COGNITIVE NEUROPSYCHOLOGY, 1999, 16 (3/4/5), 361–384

VERIDICAL AND FALSE MEMORIES IN HEALTHY OLDER ADULTS AND IN DEMENTIA OF THE ALZHEIMER'S TYPE

David A. Balota, Michael J. Cortese, Janet M. Duchek, David Adams, Henry L. Roediger III, Kathleen B. McDermott, and Benjamin E. Yerys
Washington University, St. Louis, USA

Five groups of participants (young, healthy old, healthy old-old, very mild Dementia of the Alzheimer's Type, Mild Dementia of the Alzheimer's Type) studied and were tested on six 12-item lists of words selected from the DRM (Deese, 1959; Roediger & McDermott, 1995) materials. These lists of words strongly converged semantically on a nonpresented critical word. The results indicated that both veridical recall and recognition performance decreased both as a function of age of the participants and as a function of dementia severity. However, the recall and recognition of the highly related nonpresented items actually increased as a function of age, and only slightly decreased as a function of DAT. When false memory was considered as a proportion of veridical memory, there was a clear increase as a function of both age of the participants and as a function of disease severity. The results are discussed in terms of (a) age and DAT-related changes in attention and memory performance, and (b) the underlying mechanisms that produce false memories in the DRM paradigm.

INTRODUCTION

One of the most common findings in both healthy older adults and in early stages of Dementia of the Alzheimer's Type (DAT) is a breakdown in veridical episodic memory performance (see Craik & Jennings, 1992; Nebes, 1994, for reviews). The present study will explore not only veridical memory in these populations, but also memory for information that was not presented, i.e. false memories. In particular, we will explore the false memory paradigm recently developed by Roediger and McDermott (1995) based on an initial study by Deese (1959), hereafter called the DRM paradigm, in which associatively/semantically related words converge on a critical nonpresented item. For example, consider being presented with the following list of words: THREAD, PIN, EYE, SEWING, SHARP, POINT, PRICK, THIMBLE, HAYSTACK, PAIN, HURT, INJECTION. The nonpresented critical target word is NEEDLE in this list. Roediger and McDermott reported that the likelihood of recalling the nonpresented critical word NEEDLE was as high as the probability of recalling the items that were actually presented in the middle of the list (.40 in both cases). The issues addressed in the present study are the manner in which healthy ageing and

---

Requests for reprints should be addressed to David A. Balota, Department of Psychology, Washington University, St. Louis, MO 63130, USA (E-mail: dbalota@artsci.wustl.edu).

This work was supported by Grant NIA PO1 AGO3991. Thanks are extended to the clinical core at the Washington University ADRC who provided the well-specified population for this study, Martha Storandt who provided the psychometric information, and Beth Ober, Patricia Tun, and Dan Schacter for helpful comments on an earlier version of this manuscript.

DAT influence both veridical and false memories, and the mechanisms that underlie performance in the DRM paradigm.

The present study was motivated by a series of experiments that provide evidence regarding the relative integrity and disruption of specific cognitive systems (e.g. semantic memory and attentional systems) in both healthy older adults compared to younger adults, and in DAT individuals compared to age-matched healthy control individuals. We will argue that the production of false memories in older individuals and in DAT individuals may reflect a tradeoff in the integrity of these cognitive systems, which make these populations particularly sensitive to false memories. We shall now turn to a brief review of the extant literature regarding both semantic memory systems and attentional control systems in both healthy older adults and in early stage DAT individuals.

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

In the DRM paradigm, participants are presented associatively/semantically related words that converge on a critical nonpresented target word. Of course, if there is a degradation in semantic memory in healthy older adults and/or in DAT individuals, then one might expect a decreased susceptibility to false memories. Unfortunately, the evidence regarding changes in semantic memory in both healthy older adults and in early stage DAT individuals has been somewhat controversial. For example, evidence from tasks such as verbal fluency and object naming has been taken to indicate that there is a breakdown in semantic memory structures in these populations (e.g. Albert, Heller, & Milberg, 1988; Kirshner, Webb, & Kelly, 1984; Ober, Dronkers, Koss, Delis, & Friedland, 1986; Troster, Salmon, McCullough, & Butters, 1989). However, researchers have argued that the deficits in these tasks are not simply a reflection of the integrity of semantic memory, but rather, reflect the attentional demands engaged by these tasks (see Balota, Watson, Duchek, & Ferraro, in press; Nebes, 1989; Ober & Shenaut, 1995). For example, in the verbal fluency task, the participant must generate words that begin with a specific letter and also maintain some memory of previously generated words. A breakdown in performance in this task could be due to the semantic integrity of the system, the attentional control of task goals, and/or the maintenance of previously produced words. Clearly, this task is not simply a pure measure of semantic memory structure.

In order to minimise attentional demands, researchers have turned to semantic priming tasks in which participants are asked to make simple naming or lexical decisions to targets (doctor) that follow either related (nurse) or unrelated (chair) primes. The research in the area is relatively clear: Specifically, one finds relatively intact priming effects under conditions that reflect automatic spread of activation within the memory network, whereas one finds evidence of breakdowns in both healthy older adults (Balota, Black, & Cheney, 1992) and in early stage DAT individuals (Ober & Shenaut, 1995), under conditions that reflect attentional processes. Thus, researchers have argued that the integrity of the network and automatic access to information embedded in this network are relatively intact in both healthy older adults, and in early stage DAT, and it is the attentional systems that appear to break down.

Attentional Breakdown in Healthy Ageing and in DAT Individuals

In contrast to the automatic activation of semantic pathways, recent studies indicate that there is a breakdown in processes related to inhibitory/attentional control in both healthy older adults (e.g. Duchek, Balota, & Thessing, in press; Hasher & Zacks, 1988) and in early stage DAT individuals (e.g. Spieler, Balota, & Faust, 1996; Sullivan, Faust, & Balota, 1995). It is important to note that these studies have yielded *disproportionate* breakdowns (greater than expected by overall differences in speed or accuracy), and/or qualitatively distinct patterns predicted by the inhibitory framework. For example, Spieler et al. (1996) provided evidence that there is a disproportionate breakdown in the ability to inhibit the word code when naming colours in the Stroop task in both healthy older adults, compared to young, and in DAT individuals. In addition, Balota and Duchek (1991) demonstrated that DAT individuals are more likely to keep active inappropriate meanings of ambiguous words (e.g. the BODY meaning of ORGAN) in the presence of context (e.g. MUSIC), whereas healthy older individuals appear to suppress such inappropriate interpretations (also see Faust, Balota, Duchek, Gernsbacher, & Smith, 1997). These results, coupled with the earlier cited priming studies, suggest that the activation of relevant networks is not impaired in healthy older adults or in early stages of DAT, but it is the attentional control or inhibition of partially activated pathways that appears to be compromised.

What are the implications of these observations for the DRM false memory paradigm? False memories of the critical nonpresented items in the DRM paradigm could be viewed as a situation wherein there are multiple sources of activation that need to be discriminated to make the correct response. An intact processing system should be able to discriminate between sources of activation that are directly related to what was presented compared to sources that are only indirectly related, but highly activated (see Johnson, Hashtroudi, & Lindsay, 1993, for discussion of reality and source monitoring, and Multhaup & Balota, 1997, for evidence of source deficits in DAT individuals). Based on the evidence reviewed, one might suggest that healthy older adults and DAT individuals may have difficulty controlling or inhibiting the highly activated (semantically related) nonpresented item. In this light, it is possible that a breakdown in attentional control may actually produce increased false memories. If both healthy older adults and early stage DAT individuals have intact semantic structures that produce a boost in the semantic activation for nonpresented critical items, coupled with a breakdown in attentional control systems, one might actually expect an increased likelihood of false memories in these populations. Of course, because of the group differences in memory performance, this pattern would be most likely reflected in *relative* measures in which one takes into account overall changes in veridical memory across these populations.

Norman and Schacter (1997) have recently reported evidence that healthy older adults are indeed *more* susceptible to false memories than healthy young adults (also see Tun, Wingfield, Rose, & Blanchard, 1998). Specifically, although veridical recall was lower in older adults compared to younger adults, false recall was actually higher in older adults compared to younger adults. Norman and Schacter suggested that the increased false memories in the DRM paradigm is quite consistent with age-related decreased efficiency of frontal lobe functioning. There are a number of intriguing lines of converging evidence regarding the role of frontal lobes in the creation of false memories. First Schacter, Reiman, & Curran et al. (1996b) reported evidence in a recent PET study that there is increased activation in regions of prefrontal cortex when critical nonpresented lures were presented compared to the presentation of studied words. Norman and Schacter suggested that this pattern may reflect the frontal involvement in "resisting" or opposing illusory memory for the critical lures. Obviously, this is consistent with the present view that healthy older adults and early stage DAT individuals may have some difficulty inhibiting or controlling partially activated memory representations. Second, Schacter, Curran, Galluccio, Milberg, and Bates (1996a) have recently reported evidence that an individual with a lesion in the right frontal lobe produced heightened false memories. This individual was more likely than control individuals to false alarm to information that was thematically related to the earlier studied information. Interestingly, there is a more recent study of amnesia individuals with medial temporal lobe damage which suggests that individuals with damage in this area actually are *less* likely to produce false recognition of critical words, compared to healthy controls (see Schacter, Verfaellie, & Pradere, 1996c). In this light, the false memory paradigm is ideally suited to investigate the contributions of frontal areas and medial temporal areas to the episodic memory loss in healthy ageing and in early stages of DAT. If the episodic memory loss in early stage DAT is primarily a reflection of medial temporal structures, based on the Schacter et al. arguments, one might expect relatively fewer false memories compared to healthy control individuals, whereas if there is a high degree of frontal impairment, one might actually predict relatively more false memories in these individuals, compared to healthy control individuals. Of course, the more likely case is that both structures are intimately involved in the episodic memory loss in DAT individuals.

