Neuropsychology
2001, Vol. 15, No. 2, 254–267

0894-4105/01/$5.00 DOI: 10.1037//0894-4105.15.2.254

Semantic, Phonological, and Hybrid Veridical and False Memories in Healthy Older Adults and in Individuals With Dementia of the Alzheimer Type

Jason M. Watson, David A. Balota, and Susan D. Sergent-Marshall
Washington University

Five groups of participants (young, healthy old, healthy old-old, very mild dementia of the Alzheimer type [DAT], and mild DAT) studied 12-item lists of words that converged on a critical nonpresented word (*cold*) semantically (*chill, frost, warm, ice*), phonologically (*code, told, fold, old*), or in a hybrid list of both (*chill, told, warm, old*). The results indicate that (a) veridical recall decreased with age and dementia; (b) recall of the nonpresented items increased with age and remained fairly stable across dementia; and (c) false recall varied by list type, with hybrid lists producing superadditive effects. For hybrid lists, individuals with DAT were 3 times more likely to recall the critical nonpresented word than a studied word. When false memory was considered as a proportion of veridical memory, there was an increase in relative false memory as a function of age and dementia. Results are discussed in terms of age- and dementia-related changes in attention and memory.

One of the most common findings both in healthy older adults, as compared with younger adults, and in individuals diagnosed with dementia of the Alzheimer type (DAT), as compared with age-matched controls, is an impairment in episodic memory performance (see Craik & Jennings, 1992; Nebes, 1992, for reviews). The exploration of age differences in episodic memory has recently included studies of both veridical memory (i.e., memory for information that was presented) and false memory (i.e., memory for information that was not presented). Several studies have demonstrated that, compared with younger adults, healthy older adults have decreased veridical memory and relatively increased false memory (see Balota, Cortese, et al., 1999; Norman & Schacter, 1997; Tun, Wingfield, Rosen, & Blanchard, 1998). In a similar manner, Balota, Cortese, et al. (1999) found that veridical recall decreased from healthy old individuals to individuals with early stage DAT whereas false recall remained fairly stable. In addition, using a repeated study–test procedure, Budson, Daffner, Desikan, and Schacter (2000) found that false recognition (a) consistently increased across study–test trials in individuals with DAT, (b) consistently decreased across study–test trials in young adults, and (c) at first remained constant and then decreased across study–test trials in healthy older adults.

In this study, we extend the scope of the previous studies by using a hybrid cue technique. With this technique, unlike with the Deese–Roediger–McDermott [DRM] paradigm (after Deese, 1959; Roediger & McDermott, 1995), one can investigate false memories for critical items (e.g., *cold*) following the presentation of pure lists of semantic associates (e.g., *chill, frost, warm, ice, shiver, hot, winter, sneeze, freezer, arctic, snow, frigid*), pure lists of phonological associates (e.g., *code, told, fold, old, culled, called, scold, sold, hold, coal, colt, polled*), or hybrid lists of interleaved semantic and phonological associates (e.g., *chill, told, warm, old, shiver, called, winter, sold, freezer, coal, snow, polled*). By obtaining estimates of false recall for the same critical item following both pure and hybrid lists of associates, we hoped to gain a better understanding of how healthy aging and DAT influence false memories and the mechanisms that underlie performance in the hybrid cue technique. We now turn to a brief review of the extant literature regarding the integrity of semantic and phonological networks and attentional control systems in healthy older adults and in individuals with early stage DAT.

Semantic and Phonological Memory Structure and Processes in Healthy Older Adults and in Individuals With Early Stage DAT

A number of researchers have suggested that the representation of concepts and/or the strength of associative connections in semantic networks may be impaired in healthy older adults and individuals with DAT, which could, in turn, explain their relatively poor performance on semantic tasks such as object naming and verbal fluency

Jason M. Watson, David A. Balota, and Susan D. Sergent-Marshall, Department of Psychology, Washington University.

This research was supported by National Institute on Aging Grant P01 AGO3991.

We thank the clinical core at the Washington University Alzheimer's Disease Research Center, who provided the well-specified population for this study; Martha Storandt, who provided the psychometric information; and Michael Cortese, Janet M. Duchek, and Roddy Roediger, who provided help in various aspects of this project.

Correspondence concerning this article should be addressed to David A. Balota, Department of Psychology, Campus Box 1125, Washington University, One Brookings Drive, St. Louis, Missouri 63130-4899. Electronic mail may be sent to dbalota@artsci.wustl.edu.

(Albert, Heller, & Milberg, 1988; Kirshner, Webb, & Kelly, 1984; Ober, Dronkers, Koss, Delis, & Friedland, 1986; Troster, Salmon, McCullough, & Butters, 1989). It has also been suggested that Alzheimer's disease may have a greater influence on the integrity of semantic networks than on the integrity of phonological networks (see Butters, Granholm, Salmon, Grant, & Wolfe, 1987; Mickanin, Grossman, Onishi, Auriacombe, & Clark, 1994; Monsch et al., 1994). Consistent with this argument, Salmon, Heindel, and Lange (1999) recently reported evidence that, compared with healthy older adults, individuals with DAT had greater difficulty generating exemplars from semantic categories (e.g., animals, fruits, vegetables) than generating exemplars from phonemic categories (e.g., words beginning with *f*, *a*, and *s*).

Although it is quite clear that individuals with DAT have difficulty in fluency tasks, the underlying nature of this deficit is somewhat controversial. For example, it is possible that at least some of the deficits attributed to semantic and phonological degradation reflect the attentional demands inherent in particular tasks instead of the memory representations per se (see Balota, Watson, Duchek, & Ferraro, 1999; Nebes, 1989; Ober & Shenaut, 1995). For example, in fluency tasks, a participant must retrieve words on the basis of either a semantic or a phonological cue while also maintaining some memory of previously retrieved words to avoid perseveration. As such, fluency tasks may not be pure measures of the integrity of semantic or phonological memory because a breakdown in performance could be attributed either to degradation of memory or to the attentional control of task goals. To resolve this controversy, some researchers have used priming tasks in which participants are asked to make a simple naming or lexical-decision response to targets (e.g., *cold*) that follow semantically related (e.g., *hot*), phonologically related (e.g., *sold*), or unrelated (e.g., *chair*) primes. The results from these studies have indicated relatively intact semantic and phonological priming in conditions that reflect an automatic spread of activation in semantic or phonological memory networks (e.g., Balota & Watson, 2001), whereas one finds breakdowns both in healthy older adults (Balota, Black, & Cheney, 1992) and in individuals with DAT (Ober & Shenaut, 1995) under priming conditions that demand attentional processes. Of course, even if it were the attentional demands of the fluency tasks that contribute to the observed breakdown in individuals with DAT, this still would not accommodate the differential breakdown observed by Salmon et al. (1999) in semantic and phonological fluency measures.

Attentional Control in Healthy Older Adults and in Individuals With Early Stage DAT

In contrast to the controversy regarding the structure of semantic and phonological memory, there are clear breakdowns in attentional–inhibitory control processes both in healthy older adults (e.g., Duchek, Balota, & Thessing, 1998; Hasher & Zacks, 1988) and in individuals with early stage DAT (e.g., Balota & Faust, in press; Simone & Baylis, 1997a, 1997b; Spieler, Balota, & Faust, 1996; Sullivan, Faust, & Balota, 1995). For example, Spieler et al. found that healthy older adults and individuals with DAT had more difficulty than younger adults inhibiting the highly active, but task-irrelevant, word code in Stroop color naming. Furthermore, Balota and Duchek (1991) and Faust, Balota, Duchek, Gernsbacher, and Smith (1997) found that individuals with DAT were more likely to maintain activation for the inappropriate meaning of an ambiguous word (e.g., the *body* meaning of *organ*) in the presence of biasing context (e.g., *music*), whereas healthy older adults were able to suppress such inappropriate interpretations. Of course, if an automatic spread of activation contributes to false memories following the presentation of lists of semantic or phonological associates (see Balota, Cortese, et al., 1999; Roediger, Balota, & Watson, 2001; Sommers & Lewis, 1999), an age- and/or dementia-related breakdown in attentional–inhibitory control has implications for performance in false memory paradigms. Specifically, avoiding a false memory for a critical item may require one to discriminate multiple sources of activation during retrieval. That is, an intact processing system must be able to differentiate sources of activation that map onto presented items from those that map onto highly activated, but nonpresented, critical items (see Johnson, Hashtroudi, & Lindsay, 1993, for a discussion of reality and source monitoring, and see Multhaup & Balota, 1997, for evidence of source deficits in DAT). It is possible that healthy older adults and individuals with DAT produce some breakdown in the control systems that differentiate overall activation from related list items converging on the critical nonpresented items from the item-specific information concerning the studied item (see Balota, Cortese, et al., 1999, for further discussion of this possibility).

