exploring the characteristics of illusory memories. *Memory and Cognition*, *25,* 838–848.

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

Patel, P. G., & Satz, P. (1994). The language production system and senile dementia of the Alzheimer's type: Neuropathological implications. *Aphasiology*, *8,* 1–18.

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.

Price, J. L., & Morris, J. C. (1999). Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. *Annals of Neurology*, *45,* 358–368.

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

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.

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

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., 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, doi:10.1006/jmla.1996.0018.

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

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

Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory and Language*, *40,* 1–24, doi:10.1006/jmla.1998.2611.

Sommers, M. S., & Lewis, B. P. (1999). Who really lives next door: Creating false memories with phonological neighbors. *Journal of Memory and Language*, *40,* 83–108, doi:10.1006/jmla.1998.2614.

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.

Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory and Cognition*, *27,* 494–500.

Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging*, *13,* 230–241.

Verfaellie, M., Keane, M. M., & Johnson, G. (2000). Preserved priming in auditory perceptual identification in Alzheimer's disease. *Neuropsychologia*, *38,* 1581–1592.

Watson, J. M., Balota, D. A., & Roediger III, H. L. (in press). Creating false memories with hybrid lists of semantic and phonological associates: Over-additive false memories produced by converging associative networks. *Journal of Memory and Language*.

Watson, J. M., Balota, D. A., & Sergent-Marshall, S. D. (2001). Semantic, phonological, and hybrid veridical and false memories in healthy adults and in dementia of the Alzheimer's type. *Neuropsychology*, *15,* 254–267.