In the present study, we will attempt to provide some leverage on the underlying neural systems of false memories by taking advantage of existing psychometric data on the healthy older adults and early stage DAT individuals. Specifically, we have recently demonstrated (Kanne, Balota, Storandt, McKeel, & Morris, 1998) that psychometric tests that (a priori) should reflect functioning of different neural systems (frontal vs. medial temporal vs. parietal) loaded on the predicted factors on an analysis of data from 407 individuals with DAT. In addition, we found that the relative factor scores were correlated with the underlying neuropathology that was observed in 41 DAT individuals who were autopsied. In the present study, we will take advantage of the heterogeneity of this subject population in the integrity of these systems to provide insights, via correlational analyses, into the different neurological systems (frontal vs. medial temporal) that might underlie veridical and false memories.

In addition to providing correlational data regarding psychometric performance, the present study may provide information regarding the underlying mechanism(s) producing false memories. Specifically, several distinct loci may play a role. For example, it is possible that during initial encoding, the nonpresented critical word may actually come to mind as the participant is studying the to-be-remembered list of words, as a result of automatic associative responses (see McDermott, 1997). If this is the case, then this may serve as a learning episode. If the underlying mechanism for this encoding episode is due to an automatic spread of activation, based on the evidence reviewed earlier, one might expect comparable effects of this mechanism in healthy aged individuals and in early stage DAT. Interestingly, there recently has been evidence that one can produce false memories under conditions in which list items are presented at a near-threshold presentation rate (i.e. 50 msec per item), thereby implicating an automatic activation process (see Robinson, 1998; Seamon, Luo, & Gallo, 1998).

It is also possible that at least some of the false recall in the DRM paradigm is occurring during retrieval. Specifically, it is possible that during the overt act of recall, participants generate associatively related words, and hence are likely to generate the critical word. For example, it is possible that participants might have difficulty discriminating between items that were initially presented during study, and items that were generated, but rejected, earlier during the recall period (i.e. a type of source error, see Johnson et al., 1993). One might a priori expect that as more items are generated during recall one might actually increase the likelihood of generating the nonpresented item. Because healthy older adults and early stage DAT individuals should recall fewer items compared to healthy younger adults, one might expect a relative decrease in false memories. However, as noted earlier, this pattern does not occur, at least not in the healthy older adults (e.g. Norman & Schacter, 1997).

In addition to the possible role of source memory during retrieval, it is also quite possible that there are gist-based representations that are driving false recall during retrieval (see, for example, Brainerd, Reyna, & Kneer, 1995). Although the notion of a gist-based representation is a bit vague, one might argue that such gist-based representations should be positively correlated with the number of related items that are stored in memory. Specifically, it would seem that the quality of the gist-based representation would at least in part depend on the quality of the list item representations that produce the gist being represented. In this way, one might expect more false memory as overt memory increases. If this were the case, then because of the expected overall decrease in memory performance in healthy older adults and in early stage DAT, one might expect fewer false memories in these individuals, which, at least with respect to healthy ageing, is inconsistent with some extant data (e.g. Norman & Schacter, 1997). We shall explore these and additional theoretical implications of the relation between veridical and false recall in the General Discussion.

The present experiment included five groups of participants. The first three groups were healthy young (mean age 20 years), healthy older (mean age 71 years), and healthy old-old adults (mean age 86 years). These three groups allowed the tracking of the influence of healthy ageing on veridical and false memories. In addition, we compared three groups of individuals that are relatively similar in age but have varying levels of dementia severity. Specifically, the healthy old (both young-old and old-old) individuals were compared to age-matched very mildly demented individuals, who in turn were compared to age-matched mildly demented individuals. Of course, it would be particularly important if the DRM paradigm was able to discriminate early stage DAT from healthy ageing.

During the experiment, each participant received 6 lists of 12 words for study. After each list was presented, participants were given a free recall task. After all lists were studied and recalled, a final overall yes/no recognition test was presented. The critical dependent measure in the present study is the probability of recalling and/or recognising the critical nonpresented item as a function of the participant's correct recall and/or recognition.

METHOD

Participants

A total of 159 participants were included in this study. They were divided into five groups. The first group was composed of 35 young adults with a mean age of 20.1 years, with a range from 17 to 33 years. The second and third groups were healthy older adults. The healthy older adults were divided into a younger and an older group. There were 37 participants in the younger old group with a mean age of 70.7 years, and a range from 60 to 79. The older old group included 24 participants with a mean age of 85.9, and a range from 80 to 96.

The remaining two groups of participants were composed of individuals with either very mild dementia or mild DAT. Thirty-seven older adults with very mild DAT and an average age of 77.9 participated, and their ages ranged from 66 to 91. Twenty-six participants with mild DAT and an average age of 78.5 participated, and their ages ranged from 56 to 90.

The young adults were students at Washington University who were either paid or received course credit in return for their participation. Older adults and DAT individuals were recruited from the participant pool at the Alzheimer's Disease Research Center (ADRC) at Washington University. These participants were screened for depression, severe hypertension, possible reversible dementias, and other disorders that could affect cognitive performance. DAT individuals were included or excluded based on the criteria set by the National Institute of Neurological and Communicative 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, is well documented (e.g. Berg et al., 1998). Only those DAT individuals who were classified as having either very mild dementia or mild dementia were included as participants in this study.

Psychometric Test Performance

All the healthy older adults and the DAT individuals also participated in a 2-hour battery of psychometric tests designed to assess psychological functions including language, memory, and intelligence. 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 also received the Word Fluency test, on which they were required to name as many words as possible beginning with a specified letter (e.g. *P* or *S*) in a 60sec time period (Thurstone & Thurstone, 1949). Measures of general intelligence were the Information, Block Design, and

Digit Symbol subtests of the WAIS-R (Wechsler, 1955). Visual perceptual-motor performance was assessed by the Benton Copy Test and Trail Making Form A. 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, 1946). Participants also received the WMS Mental Control test, which evaluates the ability to produce quickly a well-rehearsed letter or digit sequence, such as the alphabet, in a specified amount of time. Finally, participants also completed the Boston Naming Test (Goodglass & Kaplan, 1983), which reflects semantic/lexical retrieval processes in naming simple line drawings.

Materials

Twelve lists of 15 words each were chosen from those used by Stadler, Roediger, and McDermott (in press). Each of these lists was created by choosing the top associates of a critical word, which was itself not part of the list. Six of these lists were adapted for use as study lists, and the other six lists were used as the source for distractor words used in the recognition phase of the experiment (see Appendix for the lists used). To create the study lists used in this experiment, the three weakest associates to the critical word were dropped from each list, creating lists of 12 words. The reduction of list length was necessary to reduce the difficulty of the task, and is actually the same length that Roediger and McDermott (1995) used in their first experiment.

The recognition test was printed on a sheet of paper, and consisted of 48 words randomly ordered in 3 columns of 16. The same order was used for all participants. The words used were the critical word and 3 of the associate words for each of 12 lists. Six of these lists were the lists studied in the first phase of the experiment; the other six served only as a source of distractors.

Procedure

Participants were tested individually. For the study, and free recall phase of the experiment, participants were instructed that they were to listen as the experimenter read a list of words slowly (at approximately 2sec per word), and that they would be asked to recall the words out loud after the experimenter finished reading the list. In order to familiarise the participants with the procedure, they were first given a practice list of six related words, which were tested for immediate recall. The participants were told to recall as many items as possible from each list and they were told to take as much time as they wished to complete their recall. After each list was finished, the participant was given an immediate free recall test. This phase of the experiment was tape recorded.

For the recognition phase of the experiment, participants were given the test sheet and instructed to circle any words that they heard presented on one of the six lists. They were additionally instructed not to spend too much time on any individual word.