Eliciting False Memories in Healthy Older Adults and in Individuals With Early Stage DAT With the Hybrid Cue Technique

In this study, we obtained estimates of false recall for the same critical item following both pure and hybrid lists of associates (also see Watson, Balota, & Roediger, 2001). In this study, we were able to investigate the effect of age by including three groups of healthy control individuals (young, young-old, and old-old individuals). In addition, we were able to address the influence of DAT for age-matched groups by comparing individuals at three stages of DAT (control individuals, individuals with very mild dementia, and individuals with mild dementia). In this way, we directly addressed whether there are differential effects of age and DAT on the integrity of the semantic and phonological networks that contribute to false memories. Turning to the hybrid lists, there is convergence of semantic and phonological information on the critical nonpresented item (i.e., there is only one word that is semantically related to *frost* and phonologically related to *told*). These lists are particularly intriguing. For example, if only activation processes were invoked, one might not expect greater false memory in the hybrid list as compared with the mean of the phonological and semantic lists, because the hybrid list involves half of the phonological and semantic list cues. In contrast, because the hybrid list uniquely constrains the critical list item, one might expect a superadditive influence of these cues, thereby producing superadditive false memories (see Rubin & Wallace, 1989, for a discussion of semantic and phonological cues in a word generation task).

Method

Participants

A total of 182 participants were included in this study, representing five distinct groups. The first group was composed of 37 young adults, with a mean age of 19.3 years and a range from 17 to 22 years. The second and third groups were healthy older adults. The healthy older adults were divided into a younger group and an older group. There were 46 participants in the young-old group, with a mean age of 71.8 years and a range from 57 to 80 years. The old-old group included 38 participants, with a mean age of 86.3 years and a range from 80 to 96 years. The remaining two groups were composed of individuals with either very mild or mild DAT. There were 43 participants with very mild DAT, with an average age of 75.8 years and a range from 60 to 93 years. There were 18 participants with mild DAT, with an average age of 78.9 years and a range from 67 to 89 years.

The young adults were students at Washington University who either were paid or received course credit for their participation. Older adults and individuals with DAT were recruited from the participant pool at the Alzheimer's Disease Research Center at Washington University. These participants were screened for depression, severe hypertension, possible reversible dementias, and other disorders that could affect cognitive performance.[1] Individuals with DAT were included or excluded on the basis of the criteria set by the National Institute of Neurological and Communications Disorders and Stroke and Alzheimer's Disease and Related Disorders Association (McKhann et al., 1984). Dementia severity was scaled according to the Washington University Clinical Dementia Rating (CDR) scale. The reliability of this scale, as well as the accuracy of the diagnosis by the research team, has been well documented (e.g., Berg et al., 1998). Only those individuals with DAT who were classified as having either very mild dementia (CDR = 0.5) or mild dementia (CDR = 1.0) were included as participants in this study.

Materials

In this study, we used six critical items: *cold, dog, glass, gun, hard,* and *smoke.* Each critical item was chosen for this study because it had a large number of semantic and phonological associates. From this pool of associates, we selected 12 words for the pure list of semantic associates for each critical item and 12 words for the pure list of phonological associates for each critical item. We also created two types of hybrid lists for each critical item by interleaving six semantic and six phonological associates from its corresponding pure lists. The two lists were composed by selecting either the odd-numbered or the even-numbered serial positions in a pure list. In this way, hybrid lists contained a mixture of the same associates that were used in the pure lists. Moreover, by counterbalancing materials across participants, we obtained independent estimates of false recall for the same critical items following semantic, phonological, and hybrid lists of associates. The pure and hybrid lists of associates for each critical item are given in the Appendix.

Procedure

Participants were tested individually and were given a total of six lists of associates, which included two pure lists of semantic associates, two pure lists of phonological associates, and two hybrid lists of interleaved semantic and phonological associates. Lists were blocked together by the type of relationship to the critical item (i.e., semantic, phonological, or hybrid), and the order of these different blocks was counterbalanced across participants. At the beginning of the experiment, participants were told that they would be asked to name words that were presented to them on a computer screen and that they would later be asked to recall these words. On each trial, the following sequence of events occurred: (a) A fixation point (+) was presented in the middle of the computer screen; (b) the screen was blanked; (c) there was a 200-ms intertrial interval; (d) a to-be-named study word was presented in lowercase letters in the middle of the screen; (e) after the participant named the study word, it remained on the screen for 1,500 ms, after which the screen was blanked; (f) Steps c through e were repeated for the remaining words on a list; (g) after all 12 words for a given list had been named, the fixation point reappeared in the middle of the screen; (h) the participants were given as much time as they wished to recall as many items as possible from the previous list of 12 associates; (i) after participants stopped their recall of a given list, the experiment continued; and (j) Steps b through i were repeated for the remaining lists of associates.

Psychometric Test Performance

All of the healthy older adults and the individuals with DAT also participated in a 2-hr battery of psychometric tests designed to assess psychological functions including language, memory, and intelligence. This psychometric battery was given annually as part of an ongoing longitudinal project. Memory was assessed with the Wechsler Memory Scale (WMS; Wechsler & Stone, 1973) Associates Recall and Recognition subscales (paired-associates learning) and the Logical Memory subscale (surface-level story memory). Forward and backward digit span from the WMS were also assessed. Participants received both the Word Fluency Test, in which they named as many words as possible beginning with a specified letter (e.g., *p* or *s*) in a 60-s time period (Thurstone & Thurstone, 1949), and the Animal Naming Test (Goodglass & Kaplan, 1983), in which they named as many animals as possible during a series of six 15-s retrieval periods. Measures of general intelligence were the Information, Block Design, and Digit Symbol subtests of the Wechsler Adult Intelligence Scale (Wechsler, 1955). Visual perceptual–motor performance was assessed by the Benton Copy Test (Benton, 1963) and Trail Making Form A (Armitage, 1945). In the Benton Copy Test, participants must copy a geometric figure; in the Trail Making Form A test, participants connect numerically ordered dots that result in a specific pattern (Armitage, 1945). Participants also received the WMS Mental Control test (Wechsler & Stone, 1973), which evaluates the ability to quickly produce a well-rehearsed letter or digit sequence, such as the alphabet, in a specified amount of time. Participants com-

[1] Of the participants recruited from the Alzheimer's Disease Research Center for our study, 4 young-old participants, 5 participants with very mild DAT, and 1 participant with mild DAT were diagnosed with active mood disorder. Analyses that excluded these individuals produced identical results.

Table 1
Mean Psychometric Performance as a Function of Group

Measure	CDR = 0.0 (Young-old)	CDR = 0.0 (Old-old)	CDR = 0.5 (Very mild DAT)	CDR = 1.0 (Mild DAT)
Word Fluency Test	31.6 (10.6)	31.7 (11.9)	22.5 (8.4)	18.9 (9.3)
Animal Naming Test	21.3 (5.3)	15.4 (5.4)	14.6 (5.4)	7.8 (3.8)
Mental Control	8.0 (1.5)	7.1 (1.7)	6.5 (2.1)	6.1 (2.3)
Associate Memory	15.5 (3.5)	14.5 (3.8)	11.2 (4.1)	7.1 (1.5)
Boston Naming Test	56.6 (4.1)	52.3 (6.4)	49.1 (10.3)	38.4 (12.5)
Logical Memory	10.3 (3.3)	8.0 (2.6)	6.2 (3.8)	2.1 (2.1)
WAIS Information	22.6 (4.5)	21.1 (3.7)	18.1 (5.5)	11.7 (4.4)
Benton Copy Test	9.8 (0.5)	9.8 (0.6)	9.4 (1.0)	8.4 (1.8)
Block Design	33.5 (8.7)	28.8 (7.6)	26.5 (8.0)	17.8 (10.3)
Digit Symbol	50.1 (10.8)	42.2 (8.4)	39.1 (11.9)	25.6 (12.6)
Trail Making Form A	33.8 (9.7)	41.0 (12.1)	49.0 (21.8)	76.3 (26.5)
Digit Span	11.8 (2.0)	11.2 (2.1)	10.8 (1.8)	9.6 (1.7)

Note. Standard deviations are in parentheses. CDR = Washington University Clinical Dementia Rating scale; DAT = dementia of the Alzheimer type; WAIS = Wechsler Adult Intelligence Scale.

pleted the Boston Naming Test (Goodglass, Kaplan, & Weintraub, 1983), which reflects semantic–lexical retrieval processes in naming simple line drawings.

The means and the standard deviations from the psychometric measures are presented in Table 1 as a function of age and DAT. (The young adults did not receive psychometric testing.) A series of one-way analyses of variance (ANOVAs) with group as a between-subjects variable indicated that performance on all measures decreased with increasing age and dementia severity (all $ps < .001$). Thus, as we expected, there were clear changes across groups across a wide range of cognitive measures.

In light of recent arguments by Salmon et al. (1999), and given the importance of semantic and phonological processing for the present arguments, the influence of age and DAT on animal fluency and word fluency measures is worth noting. Salmon et al. suggested that individuals with DAT produce more of a breakdown in semantic fluency measures (e.g., animal fluency) than phonological fluency measures (e.g., word fluency). As shown in the first two rows of Table 1, this pattern partly depends on the level of age and the level of dementia severity. For example, in animal fluency performance, there was a large decrease between healthy young-old and healthy old-old adults, $t(81) = 5.03$, $p < .001$, whereas for word fluency, there was no such effect ($t < 1.00$). However, there was a large change in word fluency between age-matched healthy controls (collapsing across young-old and old-old) and the individuals with very mild DAT, $t(125) = 4.74$, $p < .001$, whereas there was no such change in animal fluency ($t < 1.00$). Finally, comparing the individuals with very mild DAT and the individuals with mild DAT, there was a reliable difference in animal fluency, $t(59) = 4.88$, $p < .001$, but this pattern was only marginal in word fluency, $t(59) = 1.46$, $p < .08$, one-tailed. Thus, the present results indicate that the conclusions that one reaches regarding phonological and semantic fluency measures are sensitive to the age of the individual (young-old vs. old-old) and the dementia severity (very mild vs. mild DAT). This pattern highlights the usefulness of including multiple levels of both age and dementia severity.[2]

Results

In each of the analysis sections, we report the results from two sets of ANOVAs. The first set addresses the influence of healthy aging and includes three groups: healthy young adults, healthy young-old adults, and healthy old-old adults. The second set of ANOVAs addresses the influence of dementia severity with three groups: healthy older adults (collapsed across young-old and old-old adults to equate age across groups), individuals with very mild dementia, and individuals with mild dementia.

Recall Performance

Figure 1 displays the proportion correct serial recall curves for each list type as a function of group. The typical bow-shaped serial position function is most apparent for the young participants. There was an effect of age and of DAT, which appeared to occur most strongly for the primacy and middle serial positions. In addition, it appeared that the semantic lists produced higher recall performance primarily at the primacy positions.

First, we consider the analyses with age as a grouping factor. The results of a 3 (age group) × 3 (list type) × 12 (serial position) mixed-factor ANOVA revealed main effects of age group, $F(2, 118) = 18.43$, $MSE = .27$, $p < .001$; list type, $F(2, 236) = 151.34$, $MSE = .09$, $p < .001$; and serial position, $F(11, 1298) = 93.16$, $MSE = .11$, $p < .001$. Three interactions were reliable, including the interaction between serial position and age group, $F(22, 1298) = 1.70$, $p < .05$; the interaction between list type and serial position, $F(22, 2596) = 3.93$, $MSE = .10$, $p < .001$; and the three-way interaction among all factors, $F(44, 2596) = 1.54$, $p < .05$. The interaction between list type and age group was not reliable ($F < 1$).

Turning to the DAT analysis, the results of the ANOVA with DAT as the grouping factor also yielded main effects of DAT group, $F(2, 142) = 51.81$, $MSE = .35$, $p < .001$; list type, $F(2, 284) = 61.32$, $MSE = .07$, $p < .001$; and serial position, $F(11, 1562) = 111.65$, $MSE = .10$, $p < .001$. Two interactions were reliable, including the interaction between list type and DAT group, $F(4, 284) = 7.48$,

[2] We also conducted a series of correlations between veridical recall, false recall, and psychometric performance. There were not any systematic patterns to these correlations.

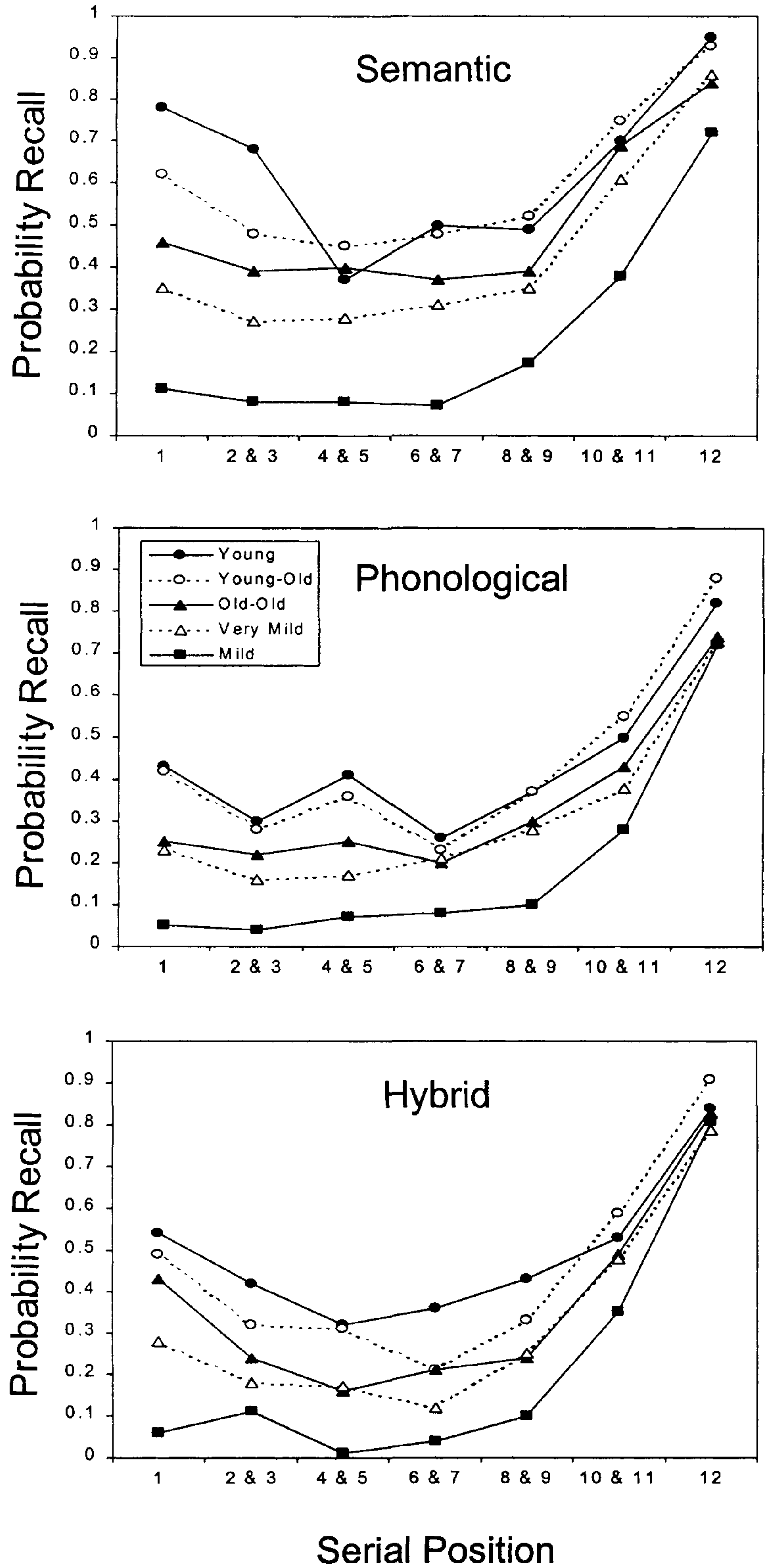

Figure 1. Mean proportion correct recall as a function of serial position, list type, and group. The means of Serial Positions 2 and 3, 4 and 5, 6 and 7, 8 and 9, and 10 and 11 are presented. Very Mild = very mild dementia; Mild = mild dementia.