RESULTS

Recall Performance

Figure 1 displays the serial recall functions as a function of group. As shown, all groups produced the standard bow-shaped serial position function, and there are clear differences across groups in overall memory performance. Specifically, one finds a large decrease in performance both as a function of healthy ageing (young is greater than the young-old, which is greater than the old-old), and DAT (the healthy old is greater than the very mild DAT, which is greater than the mild DAT). The results of a 5 (Group) by 12 (Serial Position) ANOVA yielded main effects of Group [$F(4,154) = 62.15$, $MSe = 0.15$, $P < .0001$] and Serial Position [$F(11,1694) = 80.92$, $MSe = 3.03$, $P < .0001$]. In addition, there was a reliable Group by Serial Position interaction [$F(44,1694) = 2.01$, $MSe = 0.08$, $P < .01$], which indicated that there were larger Group differences at the middle serial positions, compared to either terminal or primacy positions. For example, the difference between healthy young and mild DAT groups is 60% at Serial Position 4, whereas at Serial Position 1 and

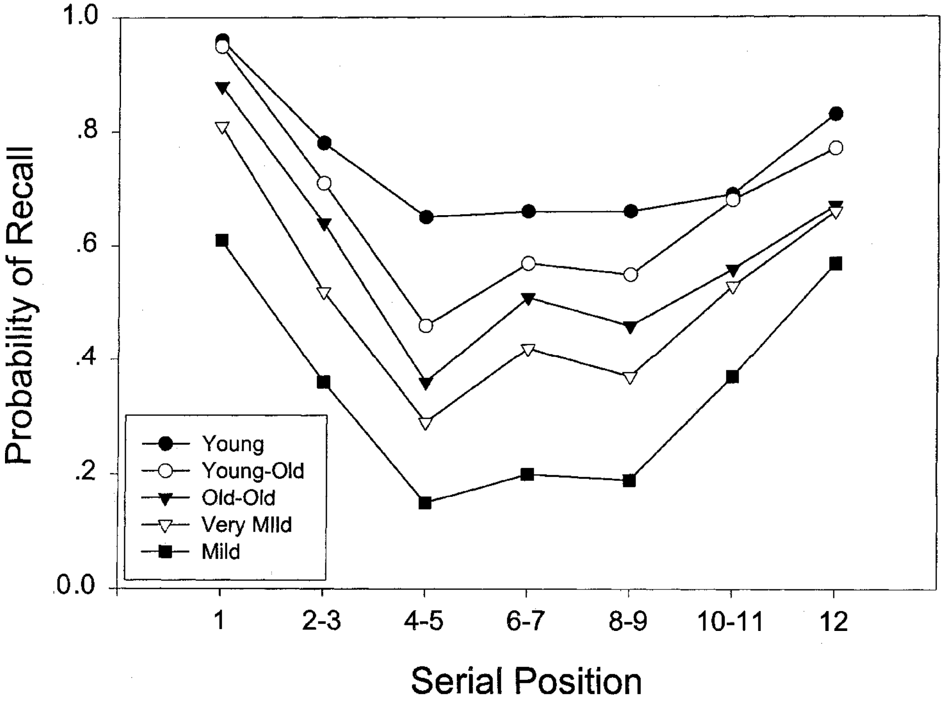


Fig. 1. *Mean proportion correct recall as a function of Serial Position and Group. The means of Serial Position 2 and 3, 4 and 5, 6 and 7, 8 and 9, 10 and 11 are presented.*

12 the differences are only 33% and 40% respectively. Of course, because of ceiling and floor effects at these serial positions, one must be cautious not to overinterpret this Group by Serial Position interaction. Overall, the results clearly indicate that the Serial Position function maintains its bowed-shape nature across groups, and there are the expected large Group differences in memory performance.

Figure 2 displays the mean proportion correct recall and the mean proportion critical intrusion rates. These results clearly indicate that there are large differences in the effects of Age and DAT on veridical recall and false recall. In order to address the influence of Age and DAT, we conducted separate ANOVAs as a function of Age and as a function of DAT. First, consider the effects of Age: There is a highly reliable drop of 26% correct in veridical recall from the healthy young adults to the old-old adults [$F(2,93) = 19.50$, $MSe = 0.011$, $P < .001$]. However, turning to the performance on the critical lure item, the results indicate that there is a small, and nonsignificant, 5% increase in recall across the groups [$F(2,93) = 1.27$, $MSe = 0.068$, $P = .29$]. There was a highly reliable interaction between Age Group and Item Type (veridical vs. critical false recall) [$F(2,93) = 6.18$, $P < .003$]. Turning to the DAT analyses, it should be noted that in this and subsequent analyses we collapsed across the young-old and the old-old groups of participants to create a single healthy older group to ensure that age was not confounded with DAT. As shown in Fig. 2, there is again a highly reliable 27% decrease in veridical recall across the healthy older group to the mild DAT group [$F(2,121) = 59.32$, $MSe = 0.014$, $P < .0001$]. However, there is only a nonsignificant 2% drop in the recall of the nonpresented critical word [$F(2,121) < 1.00$, $MSe = 0.072$]. Again, the interaction between Group and Item Type (veridical vs. critical false recall) was highly reliable [$F(2,121) = 9.71$, $MSe = 0.34$, $P < .001$]. In fact, if one considers the change across both age and DAT, one finds that

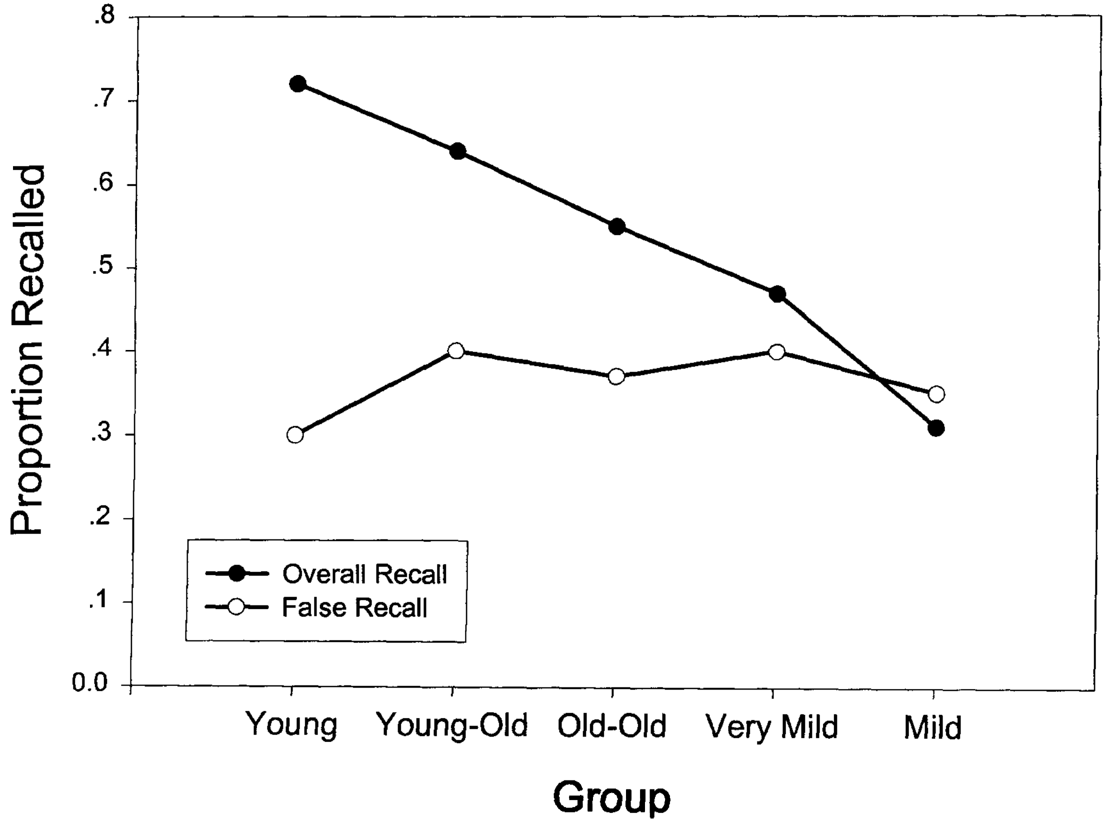


Fig. 2. *Mean proportion correct recall and mean proportion false critical recall as a function of Group.*

there is a 41% drop in veridical recall between the healthy young individuals and the mild DAT individuals [$F(4,154) = 62.15$, $MSe = 1.35$, $P < .0001$], and a nonsignificant 3% increase in recall of the nonpresented critical word [$F < 1.00$]. Of course, the interaction between Group and Item Type was highly reliable [$F(4,154) = 12.34$, $MSe = 0.48$, $P < .001$]. Clearly these results indicate that veridical recall changes dramatically across groups, whereas false recall is relatively stable.