$p < .001$, and the interaction between list type and serial position, $F(22, 3124) = 2.57$, $MSE = .09$, $p < .001$. The interaction between serial position and DAT group and the three-way interaction were not reliable.

To compare the influence of list type across groups for both veridical and false critical recall, we display both veridical and false critical recall as a function of group in Figure 2.[3] These results indicate that there were differential effects of age and DAT on veridical recall and false recall. First, we consider the effects of age: The results of a 3 (age group) × 3 (list type) × 2 (item type: false vs. veridical) ANOVA yielded main effects of list type, $F(2, 236) = 19.89$, $MSE = .05$, $p < .001$, and item type, $F(1, 118) = 4.15$, $MSE = .09$, $p < .05$. More important, there was a highly reliable interaction between item type and age group, which reflected the opposing influences of age on veridical and false recall. Specifically, there was a reliable 11% drop in veridical recall from the young to the old-old adults, $F(2, 118) = 18.43$, $MSE = .008$, $p < .001$. Post hoc t tests revealed that young (49%) and young-old (46%) adults produced the highest and comparable levels of veridical recall, but this measure dropped significantly ($p < .001$) for the old-old adults (38%). In contrast, false recall increased across groups, $F(2, 118) = 5.31$, $MSE = .05$, $p < .01$. Post hoc tests revealed that the young (31%) and young-old (41%) adults differed marginally ($p < .10$) but that the young-old and old-old adults (47%) did not reliably differ. The difference in false recall from young to old-old was significant ($p < .01$).

The ANOVA with age as a grouping factor also yielded a reliable two-way interaction between list type and item type, $F(2, 236) = 58.37$, $MSE = .06$, $p < .001$. Veridical recall was highest for semantic lists (55%) and dropped significantly (both ps $< .001$) for phonological and hybrid lists (38% and 40%, respectively). False recall, however, was lowest for semantic lists (26%), increased somewhat ($p < .06$) for phonological lists (35%), and then reliably increased ($p < .001$) for hybrid lists (59%). Neither the two-way interaction between list type and age group, $F(4, 236) = 1.22$, nor the three-way interaction among all factors, $F(4, 236) = 1.19$, was reliable.

Turning to the ANOVA with DAT as a grouping factor, there were main effects of DAT group, $F(2, 142) = 10.25$, $MSE = .10$, $p < .001$; list type, $F(2, 284) = 7.97$, $MSE = .07$, $p < .001$; and item type, $F(1, 142) = 25.01$, $MSE = .08$, $p < .001$. More important, there was an interaction between item type and DAT group, $F(2, 142) = 6.65$, $p < .01$, which indicated that there were different effects of DAT on the rates of veridical and false recall. Specifically, veridical recall decreased as a function of DAT, $F(2, 142) = 51.81$, $MSE = .50$, $p < .001$. Healthy older individuals produced the highest level of veridical recall (42%), individuals with very mild DAT were significantly lower ($p < .001$) at 33%, and individuals with mild DAT were significantly lower ($p < .001$) than both groups at 18%. However, healthy older adults (44%), individuals with very mild DAT (45%), and individuals with mild DAT (39%) did not differ in their likelihood of recalling the critical nonpresented word, $F(2, 142) < 1.00$.

The overall ANOVA with DAT as a grouping factor also yielded a reliable interaction between list type and item type, $F(2, 284) = 25.67$, $MSE = .07$, $p < .001$. This interaction was produced by a similar pattern to that found in the age analysis. Specifically, veridical recall was significantly higher for the semantic lists (45%) than for the phonological (31%) or hybrid (32%) lists (both ps $< .001$). False recall, in contrast, was significantly higher for the hybrid lists (60%) than for the semantic (32%) or phonological (39%) lists (both ps $< .001$). Neither the two-way interaction between list type and DAT group, $F(4, 284) = 1.84$, nor the three-way interaction, $F(4, 284) = 1.36$, was significant.

Recall Position of the Critical Nonpresented Items

In addition to analyzing the rate of false and veridical recall for the different groups of participants, we analyzed the output position of the critical item. Because there were different levels of recall for each group, we adjusted this estimate by dividing the output position of the critical nonpresented item by the total number of items recalled on that list. In this way, we provide an estimate of the relative output position of the critical item. Table 2 displays the relative output position of the critical nonpresented item as a function of list type and group. These proportions indicate that there was remarkable consistency in the relative output position for the critical nonpresented items, even though there were marked differences in the number of words recalled. A 3 (list type) × 5 (group) ANOVA revealed no main effects or interaction (all Fs < 1.00). Thus, the relative output position of the critical nonpresented word was about two thirds into the recalled items, independent of group, type of list, and number of items recalled. This pattern is consistent with the analyses reported by others (e.g., Balota, Cortese, et al., 1999; McEvoy, Nelson, & Komatsu, 1999; Roediger & McDermott, 1995).

Noncritical Item Intrusions

It is possible that the relatively high level of critical item intrusions across groups, in comparison to veridical recall, was simply due to an increased likelihood of producing intrusions. To address this possibility, we also analyzed intrusions of noncritical items. Because there are a number of distinct types of intrusions that may reflect different processes, we coded nine different types of intrusions: (a) semantically related but noncritical intrusions; (b) phonologically related but noncritical intrusions; (c) cross-list intrusions from previous semantic lists; (d) cross-list intrusions from previous phonological lists; (e) cross-list intrusions from previous hybrid lists; (f) cross-list critical item

[3] Although it is customary to graph between-group comparisons with bar graphs, we purposefully plotted these data in line graphs to highlight the apparent continuous change in performance across groups. It is possible that some of the changes observed in the healthy older groups may reflect undetected very early stage Alzheimer's disease.

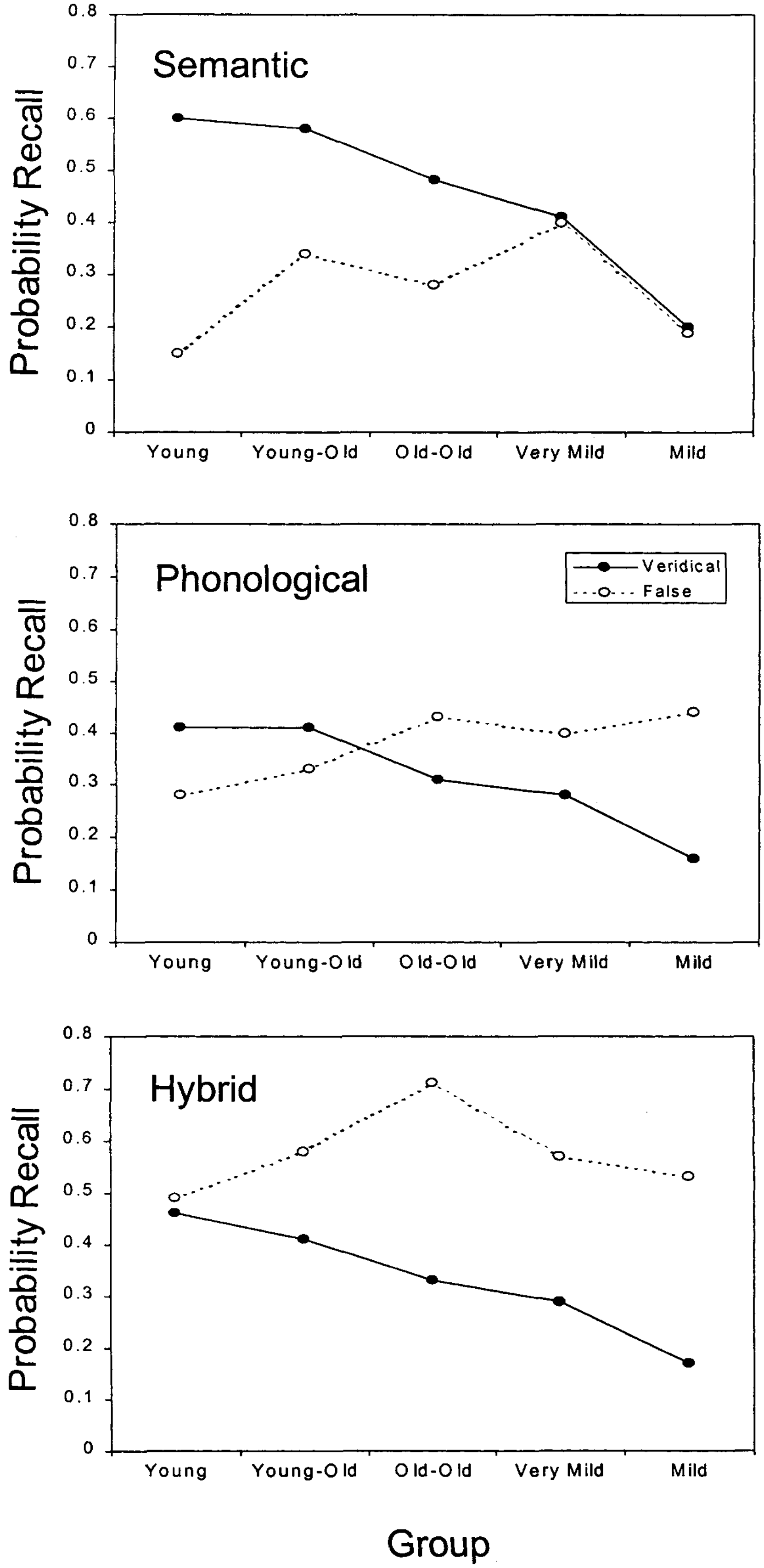

Figure 2. Mean proportion of veridical recall and mean proportion of false recall of the critical nonpresented item as a function of list type and group. Very Mild = very mild dementia; Mild = mild dementia.

Table 2
Relative Output Position for the Critical Item as a Function of List Type and Group

List type	Young	Young-old	Old-old	Very mild DAT	Mild DAT	*M*
Semantic	.75	.80	.59	.70	.83	.72
Phonological	.68	.55	.61	.57	.55	.59
Hybrid	.56	.57	.57	.47	.71	.56
M	.66	.64	.59	.58	.70	

Note. DAT = dementia of the Alzheimer type.

intrusions from previous semantic lists (i.e., recalling the critical item from a previous list); (g) cross-list critical item intrusions from previous phonological lists; (h) cross-list critical item intrusions from previous hybrid lists; and (i) other, apparently unrelated, intrusions.

The average numbers of intrusions per group for each list type are presented in Table 3. The first aspect of this table to note is that, overall, all participants produced very few noncritical intrusions per list. For each of the nine types of intrusions, we conducted separate 3 (list type) $\times$ 5 (group) ANOVAs. However, the results of these analyses should be viewed with some caution because of the low level of intrusions.

First, we consider semantically related intrusions: The analyses revealed that there was a significant main effect of list type, $F(2, 354) = 11.90$, $MSE = 0.04$, $p < .001$. Noncritical semantic intrusions were significantly more likely to occur on pure semantic (0.56) and hybrid (0.32) lists than on pure phonological lists (0.03). The main effect of group was marginally significant, $F(4, 177) = 2.06$, $MSE = 0.01$, $p < .09$, but the interaction between the two factors was not reliable. The findings for the phonologically related intrusions followed a similar pattern. The reliable main effect of list type, $F(2, 354) = 39.95$, $MSE = 0.06$, $p < .001$, revealed that noncritical phonological intrusions were significantly more likely to occur on pure phonological lists (1.27) than on hybrid lists (0.46), and hybrid lists elicited significantly more phonological intrusions than semantic lists (0.07). Once again, there was no main effect of group and no interaction.

The cross-list intrusions (i.e., intrusions of list words from prior lists) provide an indicant of susceptibility to

Table 3
Mean Number of Noncritical Intrusions per List as a Function of List Type, Group, and Intrusion Type

Intrusion type	Young	Young-old	Old-old	Very mild DAT	Mild DAT
	Intrusions on semantic lists				
Semantic	0.09	0.17	0.09	0.15	0.06
Phonological	0.00	0.00	0.04	0.00	0.03
Cross-list–semantic list	0.00	0.03	0.01	0.06	0.03
Cross-list–phonological list	0.00	0.00	0.00	0.00	0.00
Cross-list–hybrid list	0.00	0.01	0.05	0.07	0.03
Cross-critical–semantic list	0.00	0.01	0.00	0.00	0.00
Cross-critical–phonological list	0.00	0.00	0.00	0.00	0.00
Cross-critical–hybrid list	0.00	0.00	0.01	0.05	0.00
Other	0.00	0.04	0.00	0.09	0.08
	Intrusions on phonological lists				
Semantic	0.00	0.00	0.01	0.02	0.00
Phonological	0.22	0.24	0.29	0.30	0.22
Cross-list–semantic list	0.00	0.05	0.08	0.00	0.08
Cross-list–phonological list	0.11	0.17	0.07	0.08	0.03
Cross-list–hybrid list	0.00	0.00	0.08	0.00	0.00
Cross-critical–semantic list	0.00	0.01	0.04	0.00	0.03
Cross-critical–phonological list	0.00	0.01	0.01	0.05	0.00
Cross-critical–hybrid list	0.00	0.00	0.00	0.00	0.00
Other	0.02	0.13	0.03	0.06	0.11
	Intrusions on hybrid lists				
Semantic	0.01	0.04	0.08	0.13	0.06
Phonological	0.07	0.05	0.13	0.10	0.11
Cross-list–semantic list	0.00	0.01	0.00	0.00	0.03
Cross-list–phonological list	0.00	0.11	0.09	0.17	0.00
Cross-list–hybrid list	0.01	0.05	0.02	0.10	0.03
Cross-critical–semantic list	0.00	0.00	0.00	0.00	0.00
Cross-critical–phonological list	0.00	0.02	0.01	0.06	0.00
Cross-critical–hybrid list	0.00	0.01	0.00	0.02	0.03
Other	0.00	0.08	0.02	0.18	0.08

Note. DAT = dementia of the Alzheimer type.

proactive interference. For cross-list intrusions from semantic lists, there was a marginally significant main effect of list type, $F(2, 354) = 2.87$, $MSE = 0.02$, $p < .06$. Cross-list intrusions from previous semantic lists occurred significantly more frequently on phonological (0.21) and semantic (0.13) lists than on hybrid lists (0.04). There was no main effect of group and no reliable interaction between list type and group. For cross-list intrusions from phonological lists, again, the main effect of list type was reliable, $F(2, 354) = 8.20$, $MSE = 0.05$, $p < .001$. The main effect of group approached significance, $F(4, 177) = 2.06$, $MSE = 0.02$, $p < .09$, as did the interaction between the two factors, $F(8, 354) = 1.93$, $p < .06$. Cross-list intrusions from prior phonological lists never occurred on semantic lists for any of the groups; however, the frequency of occurrence of this type of intrusion differed greatly across groups for hybrid and phonological lists. Finally, the analyses of cross-list intrusions from previous hybrid lists revealed no reliable effects (all Fs < 1.00).