In addition to examining critical word intrusions, we also considered intrusions of noncritical items. Because there are a number of distinct types of intrusions that may reflect different processes, we coded five different types of noncritical intrusions: (1) cross-list intrusions from previous lists; (2) cross-list critical lure intrusions (i.e. recalling a critical intrusion from a previous list); (3) semantically related, but noncritical intrusions; (4) phonologically related intrusions (most of which were rhymes); (5) "other," apparently unrelated, intrusions. The average number of intrusions per participant per list are displayed in Table 1. First, consider the cross-list intrusions. These reflect proactive interference from previous lists. The results indicate that both Age [$F(2,81) = 4.28$, $MSe = 0.047$, $P < .02$] and DAT [$F(2,116) = 11.04$, $MSe = 0.074$, $P < .0001$] affected susceptibility to proactive interference. Second, the cross-critical intrusions reflect the degree to which a highly activated critical item remained active across lists. Here there is little evidence of such intrusions in any group of participants. (Because of the low values, analyses were not conducted on this measure). The semantic intrusions reflected a semantically related word being produced on a given list. As shown in Table 1, there is an increase in semantically related intrusions as a function of Age [$F(2,81) = 5.24$,

Table 1. *Mean Number of Noncritical Intrusions per List as a Function of Group and Intrusion Type*

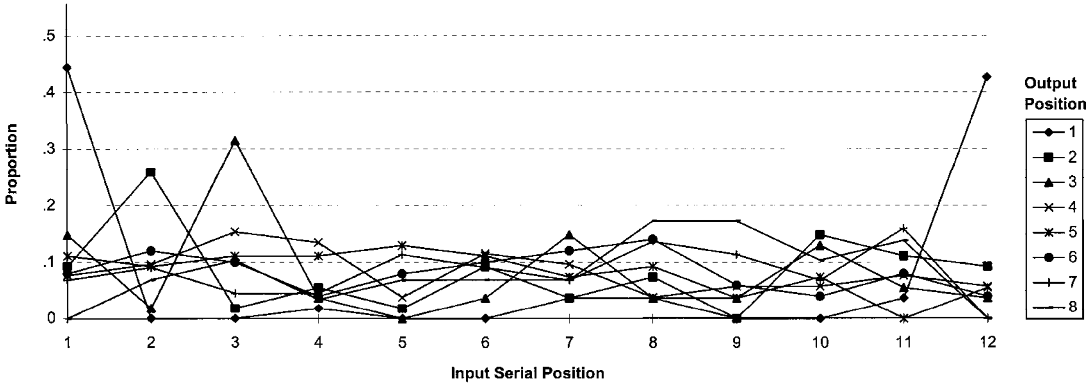
	Young	*Young Old*	*Old-Old*	*Very Mild*	*Mild*
Cross-list	.01	.04	.18	.21	.44
Cross-critical	.00	.00	.00	.00	.05
Semantic	.04	.15	.20	.18	.29
Phonological	.00	.00	.02	.03	.03
Other	.00	.01	.05	.08	.25
Total	.05	.20	.45	.50	1.06

$MSe = 0.029$, $P < .01$], but the effect of DAT did not reach significance [$F(2,116) = 1.46$, $MSe = 0.087$]. Also, as shown in Table 1, there were relatively few phonological intrusions in these data, which did not vary as a function of Age or DAT. (Again, because of the low values, analyses were not conducted.) Finally, the "other" intrusions reflected the output of relatively unrelated words. Although we could not conduct analyses as a function of Age because of the low levels in this category in the young adults, there was a reliable effect of DAT [$F(2,116) = 7.75$, $MSe = 0.057$, $P < .002$]. As shown here, the mild DAT individuals were likely to produce unrelated intrusions. However, further analyses of these items indicated that this was primarily produced by three participants, and when these three Mildly demented individuals were eliminated, the means were very similar across the two levels of DAT, i.e. 0.08 for the very mild DAT, and 0.11 for the mild DAT. In sum, there is clear evidence of a general tendency to produce noncritical intrusions as a function of Age and DAT status, but it is also clear that these errors are not random outputs of words, but reflected activated representations, either due to earlier list recall (proactive interference) or words that are semantically related to the current list.

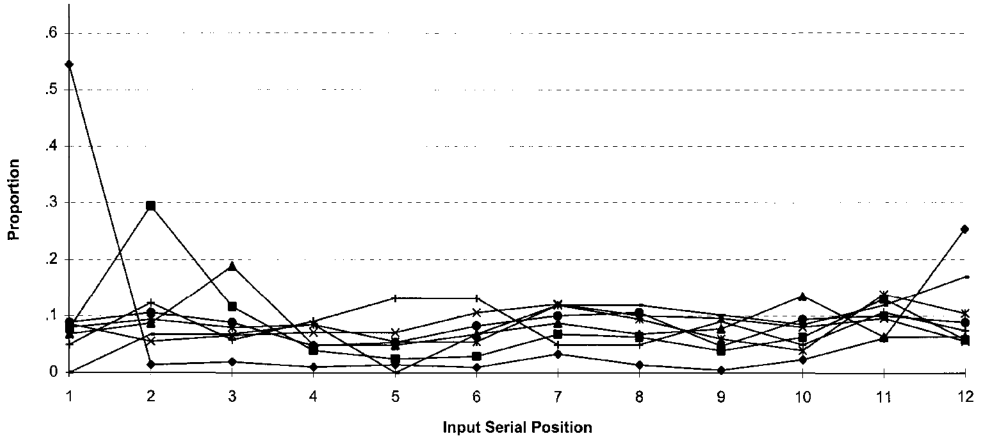
One intriguing aspect of the recall data is that across groups, it appears that as overt recall decreases, the critical intrusions remain relatively stable. This is an interesting pattern because one might expect that as participants recall more items, there should be a higher incidence of critical lure intrusions. In order to address this issue, we investigated the correlation between overall recall per participant and their false recall. The correlations indicated that there was some evidence of such a relation in the healthy old-old group ($r = .41$, $P < .05$) and the very mild DAT group ($r = .48$, $P < .01$), but none of the remaining groups produced reliable correlations, i.e. young ($r = -.20$), young-old ($-.06$), mild DAT ($+.04$). Most importantly, a correlation across all participants did not approach significance, $r = .06$. (It is worth noting here that Stadler, Roediger, & McDermott, in press, recently reported nonsignificant correlations between veridical recall and false recall across a large group of young adults and a larger set of lists.) Thus, above and beyond relatively limited correlations within specific groups, it does not appear that recalling items is a strong predictor of falsely recalling the critical nonpresented word.

We also tabulated the relative output position of both items correctly recalled and the critical intrusions. Because these analyses were based on tape recordings of the participants' verbal output, and some of the tapes were not audible (for 14 out of the 159 participants), these analyses were based on 26 young adults, 36 young-old adults, 22 old-old adults, 36 very mild DAT individuals, and 25 mild DAT individuals. There were two goals in this analysis: First, with respect to correct recall, it is possible that the present groups of participants may produce different retrieval strategies that might be reflected in the output positions. In order to address this possibility, we averaged the actual input position for each of the output positions. Of course, because there were large differences across groups, we could only do this analysis for the recall lengths of five, six, seven, and eight, in which there were sufficient number of observations per group. Moreover, to obtain stable estimates, we collapsed across the two healthy older groups and the two DAT groups. These data are displayed in Figs 3a (young), 3b (old), and 3c (DAT), as proportions of each input position that were produced in different output positions. There are three points to note from these data. First, the younger adults appear to be relatively more likely to recall the last item in the first position than the remaining two groups of participants. This pattern would be quite efficient if output of the last item would not disrupt the

(a) Younger Individuals



(b) Healthy Older Individuals



(c) DAT Individuals

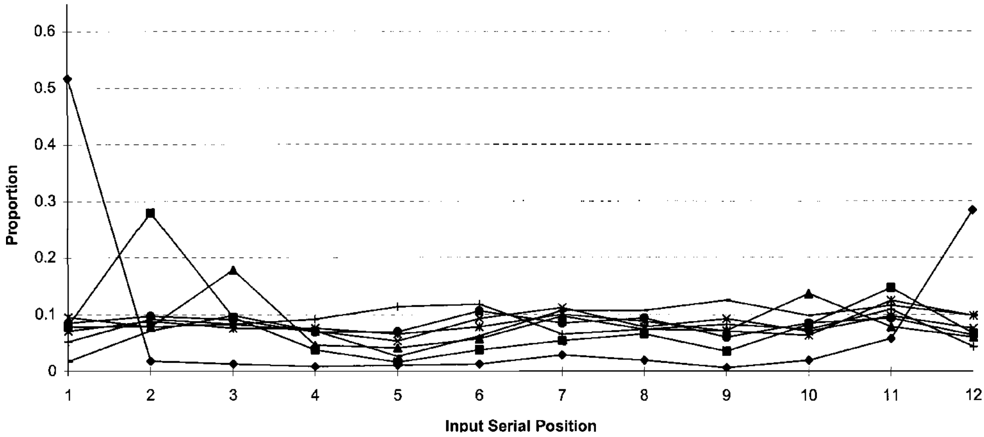


Fig. 3. *Mean proportion of each input position as a function of the output position during recall. These data are based on list recall lengths of 6, 7, and 8. The Healthy Old (Fig. 3b) is collapsed across Young-Old and Old-Old, and the DAT (Fig. 3c) is collapsed across Very Mild and Mild DAT.*

rehearsal pattern for the first few list positions, i.e. a type of memory triage. It appears that the healthy old and the DAT individuals are more likely to hold onto a rehearsal pattern that may not take advantage of the strong trace for the last presented item. In addition, as shown in Fig. 3a, one can see that for the young adults there is a consistent relation between the first three input serial positions producing high recall at the first three output positions. Turning to the healthy older adults and the DAT individuals, this pattern is only found for the first two serial positions, with a weaker effect for the third serial position being output at the third position. One possible account for this pattern is that the young adults have a three-item rehearsal set as opposed to a two-item rehearsal set in the healthy old and DAT individuals. Although the present analyses are only descriptive, it is intriguing that there are these different patterns of output when overall recall performance is equated across groups. We provide these data primarily because they provide a unique way to look at age-related and DAT-related changes in recall output and strategies.