As shown in Table 3, critical item intrusions from prior semantic lists rarely occurred, and the analyses revealed that the main effect of list type approached significance, $F(2, 354) = 2.31$, $MSE = 0.005$, $p = .10$, but neither the main effect of group nor the interaction was reliable (both Fs < 1.00). The cross-list critical intrusions for the phonological lists did reveal a significant effect of group, $F(4, 177) = 2.75$, $MSE = 0.003$, $p < .05$. These intrusions occurred significantly more often for the individuals with very mild DAT (0.04) than for any of the other groups. Neither the main effect of list type nor the interaction was reliable. Cross-list critical intrusions for the hybrid lists yielded a similar pattern, $F(1, 177) = 2.18$, $MSE = 0.002$, $p < .08$, with the individuals with very mild DAT (0.02) producing more of these intrusions than the other groups. Again, the main effect of list type and the interaction were not reliable. Finally, analyses of "other" intrusions revealed a significant main effect of group, $F(4, 176) = 4.58$, $MSE = 0.02$, $p < .01$. Young-old adults (0.08), individuals with very mild DAT (0.11), and individuals with mild DAT (0.09) produced more "other" intrusions than did old-old (0.02) and young (0.01) adults. Neither the main effect of list type nor the interaction was reliable.

Naming Performance

An analysis was also conducted on mean naming response latency during the encoding task. This analysis yielded a main effect of list type, with words from phonological lists taking slightly longer to name (868 ms) than words from either hybrid or semantic lists (843 ms and 840 ms, respectively), $F(2, 354) = 4.05$, $MSE = 15{,}713.37$, $p < .02$. There was also a main effect of group, with longer response latencies associated with increasing age and increasing DAT, $F(4, 177) = 12.15$, $MSE = 100{,}499.70$, $p < .001$. Young adults named list words at an average of 689 ms per word, young-old adults at 833 ms, old-old adults at 916 ms, individuals with very mild DAT at 884 ms, and individuals with mild DAT at 1,007 ms. The interaction between the two factors was not reliable ($F < 1.00$).

Discussion

This experiment yielded three main findings. First, veridical recall decreased in healthy older adults as compared with healthy young adults, whereas false recall increased as a function of age. Second, veridical recall decreased in individuals with DAT as compared with age-matched healthy older adults, whereas false recall remained fairly stable. Third, compared with pure semantic and pure phonological lists, hybrid lists produced higher false recall of the critical nonpresented word for all groups. We now turn to a discussion of the implications of these results for the understanding of memory changes that occur in both healthy older adults and individuals with DAT and the mechanisms that underlie false memories in the hybrid cue technique.

Veridical and False Memory in Healthy Aging and in DAT Performance

The results from the healthy young adults and older adults replicate and extend the findings of Balota, Cortese, et al. (1999), Norman and Schacter (1997), and Tun et al. (1998). Specifically, following pure lists of semantic associates, veridical recall decreased as a function of age, whereas false recall increased as a function of age. This pattern was also produced by healthy older adults following pure lists of phonological associates and hybrid lists of interleaved semantic and phonological associates. These results suggest that an increased susceptibility to false memories in older adults clearly extends beyond the semantic associates used in the DRM lists (also see Bartlett, Strater, & Fulton, 1991; Cohen & Faulkner, 1986; Koutstaal & Schacter, 1997; Rankin & Kausler, 1979; Schacter, Israel, & Racine, 1999).

Because there are large differences across groups in veridical recall, one might argue that it is not appropriate simply to compare groups on absolute levels of false recall. Hence, it might be more appropriate to measure false recall in relation to overall veridical recall. The notion here is that if an individual does not recall many words, then there is (a) little likelihood of recalling the critical nonpresented item, simply due to sheer number of retrieved words; (b) little support from other items in memory for the critical nonpresented item; and (c) a decreased likelihood of related words influencing the recall of the critical nonpresented item during retrieval. Rather than comparing young and old adults on their absolute amount of false recall, one can adjust for age differences in episodic memory performance by dividing participants' false recall by their veridical recall. These data are displayed in Figure 3, along with the data from Balota, Cortese, et al.'s (1999) study of false recall with semantically related lists, for comparison purposes. There are three points to note in Figure 3. First, the results from the present semantic lists produced performance remarkably similar to the results from Balota, Cortese, et al.'s study, even though the earlier study used items from the DRM lists. In addition, focusing on only the results from the

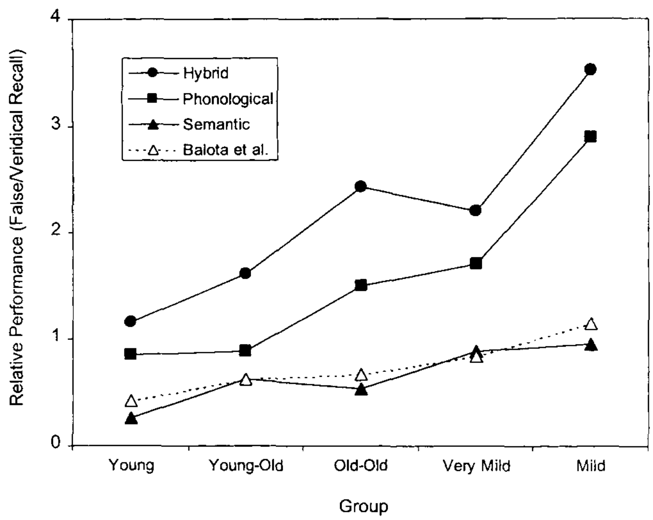

Figure 3. Mean proportion of relative false recall (false/veridical) as a function of list type and group. The dashed line indicates data from the semantic lists in Balota, Cortese, et al.'s (1999) study. Very Mild = very mild dementia; Mild = mild dementia.

present study, there was a clear increase in the relative false memory both as a function of age, $F(2, 118) = 11.69$, $MSE = 1.32$, $p < .001$, and as a function of DAT severity, $F(2, 140) = 13.02$, $MSE = 2.29$, $p < .001$. There is also evidence of an Age Group × List Type interaction, $F(4, 236) = 3.22$, $MSE = 1.08$, $p < .05$, reflecting the larger age-related increase in the relative false memory in the hybrid lists than in the phonological and semantic lists. The DAT × List Type interaction did not approach significance, $F(4, 280) = 1.25$, $MSE = 2.46$, $p = .29$, and hence, DAT does not appear to modulate the sensitivity to semantic and/or phonological information in producing false memories.

The relative increase in false recall across groups may not be due to overall changes in episodic memory but may be a reflection of changes in the efficiency of an attentional control system that maintains the current task demands, that is, episodic retrieval of specific list items (see Balota, Cortese, et al., 1999). When given a list of semantic and/or phonological associates, a participant needs to discriminate between two different sources of information: (a) highly activated convergence on a critical nonpresented word and (b) item-specific information due to initial encoding. Because these two sources of activation may compete for output during retrieval, an efficient attentional control system needs to be able to differentiate sources of activation that map onto studied items from those that map onto nonpresented critical items. In this light, if there is a group-related breakdown in attentional control systems (i.e., the system that maintains the integrity of the critical task set), then healthy older adults, and to a greater extent individuals with DAT, may be less likely to correctly discriminate between these sources of activation, thereby increasing the likelihood of producing false memories following lists of semantic and/or phonological associates. As we noted in the introduction, there is clear evidence of attentional breakdowns in these groups, and yet there is also evidence of intact activation processes as reflected by semantic and phonological priming tasks (Balota & Duchek, 1991; Balota & Watson, 2001; Duchek & Balota, 1993; Nebes, 1989; Shenaut & Ober, 1996). Thus, there should be activation for nonpresented critical items that competes for output against context-specific information for studied items, but there may be a breakdown in the attentional system that needs to discriminate between sources of information.

Although we have been discussing the present results within an attentional control framework, there is a relatively simple alternative account. Specifically, it is possible that these results may reflect an imbalance in the strength of the activation for nonpresented critical items and context-specific information for studied items. Both healthy older adults and individuals with DAT produced large decrements in episodic memory as reflected by their veridical recall performance. Nonpresented critical words may be more likely to intrude in recall performance simply because there is an imbalance in the strength of the item-specific information and the overall familiarity for the nonpresented items.

To address this alternative account of the present results, we collapsed across the three different types of lists to obtain a single veridical recall estimate and a single false recall estimate for each participant. We then conducted a regression analysis in which we predicted overall false recall, via participant group, after overall veridical recall was partialed out. This analysis indicated that participant group reliably predicted false recall after veridical recall was entered in a hierarchical multiple regression, $t(179) = 2.88$, $p < .005$. Thus, there was an overall increase in false recall across groups even after veridical recall had been partialed out. In this light, it is interesting to note that Schacter, Verfaellie, and Anes (1997) reported that individuals with severe item-specific breakdowns (i.e., mostly individuals with amnesia due to Korsakoff's syndrome) produced decreased semantic, orthographic, and/or phonological false recognition. Schacter et al.'s results, coupled with the present regression analysis, suggest that overall changes in memory cannot totally accommodate the relative increase in the false memories in healthy aging and in individuals with early stage DAT.

Implications of the Present Results for Understanding the Mechanisms Underlying False Memories in the Hybrid Cue Technique

The present results have implications for understanding not only age-related and dementia-related changes in memory performance but also the mechanisms that underlie false memories elicited by the hybrid cue technique. Specifically, the present pure-list results suggest a role for activation within semantic and phonological networks that converges on a nonpresented critical item (Balota, Cortese, et al., 1999; McDermott & Watson, in press; Roediger, Balota, & Robinson, 2001; Roediger, Balota, & Watson, 2001; Seamon, Luo, & Gallo, 1998; Sommers & Lewis, 1999). In this same light, the present results suggest that one clearly needs to make some additional assumptions regarding a spreading activation account. Specifically, the present design allows for an interesting test of simple additivity of semantic and phonological influences on false recall. That is, if false recall simply reflects activation processes from semantic and phonological lists, then the hybrid list should produce the same level of false recall as the mean of the phonological and semantic lists. This prediction follows from the fact that the hybrid list was made up of the same number of words, with half being from the semantic list and half being from the phonological list. The results of a series of *t* tests indicated that for every group of participants, the hybrid list produced significantly higher false recall than the mean estimate from the semantic and phonological lists (all $ps \leq .05$), with no reliable interactions across groups. Thus, the present results clearly indicate superadditive influences of semantic and phonological information in the hybrid lists.

There are at least two ways to accommodate the superadditive influences of semantic and phonological information. First, this pattern may be produced by the direction of attention to a critical item by convergent meaning and rhyme cues in the context of a hybrid list. That is, because phonology and semantics are usually uncorrelated in language (i.e., the linguistic universal referred to as arbitrariness), once one provides multiple constraints from both phonology and semantics, the word is uniquely constrained. This convergence of activation may direct attention to the unique code, which then produces a greater encoding event (see Rubin & Wallace, 1989, for similar arguments). However, if this were the case, then one might expect a decreased false memory effect in the hybrid list for the healthy older adults and for individuals with DAT, given the documented attentional declines in these groups.

Second, it is also possible that this superadditive pattern may be a reflection of activation patterns in distinct lexical representations that are at different points in sensitivity in the activation functions. Consistent with this notion, Balota, Pollatsek, and Rayner (1986) and Becker and Killion (1977) reported superadditive interactions in eye fixation durations during reading and primed lexical decisions, respectively, when there are multiple constraints from semantic context and visual information. If there are distinct orthographic, semantic, and phonological lexical representations that accumulate activation (see, e.g., Besner, 1990), it is possible that a mixture of codes produces the most effective activation pattern. More specifically, consider the possibility that the activation function is at a peak sensitivity range with half of the list items (e.g., six words per list). After this level, there is a decreasing influence of additional words within the same domain because the activation is approaching asymptote for that representation. Thus, the benefit of six additional words within the same code in the pure lists does not produce a maximal increase in activation. In contrast, for the hybrid list conditions, the activation functions are in the maximal sensitivity range for both the phonological and semantic representations. Hence, when these two sources combine, one finds a superadditive influence on performance. Although this account is clearly speculative, the present results indicate that this superadditive pattern occurs across groups of individuals with a wide range of cognitive performance. It is also interesting to note that the present superadditive memory errors are consistent with the predictions from an interactive activation model of speech production developed by Dell and O'Seaghdha (1992). Their model nicely predicts that semantic and phonological information can combine to produce the observed superadditive influences on speech errors, for example, substituting the word *start* for *stop* on the basis of semantic and phonological features.

Finally, the present results also have implications for the notion that the act of retrieval provides some support for false recall. That is, one might argue that during the process of generating words at retrieval, the participant actually generates the critical item because of previously generated related words. Once the critical item is generated, it is likely to be output later because of a source error. In addition, one might argue that in outputting related words there is activation spreading from these words to the critical nonpresented item. However, as noted by Balota, Cortese, et al. (1999), the present type of between-group comparisons places some constraints on this interpretation. Specifically, one of the most striking aspects of these data is that although veridical recall decreased dramatically from young adults (50%) to individuals with DAT (18%), false recall actually increased. Hence, the act of recalling related words that converge on a critical nonpresented word clearly cannot accommodate the present results.

Conclusion

The present results indicate that the relative incidence of false memories to veridical memories increases as a function of both age and DAT. We have argued that the present results are consistent with the notion that there are breakdowns in attentional control systems that select among activated pathways due to spreading activation and item-specific information due to memory encoding. These results also suggest that there is clear sensitivity to multiple phonological and semantic constraints even in individuals with mild DAT. Finally, turning to the implications of the present results for mechanisms underlying false memories with the hybrid cue technique, the present superadditive results suggest that the convergence of independent dimensions of words (e.g., semantics and phonology) affords a heightened susceptibility to false memories.

References

Albert, M. S., Heller, H. S., & Milberg, W. (1988). Changes in naming ability with age. *Psychology and Aging, 3,* 173–178.

Armitage, S. G. (1945). An analysis of certain psychological tests used for the evaluation of brain injury. *Psychological Monographs, 60*(1, Whole No. 177), 1–48.

Balota, D. A., Black, S., & Cheney, M. (1992). Automatic and attentional processing in young and old adults: A reevaluation of the two-process model of semantic priming. *Journal of Exper-*

imental Psychology: Human Perception and Performance, 18, 485–502.

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

Balota, D. A., & Faust, M. E. (in press). Attention in dementia of the Alzheimer's type. In F. Boller & S. Cappa (Eds.), *Handbook of neuropsychology* (Vol. 6). Amsterdam: Elsevier Science.

Balota, D. A., Pollatsek, A., & Rayner, K. (1986). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17,* 364–390.

Balota, D. A., & Watson, J. M. (2001). *Semantic and phonological priming in young, healthy older adults, and in individuals with dementia of the Alzheimer's type.* Manuscript in preparation.

Balota, D. A., Watson, J. M., Duchek, J. M., & Ferraro, R. F. (1999). Cross-modal priming with ambiguous and unambiguous words in young, healthy older adults, and in individuals with dementia of the Alzheimer's type: Explorations of semantic memory. *Journal of the International Neuropsychological Society, 5,* 626–640.

Bartlett, J. C., Strater, L., & Fulton, A. (1991). False recency and false fame of faces in young adulthood and old age. *Memory & Cognition, 19,* 177–188.

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 389–401.

Benton, A. L. (1963). *The revised visual retention test: Clinical and experimental applications.* New York: Psychological Corporation.

Berg, L., McKeel, D. W., Jr., Miller, J. P., Storandt, M., Rubin, E. H., Morris, J. C., Baty, J., Coats, 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.

Besner, D. (1990). Does the reading system need a lexicon? In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 73–96). Hillsdale, NJ: Erlbaum.

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.

Butters, N., Granholm, E. L., Salmon, D. P., Grant, I., & Wolfe, J. (1987). Episodic and semantic memory: A comparison of amnesic and demented patients. *Journal of Clinical and Experimental Neuropsychology, 9,* 479–497.