In addition to analysing the output position of the list items as a function of input position, we also analysed the output position of the critical lure. 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 lure. These proportions indicated that there was remarkable consistency in the relative output position for the critical nonpresented items. Specifically, the relative position as a function of total number of items recalled were .66 (young), .70 (young-old), .66 (old-old), .60 (very mild DAT), and .64 (mild DAT). A one-way ANOVA indicated that this effect did not approach significance [*F*(4,116) = 1.36, *MS*E = 0.03]. Thus, the *relative* output position of the critical nonpresented word was about two thirds into the recalled items, independent of group and number of items recalled. This pattern is consistent with the analyses reported by Roediger and McDermott (1995) concerning output position.

Recognition Performance

The mean percentage of hits, false alarms to noncritical items, and false alarms to critical items on the recognition test are presented in Table 2[1]. There are two points to note in this table. First, consider the effects of Age. There is relatively little change in hit rate [*F*(2,93) = 0.73, *MS*e = 0.018] and a small, yet reliable, change in the noncritical (standard) false alarm rate (4%) across Age groups [*F*(2,93) = 5.33, *MS*e = 0.003, *P* < .01]. This pattern is consistent with the observation that older adults have relatively intact recognition performance, at least compared to recall performance (see Craik & Jennings, 1994). However, a different pattern emerges from the false alarms to the critical lures. Specifically, there is a large *increase* in false alarms to the critical lures (.23) as a function of Age [*F*(2,93) = 7.46, *MS*e = 0.076, *P* < .001]. Now, consider the DAT individuals. There is dramatic .33 drop-off in hit rate in the mild DAT [*F*(2,121) = 38.39, *MS*e = 0.026, *P* < .0001], whereas the increase in the noncritical false alarm

Table 2. *Mean Proportion Hit Rate, Noncritical False Alarm Rate, and Critical Lure False Alarm Rate as a Function of Group*

	Young	*Young-Old*	*Old-Old*	*Very Mild*	*Mild*
Hits	.80	.76	.77	.68	.44
Noncritical false alarms	.02	.04	.06	.09	.14
Critical lure false alarms	.58	.80	.81	.79	.51

---

[1] In the present experiment, the recognition test always followed the recall test, and so this measure is somewhat contaminated by earlier recall performance. However, it should also be noted that the effect of recall on recognition is relatively inconsistent in the false memory literature (see Roediger, McDermott, & Robinson, 1998). This potential contamination is a bit more serious when there are large differences across groups in recall performance, as in the present study. Thus, one must be cautious not to overinterpret the present recognition data.

rate is relatively small (.09), albeit significant [*F*(2,121) = 13.36, *MS*e = 0.006, *P* < .001]. The lack of a large increase in false alarm rates for these participants is important because it appears that even mildly demented individuals are relatively *cautious* in making false alarms. Turning to performance on the critical lure, although the very mildly demented individuals are producing a high false alarm rate to this word, the false alarm rate decreases in the mildly demented individuals [*F*(2,121) = 16.16, *MS*e = 0.055, *P* < .0001]. However, it should be pointed out that the false alarm rate to the critical lure is quite high (.51) compared to the noncritical lures (.14), even in the mildly demented individuals.

Figure 4 displays a measure of recognition performance (hits minus noncritical false alarms) and an adjusted measure of critical false alarm performance (critical false alarms minus noncritical false alarms) as a function of Group. As shown in Fig. 4, there is a cross-over interaction in the healthy older adult groups; specifically, although accurate recognition performance remains relatively stable, false recognition of the critical lure increases across groups. Moreover, although both correct and false recognition measures decrease as a function of DAT, the difference between correct recognition and false recognition remain relatively stable across the groups. An analysis of the correct recognition measure yielded a marginally reliable main effect of Age [*F*(2,93) = 2.71, *MS*e = 0.018, *P* < .07] and a main effect of DAT [*F*(2,121) = 62.56, *MS*e = 0.025, *P* < .0001]. Turning to the critical false alarm measure, there is a main effect of Age [*F*(2,93) = 5.84, *MS*e = 0.072, *P* < .01], which indicates that there is an increase in false recognition in the older adults, and a main effect of DAT [*F*(2,121) = 27.06, *MS*e = 0.052, *P* < .0001], which indicates that there is a decrease in false recognition in the DAT individuals compared to the healthy older adults. Finally, there is a highly reliable interaction between Age and Item Type (veridical vs. false recognition [*F*(2,93) = 9.71, *MS*e = 0.37, *P* < .0001]. However, because the size

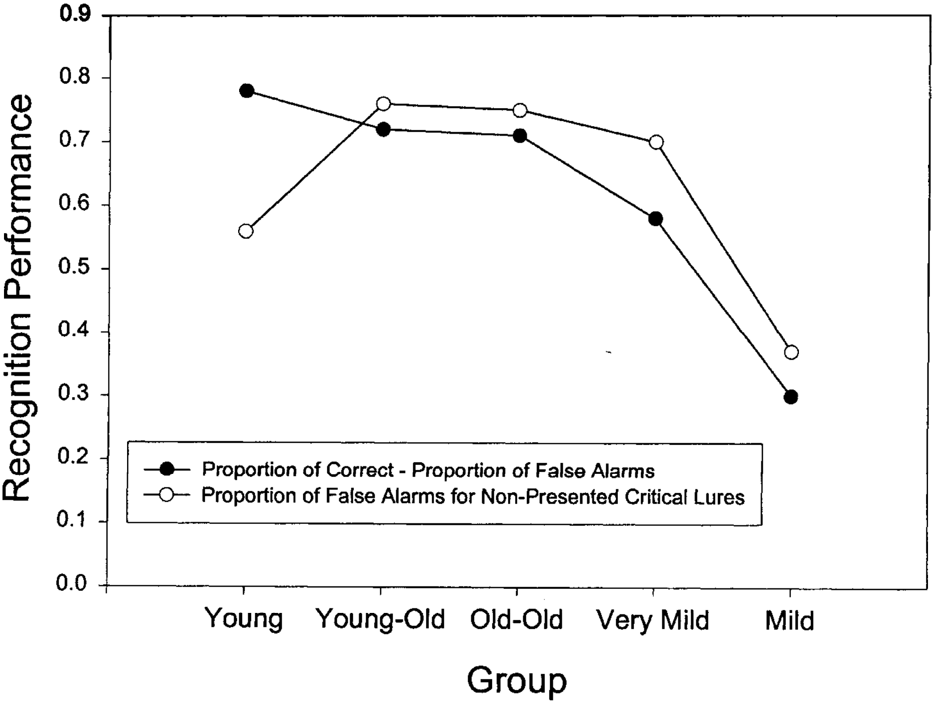


Fig. 4. *Mean recognition memory (hits minus noncritical false alarms) and false recognition (critical lures minus noncritical lures) as a function of Group.*

of the difference between veridical and false recognition did not change as a function of disease status, the interaction between DAT and Item Type did not reach significance [$F(2,121) = 1.65$, $MSe = 0.073$, $P = .20$].

Psychometric Performance

The means and standard deviations from the psychometric measures are presented in Table 3 as a function of Age and DAT. (The young adults did not receive psychometric testing.) A series of one-way ANOVAs with Group as a between-subjects factor indicated that performance on all measures decreased with increasing age and dementia severity, all Ps < .003. Thus, as expected, there are clear changes across groups across a wide range of cognitive measures.

The correlations between percentage correct and percentage false recall and the psychometric measures are presented in Table 4, as a function of Group. Also, as shown in this table, we have organised the psychometric tests as a function of differences in possible underlying neural systems (i.e. medial temporal, parietal, and frontal). As noted earlier, the selection of this particular grouping is based on recent work by Kanne et al. (1998). Because of the evidence suggesting a role of frontal structures in false recall, we expected that decreased frontal functioning would be associated with increased susceptibility to false recall, i.e. negative correlations. We also expected that medial temporal measures should be relatively more positively correlated with veridical recall, although there should be at least some correlations with the frontal measures. As shown in Table 4, these correlations did not support these predictions. Specifically, most psychometric measures were positively correlated with veridical recall, independent of the nature of the psychometric task. On the other hand, the pattern of correlations between the psychometric tests and false recall was much less consistent. The only groups that produced the predicted negative correlations were the healthy young-old participants and the mildly demented individuals. Unfortunately, both of these groups also produced some evidence of negative correlations with the medial temporal measures and false recall. Thus, there was relatively little support for the predicted correlations with the present psychometric tests[2].

Table 3. *Psychometric Means (SD) as a Function of Group*

	CDR 0 (Young-Old)	*CDR 0 (Old-Old)*	*CDR 0.5 (Very Mild DAT)*	*CDR 1 (Mild DAT)*
WAIS Information	23.8 (2.9)	20.6 (3.8)	17.6 (5.6)	12.0 (4.2)
Boston Naming	57.6 (3.0)	51.8 (8.7)	49.1 (11.8)	37.7 (12.6)
Logical Memory	10.3 (2.6)	8.2 (3.2)	6.2 (3.8)	2.6 (1.5)
Associate Memory	15.1 (3.5)	12.3 (3.7)	11.4 (3.8)	7.5 (1.8)
Benton Copy	9.8 (0.7)	9.4 (0.8)	9.1 (1.3)	8.2 (2.1)
Trailmaking A[a]	38.0 (11.6)	45.2 (11.8)	45.3 (23.4)	78.3 (32.2)
Block Design	35.3 (7.4)	29.0 (6.7)	24.1 (7.5)	16.2 (8.2)
Digit Symbol	50.6 (9.8)	38.0 (9.6)	39.7 (12.6)	23.3 (9.1)
Digit Span	12.1 (2.2)	11.1 (2.6)	10.5 (2.2)	9.4 (1.9)
Word Fluency	31.2 (11.2)	26.9 (10.3)	25.6 (12.2)	19.8 (9.0)
Mental Control	8.0 (1.3)	7.5 (1.8)	7.1 (1.9)	6.2 (2.1)

[a] Represents time in seconds.

---

[2] In addition to the correlations within groups displayed in Table 4, we also conducted overall correlations independent of group. Moreover, we also conducted correlations in which we converted each psychometric score to a z-score and then obtained a relative measure of each set of scores (i.e. mean of frontal, mean of medial temporal, and mean of parietal scores) divided by overall z-scores across all tasks. None of these individual correlations produced the predicted pattern of correlations between frontal measures and false recall.

Table 4. *Correlations between Percentage Correct and Percentage False Recall and Psychometric Measures by Group*

	CDR 0 (Young-Old)	*CR 0 (Old-Old)*	*CDR 0.5 (Very Mild DAT)*	*CDR 1 (Mild DAT)*
Percentage Correct Recall				
Medial Temporal				
WAIS Information	.12	.55*	.71	.31
Boston Naming	.14	.47*	.69*	.35
Logical Memory	.13	.60*	.68*	.69*
Associate Memory	.24	.75*	.50*	.51*
Parietal				
Benton Copy	.39*	.29	.23	−.02
Trailmaking A	−.15	−.50*	−.48*	.16
Block Design	.23	.62*	.44*	.18
Digit Symbol	.37*	.62*	.73*	.02
Frontal				
Digit Span	.27	.44*	.40*	.38
Word Fluency	.46*	.72*	.57*	.27
Mental Control	.11	.21	.50*	−.01
Percentage False Recall				
Medial Temporal				
WAIS Information	−.65	.32	.26	−.24
Boston Naming	−.42*	.25	.28	−.47*
Logical Memory	−.20	.08	.23	−.10
Associate Memory	−.15	.28	.15	.20
Parietal				
Benton Copy	.15	.49*	.17	−.23
Trailmaking A	.00	−.14	−.40*	.24
Block Design	.16	.48*	.28	−.36
Digit Symbol	.14	.53*	.38*	−.25
Frontal				
Digit Span	−.07	.08	.11	−.46*
Word Fluency	.04	.17	.60*	−.05
Mental Control	−.19	.21	.38*	−.20

*$P < .05$

GENERAL DISCUSSION

The results of the present experiment are clear. Both healthy ageing and DAT produced large changes in veridical recall performance. However, there was relatively little influence of either age or DAT on false recall performance. A similar pattern was found in recognition performance (although see Footnote 1). We shall now turn to a discussion of the implications of these results for the understanding of the memory changes that occur in both healthy older adults and in DAT individuals, and the mechanisms that may underlie false memories in the DRM paradigm.

Age Differences in False Memory Performance

The results from the young adults and the young-old adults nicely replicate the findings of Norman and Schacter (1997) and Tun et al. (1998) with the DRM materials. Moreover, the addition of the old-old adults extends this pattern to a group of individuals in their mid 80s. Specifically, in the present study, veridical recall decreased as a function of age, whereas false recall and recognition increased as a function of age (although the effect in recall was not reliable). The fact that this pattern has been replicated in three different laboratories clearly solidifies the robustness of this finding. Interestingly, even under conditions in which there is not a reliable increase in the false recall or recognition performance, if one considers false memory relative to the overall veridical memory, there is evidence of an age-related change in the *relative* presence of false memory performance in healthy older adults, compared to healthy young adults (see McEvoy & Nelson, 1998). It is important to note here that this pattern does not appear to be limited to the DRM materials, because older adults have shown increased susceptibility to memory distortions and illusions across a wide set of experimental paradigms (e.g. Bartlett, Strater, & Fulton, 1991; Cohen & Falkner, 1989; Dywan & Jacoby, 1990; Jacoby, this issue, in press; Kensinger & Schacter, this issue; Koutstaal & Schacter, 1997; Rankin & Kausler, 1979; Schacter, Israel, & Racine, 1999).

The more intriguing question regarding these data concerns the underlying mechanisms. There have been a number of such mechanisms recently proposed in the literature. For example, Norman and Schacter (1997) have suggested that, compared to younger adults, older adults might have less specific detail regarding perceptual characteristics of the list items. Hence, older adults may be more susceptible to confusing correctly recalled items with highly activated critical nonpresented items, thereby producing a type of source misattribution error (Johnson et al., 1993). Younger adults presumably have more perceptual information available to make the appropriate source attribution. This possibility is consistent with the work of Hashtroudi, Johnson, and Chrosniak (1990), who found that confusions in perceptual details appear to increase the likelihood of source errors in older adults, compared to younger adults. Tun et al. (1998) have suggested that there is an age-related increase in gist-based processing. Specifically, as individuals age, and the item-specific information decreases, there is an increased reliance on other sources, such as overall gist. Because the critical items in the DRM materials are highly activated by gist-based information, this would clearly produce the observed pattern. This view is also consistent with the work by Reder, Wible, and Martin (1986), who argued that older adults are more likely to rely on "plausibility" in making their judgements regarding previously presented sentences, whereas younger adults were more likely to recall item-specific information. Reder et al. suggested that this may be due to differences in attentional resources across groups, i.e. older adults have fewer resources, and it is less demanding to retrieve plausibility information than item-specific-information. Similarly, Rabinowitz, Craik, and Ackerman (1982) have noted that older adults are more likely to encode generic information than younger adults. All of these theoretical perspectives suggest that older adults are more likely to rely on overall list relations than on item-specific information.

Why might older adults rely more heavily on relations across items than on deindividuating information for specifically encoded items? We have suggested that this pattern may not simply be a reflection of memory specific changes in older adults, but rather may be reflective of attentional control changes. Specifically, with the DRM materials, the participant needs to discriminate between two types of information: (a) highly activated semantic convergence on a given nonpresented representation and (b) activated item-specific information due to initial encoding. As shown in Fig. 5, one could envisage these two sources competing for correct output during an episodic memory task. An efficient attentional control system needs to be able to discriminate between these two sources of information, i.e. accentuating (as indicated by the plus sign) the item-specific recollection, and inhibiting (as indicated via the minus sign) the pathway that produces high availability due to the spreading activation across items that converge on critical nonpresented words. In this light, there appears to be an analogy to the Stroop task, wherein the participant must select between a strong word processing pathway and a weaker, but task-relevant, colour naming pathway. One might argue that the critical nonpresented item is like the word pathway in Stroop, wherein the participant must control the tendency to respond with the highly activated representation. As noted earlier, there is clear evidence that there are age-related effects in both the Stroop task, along with other non-memory tasks such as in