Cohen, G., & Faulkner, D. (1986). Memory for proper names: Age differences in retrieval. *British Journal of Developmental Psychology, 4,* 187–197.

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.

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

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42,* 287–314.

Duchek, J. M., & Balota, D. A. (1993). Sparing of activation processes in older adults. In J. Cerella & W. Hoyer (Eds.), *Adult information processing: Limits on loss* (pp. 383–406). San Diego, CA: Academic Press.

Duchek, J. M., Balota, D. A., & Thessing, V. C. (1998). Inhibition of visual and conceptual information during reading in healthy aging and Alzheimer's disease. *Aging, Neuropsychology, and Cognition, 5,* 169–181.

Faust, M. E., Balota, D. A., Duchek, J. M., Gernsbacher, M. A., & Smith, S. (1997). Inhibitory control during sentence comprehension in individuals with dementia of the Alzheimer type. *Brain and Language, 57,* 225–253.

Goodglass, H., & Kaplan, E. (1983). *Boston Diagnostic Aphasia Examination booklet, III, Oral Expression, J. Animal Naming-(Fluency in Controlled Association).* Philadelphia: Lea & Febiger.

Goodglass, H., Kaplan, E., & Weintraub, S. (1983). *Boston Naming Test scoring booklet.* Philadelphia: Lea & Febiger.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 193–225). San Diego, CA: Academic Press.

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

Kirshner, H. S., Webb, W. G., & Kelly, M. P. (1984). The naming disorder of dementia. *Neuropsychologia, 22,* 23–30.

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.

McDermott, K. B., & Watson, J. M. (in press). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory and Language.*

McEvoy, C. L., Nelson, D. L., & Komatsu, T. (1999). What's the connection between true and false memories? The differential roles of interitem associations in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1177–1194.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS–ADRDA work group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology, 34,* 934–939.

Mickanin, J., Grossman, M., Onishi, K., Auriacombe, S., & Clark, C. (1994). Verbal and nonverbal fluency in patients with probable Alzheimer's disease. *Neuropsychology, 8,* 385–394.

Monsch, A. U., Bondi, M. W., Paulsen, J. S., Brugger, P., Butters, N., Salmon, D. P., & Swenson, M. (1994). A comparison of category and letter fluency in Alzheimer's disease and Huntington's disease. *Neuropsychology, 8,* 25–30.

Multhaup, K. S., & Balota, D. A. (1997). Generation effects and source memory in healthy older adults and individuals with dementia of the Alzheimer type. *Neuropsychology, 11,* 382–391.

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

Nebes, R. D. (1992). Cognitive dysfunction in Alzheimer's disease. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 373–446). Hillsdale, NJ: Erlbaum.

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

Ober, B. A., Dronkers, N. F., Koss, E., Delis, D. C., & Friedland, R. P. (1986). Retrieval from semantic memory in Alzheimer-type dementia. *Journal of Clinical and Experimental Neuropsychology, 8,* 75–92.

Ober, B. A., & Shenaut, G. K. (1995). Semantic priming in Alzheimer's disease: Meta-analysis and theoretical evaluation. In P. A. Allen & T. R. Bashore (Eds.), *Age differences in word and language processing* (pp. 247–271). Amsterdam: Elsevier.

Rankin, J. S., & Kausler, D. H. (1979). Adult age differences in false recognition. *Journal of Gerontology, 34*, 58–65.

Roediger, H. L., Balota, D. A., & Robinson, K. J. (2001). *The automatic mechanisms of false memory.* Manuscript in preparation.

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.

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.

Rubin, D. C., & Wallace, W. T. (1989). Rhyme and reason: Analyses of dual retrieval cues. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 698–709.

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

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

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

Shenaut, G. K., & Ober, B. A. (1996). Methodological control of semantic priming in Alzheimer's disease. *Psychology and Aging, 11*, 443–448.

Simone, P. M., & Baylis, G. C. (1997a). The role of attention in a spatial memory task in Alzheimer disease patients. *Alzheimer Disease and Associated Disorders, 11*, 140–152.

Simone, P. M., & Baylis, G. C. (1997b). Sustained attention in a reaching task: Effects of normal aging and Alzheimer's disease. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 595–608.

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.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in younger adults, healthy older adults, and individuals with senile dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 461–479.

Sullivan, M. P., Faust, M. E., & Balota, D. A. (1995). Identity negative priming in older adults and individuals with dementia of the Alzheimer type. *Neuropsychology, 9*, 537–555.

Thurstone, L. E., & Thurstone, T. G. (1949). *Examiner manual for the SRT Primary Mental Abilities.* Chicago: Science Research.

Troster, A. I., Salmon, D. P., McCullough, D., & Butters, N. (1989). A comparison of category fluency deficits associated with Alzheimer's and Huntington's disease. *Brain and Language, 37*, 500–513.

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.

Watson, J. M., Balota, D. A., & Roediger, H. L. (2001). *The role of semantic and phonological activation in the creation of false memories.* Manuscript in preparation.

Wechsler, D. (1955). *WAIS manual.* New York: Psychological Corporation.

Wechsler, D., & Stone, C. P. (1973). *Manual: Wechsler Memory Scale.* New York: Psychological Corporation.

Appendix

Pure and Hybrid Lists of Associates for Each Critical Item Used in Study and Free Recall

Semantic Lists

cold: chill, frost, warm, ice, shiver, hot, winter, sneeze, freezer, Arctic, snow, frigid

dog: hound, puppy, Lassie, mutt, pet, beware, cat, paw, animal, poodle, flea, bark

glass: bottle, jar, shatter, fragile, mirror, lens, crystal, prism, break, cup, mug, window

gun: pistol, revolver, holster, bang, bullet, rifle, military, powder, shoot, trigger, murder, weapon

hard: rigid, difficult, easy, concrete, cement, tough, stiff, simple, complex, firm, solid, soft

smoke: nicotine, puff, cigar, tobacco, pipe, fire, fumes, chimney, ashtray, cigarette, habit, Marlboro

Phonological Lists

cold: code, told, fold, old, culled, called, scold, sold, hold, coal, colt, polled

dog: log, doll, dug, frog, bog, dig, daub, hog, dock, dawn, fog, jog

glass: class, pass, brass, sass, glaze, grass, bass, lass, mass, gas, gloss, glance

gun: pun, done, gum, gone, one, bun, gut, gush, run, ton, gain, fun

hard: bard, herd, harm, yard, charred, hark, hoard, tarred, card, hired, harp, heart

smoke: smack, smock, cloak, yoke, stroke, stoke, smote, smirk, choke, oak, soak, spoke

Hybrid Lists (Semantic–Phonological)

cold: chill, told, warm, old, shiver, called, winter, sold, freezer, coal, snow, polled

dog: hound, doll, Lassie, frog, pet, dig, cat, hog, animal, dawn, flea, jog

glass: bottle, pass, shatter, sass, mirror, grass, crystal, lass, break, gas, mug, glance

gun: pistol, done, holster, gone, bullet, bun, military, gush, shoot, ton, murder, fun

hard: rigid, herd, easy, yard, cement, hark, stiff, tarred, complex, hired, solid, heart

smoke: nicotine, smock, cigar, yoke, pipe, stoke, fumes, smirk, ashtray, oak, habit, spoke

Hybrid Lists (Phonological–Semantic)

cold: code, frost, fold, ice, culled, hot, scold, sneeze, hold, Arctic, colt, frigid

dog: log, puppy, dug, mutt, bog, beware, daub, paw, dock, poodle, fog, bark

glass: class, jar, brass, fragile, glaze, lens, bass, prism, mass, cup, gloss, window

gun: pun, revolver, gum, bang, one, rifle, gut, powder, run, trigger, gain, weapon

hard: bard, difficult, harm, concrete, charred, tough, hoard, simple, card, firm, harp, soft

smoke: smack, puff, cloak, tobacco, stroke, fire, smote, chimney, choke, cigarette, soak, Marlboro

Received July 24, 2000
Revision received November 3, 2000
Accepted November 5, 2000 ■