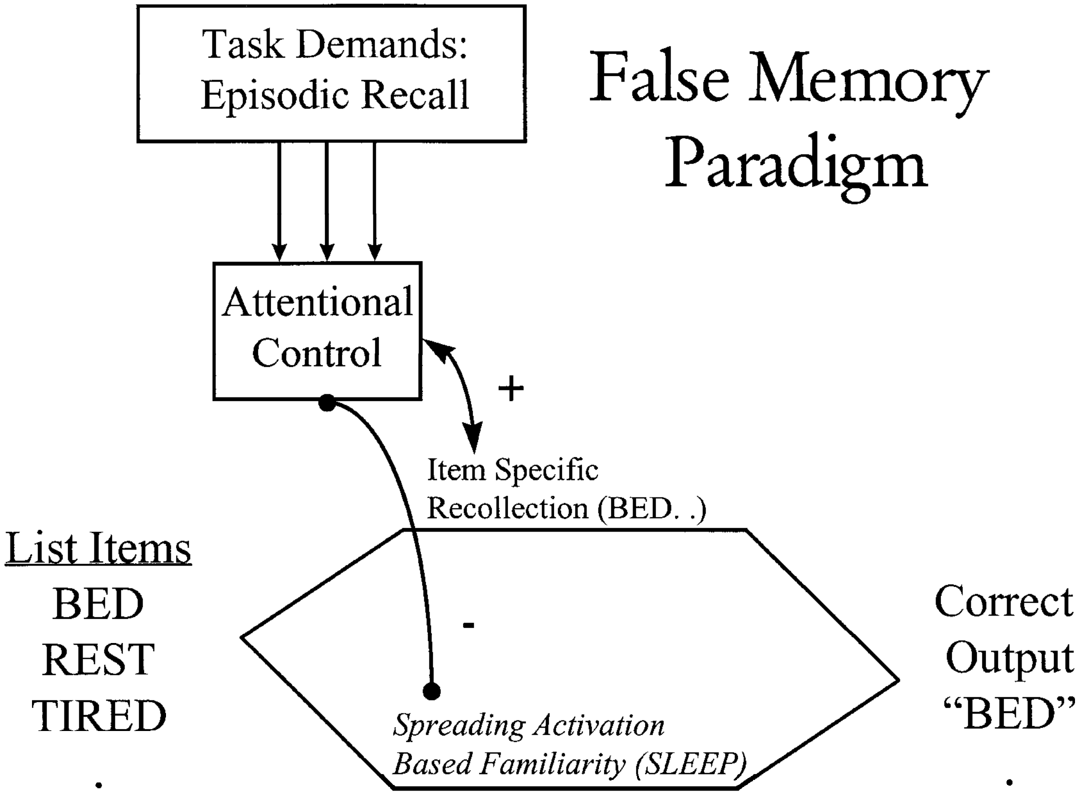


Fig. 5. *An attentional control framework for interpreting performance in the DRM recall paradigm. Deficits as a function of Age and DAT might be expected in the attentional control module.*

picture word interference paradigms (e.g. Duchek, Balota, Faust, & Ferraro, 1995), and the speeded naming of low-frequency exception words (e.g. Balota & Ferraro, 1993). Before further discussion of this framework, we shall turn to a discussion of the results from Alzheimer’s Disease individuals.

Veridical and False Recall in DAT

The results from the DAT individuals appear to be in large part an extension of the changes we find in healthy older adults. Specifically, veridical recall continued to decrease across levels of dementia; however, the false recall remained relatively stable (see Fig. 2). Turning to the recognition data (see Fig. 4) there is a similar pattern, with the exception of the mildly demented individuals, in whom there is a decrease in both veridical recognition and false recognition. However, it is important to note that both the very mild and the mildly demented individuals produced a higher probability of false recognition than of correct recognition. Overall, these results indicate that relative to veridical memory, the DAT individuals are even more susceptible to false memories than are healthy older adults.

The DAT individuals not only falsely recalled the critical nonpresented items at disproportionately high levels relative to their overall recall performance but, as shown in Table 1, they were also more likely to produce cross-list intrusions (reflecting an increased susceptibility to proactive interference) and semantically related intrusions. Hence, one must entertain the possibility that the increase in false memories in the DAT individuals simply reflects a criterion shift in these participants. However, if this were the case, then one might expect a relatively high false alarm rate on the recognition test. Interestingly, however, both the very mild and mild DAT individuals produced relatively low false alarm rates to noncritical lures, .09 and .14 respectively. Thus, at least in the recognition test, the DAT individuals were resistant to respond "yes" to a lure stimulus. Of course, this cautiousness in responding primarily applies to the noncritical lures, because the critical lures produced a false alarm rate of .51[3]. Thus, although there is an overall increase in cross-list intrusions and semantically related intrusions, we believe these errors reflect relatively high activation for competing sources that need to be controlled via attentional systems. In this light, we do not believe that the present results are simply a reflection of a criterion shift in the DAT individuals.

As in the healthy older adults, there is also evidence that DAT individuals produce breakdowns in attentional control processes that, as suggested earlier (and in Fig. 5), may lead to a relative increase in false recall. For example, DAT individuals are more likely to produce failures to resist highly activated, but competing, representations. The Stroop task again nicely documents this problem. Specifically, DAT individuals are more likely to produce the incongruent word responses under colour naming instructions, compared to healthy age-matched controls (Spieler et al., 1996). Thus, if there is a further breakdown in the attentional control system that controls highly activated pathways in DAT individuals, then (based on the framework displayed in Fig. 5), one might expect a further increase in false memories in these participants. Of course, this assumes that there is intact spreading activation for the critical nonpresented item. As noted earlier, the evidence from the semantic priming literature indicates that the automatic activation process is relatively intact in both healthy ageing and in early stage DAT individuals (Balota & Duchek, 1991; Duchek & Balota, 1993; Nebes, 1989; Shenaut & Ober, 1996). Thus, there should be activation for the critical nonpresented item that competes for output against contextualised item-specific information[4].

Although we have been discussing the present healthy ageing and DAT results within the attentional control framework displayed in Fig. 5, there is an important alternative account that must be discussed. Specifically, it is possible that one does not need to invoke changes in attentional control systems, but rather that these results simply reflect an imbalance in the strength of the two sources of information. There is already considerable extant literature that both healthy older adults and to a greater extent DAT individuals produce large decrements in episodic memory performance. This, of course, was also reflected in the present recall performance. Within the framework displayed in Fig. 5, this would be viewed as suggesting that the item-specific recollection pathway is impaired in these groups of participants. Also, as noted, there is considerable evidence that the automatic spreading activation is relatively intact in both healthy older adults and in early stage DAT. Thus, the spreading activation-based pathway in Fig. 5 should be relatively intact. The present results do not demand a change in attentional control systems, but rather may simply reflect changes in the balance of the strength of the two underlying

---

[3] It is also worth noting here that if one adjusts the false recall data by the overall intrusion rate, then one would also need to adjust the veridical recall for the intrusion rate, because a simple shift in output criterion is likely to reflect both correct recall and false recall. Of course, if one makes the same adjustment for both false recall and veridical recall, then this will not have any effect on the important observation in Fig. 2 that veridical and false recall change in a dramatically different manner across groups.

[4] It is worth noting that the present discussion focuses on group-related changes in attentional control processes at retrieval in producing the relative increase in false memories across groups. However, it is important to note that these group-related differences in attention would also change the nature of the encoded traces across groups. In fact, one might expect decreased semantic elaborative encoding across the participant groups due to changes in attentional control systems at encoding across the participant groups. Interestingly, there is evidence that deeper, more elaborative encoding processes actually increase rather than decrease false recall (Thaper, McDermott, & Fong, 1998; Toglia, Neuschatz, & Goodwin, 1999). Because the older adults and DAT individuals may be expected to engage in less elaborative processing, the present results might actually *underestimate* the relative increase in false memories to veridical memories in these groups.

processing pathways, i.e. the item-specific recollection pathway decreases in strength across groups, whereas, the spreading activation-based pathway remains relatively stable across groups. Therefore, because of the potential competition between the two pathways, memory performance is more driven by the spreading activation-based pathway, as the recollection-based pathway decreases as a function of both healthy ageing and DAT.

In order to address this alternative account of the present results, and also to shed some light on past false memory studies, we attempted to equate the strength of the recollection-based pathway across our participants. Thus, we attempted to match participants on both mean performance (and range) across groups on their veridical recall. If indeed there is an attentional control problem in controlling the spreading activation-based pathway as a function of age and DAT, then one should find a relative increase in false recall performance across groups of participants who are matched on veridical recall. These data are shown in Fig. 6, and clearly support this prediction. Specifically, if one equates veridical recall across groups by matching participants, the false recall performance actually increases for these same participants. This increase in false recall was reliable in the comparison of the healthy young vs. old [$t(59) = 2.27$, $P < .05$]; however, because of the decreased power (due to selective matching) this comparison did not reach significance in the comparison of the healthy old vs. very mild DAT or in the comparison of the very mild vs. mild DAT individuals. However, when one includes these later two comparisons in an overall ANOVA (with no participant represented twice in the analyses), there is a marginally reliable increase [$F(1,78) = 3.61$, $P < .06$]. In addition to these analyses, and in order to avoid the problem with differential subject selection, we also conducted a regression analysis in which we predicted false recall after veridical recall was partialled out. This analysis indicated that participant group reliably predicted false recall after we had partialled out the variance due to veridical recall [$t(156) = 2.23$, $P < .05$]. Thus, these results indicate that the increased susceptibility to false recall is not simply a reflection of a decrease in the item-specific recollection-based pathway, because there is an overall increase in false recall even when participants were equated on veridical recall. Rather, these results are more consistent with the notion that there is a change across groups in attentional control of relevant processing pathways.

Potential Neurological Underpinnings for False Memories

What might be the neural underpinnings for the breakdowns in attentional control systems in healthy older adults and in early stage DAT individuals? There has been some suggestion in the literature that older adults have more difficulty in tasks that reflect frontal functioning rather than medial temporal or parietal functioning (e.g. Craik, Morris, Morris, & Loewen, 1990; Moscovitch & Winocur, 1995; Shimamura & Jurica, 1994). There is also evidence that there is an increased influence of age on volumetric measures of frontal areas of the brain (e.g. West, 1996, for a review), and recent neuropathological studies of DAT (Morris et al., 1996) indicate a relatively high presence of plaques in frontal areas. As reviewed in the Introduction, false recognition and confabulation effects have been shown to increase in individuals with frontal damage (e.g. Schacter et al., 1996a). In addition, tasks that involve selection from a number of activated pathways, such as the Stroop and Wisconsin Card Sorting tasks, appear to load more on frontal systems (e.g. Dempster, 1992). All of these observations converge on the notion that breakdowns in frontal systems may be related to the observed relative increase in false memories. Although this account is clearly tempting, an analysis of the selected psychometric tests did not provide much support for this prediction. Specifically, although medial temporal measures were nicely correlated with veridical recall, there was relatively little correlation between the frontal measures and false recall. Unfortunately, as discussed earlier, the false recall measures were relatively uncorrelated with most measures of either memory or attention. Of course, it is quite possible that the selected set of psychometric tasks to measure frontal functioning (which were limited in scope due to the longitudinal

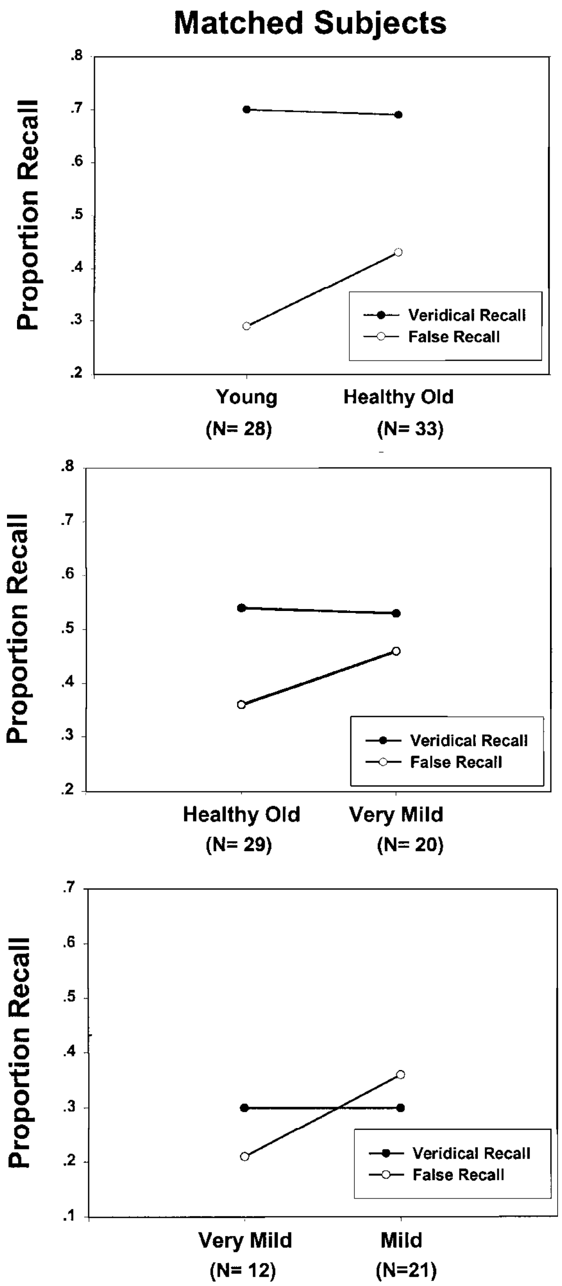


Fig. 6. *Matched subjects as a function of Age (top panel), Age and Very Mild DAT (middle panel), and Very Mild and Mild DAT (bottom panel). The number of participants in each group are in parentheses.*

studies in which these subjects are participating) did not capture the critical operations, which still could be mediated by frontal systems. Thus, it would be quite useful to use the psychometric tasks that Glisky, Polster, and Routhieaux (1995) have recently used to determine if one could predict false recall performance in the DRM paradigm via frontal functioning as evidenced from a more targeted neuropsychological battery. In this light, it should be noted that Henkel, Johnson, and De Leonardis (1998) recently provided evidence that false memories were predicted by both frontal and medial temporal psychometric measures. Clearly further work is needed to examine the contribution of frontal systems in the production of false memories in the DRM paradigm across healthy ageing and DAT.

Implications of the Present Results for Understanding the Mechanisms Underlying False Recall and Recognition

The present results not only have implications for understanding age-related and dementia-related changes in memory performance, but these results also have implications for understanding the mechanisms that might underlie false memories in the DRM paradigm. Following Deese (1959) and Roediger and McDermott (1995), we have suggested that false memories may reflect a type of automatic implicit association that occurs during list presentation. The underlying mechanism for this might be a spread of activation within a semantic network, which converges on a nonpresented critical representation. This activation is likely to summate across the related words until some critical threshold is reached. In fact, Balota and Paul (1996) have demonstrated that automatic semantic priming effects are additive across multiple related words in both lexical decision and naming tasks. Recent work by Robinson (1998) and Seamon et al. (1998) has indicated that conscious processing of the list words is not a necessary condition for producing false memories. For example, Robinson found evidence for robust false recall and recognition under presentation conditions that greatly minimised conscious processing (50msec per word with a backward masking stimulus). Under these conditions, participants were not consciously generating the nonpresented items at encoding and so this does not appear to be the primary locus of the false memories. Finally, it is noteworthy that Stadler, McDermott, Roediger, Miller, and Cowan (1999) have recently found that divided attention tasks disrupt veridical recall more than false recall. These studies all appear to converge on the notion that the conscious generation of the critical nonpresented item during list presentation is not a necessary condition to produce false recall.

What are the implications of the present results for the notion that retrieval is the locus of false recall? One might argue that in the process of generating words at retrieval, the participant actually generates the critical item because of the previously generated related words. Once the critical item is generated, it is likely to be output because of the high activation due to the convergence of the semantic associates. However, the present results also produce some constraints on this interpretation. That is, one of the most striking aspects of the present data is that although veridical recall decreases by nearly 50% across the present groups of participants, false recall is invariant across these same groups. This is important because it suggests that the overt act of recalling associates does not produce the false recalls.

CONCLUSIONS

The present results clearly indicate that the relative incidence of false memories to veridical memories increases as a function of both age and DAT. We have argued that part of this effect may be due to breakdowns in attentional control systems that select among activated pathways due to spreading activation and item-specific information due to memory encoding. Clearly, this pattern of results can be informative regarding age-related and DAT-related changes in memory performance and also can provide insight to basic memory mechanisms that produce false memories in healthy young adults. Of course, one might also ask whether there may be ways of increasing the functioning of the attentional control system in both healthy older adults and in DAT individuals. For example, Multhaup (1995) has shown that older adults are less likely to fall prey to illusions of fame when provided with a more exhaustive set of response alternatives at retrieval. It is also possible that older adults may benefit more than younger adults if there is a warning signal that indicates the presence of such highly related distractors, or possibly if they are given more time during the retrieval task. Given the potential negative consequences of false memories, it is important not only to understand this phenomenon but also consider how to limit it in populations that appear to be particularly susceptible (see Schacter et al., 1998, for further discussion of these issues).

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. (1946). An analysis of certain psychological tests used for the evaluation of the brain injury. *Psychology Monographs, 60,* (Whole No. 277).

Balota, D.A., Black, S., & Cheney, M. (1992). Automatic and attentional processes in young and old adults: A re-evaluation of the two-process model of semantic priming. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 485–502.

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

Balota, D.A., & Ferraro, F.R. (1993). A dissociation of frequency and regularity effects in pronunciation performance across young adults, older adults, and individuals with senile dementia of the Alzheimer type. *Journal of Memory and Language, 32,* 573–592.

Balota, D.A., & Paul, S.T. (1996). Summation of activation: Evidence from multiple primes that converge and diverge within semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 336–345.

Balota, D.A., Watson, J.M., Duchek, J.M., & Ferraro, R.F. (in press). 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 International Neuropsychological Society*.

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

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.

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

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.), *Handbook of aging and cognition* (pp. 451–110). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Craik, F.I.M., Morris, L.W., Morris, R.G., & Loewen, E.R. (1990). Relations between source amnesia and frontal lobe functioning in older adults. *Psychology and Aging, 5,* 148–151.

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

Dempster, F.N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review, 12,* 45–75.

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).

Duchek, J.M., Balota, D.A., Faust, M.E., & Ferraro, F.R. (1995). Inhibitory processes in young and older adults in a picture-word task. *Aging and Concern, 2,* 156–167.

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

Dywan, J., & Jacoby, L.L. (1990). Effects of aging on source monitoring: Differences in susceptibility to false fame. *Psychology and Aging, 5, 379*–387.

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.

Glisky, E.L., Polster, M.R., & Routhieaux, B.C. (1995). Double dissociation between item and source memory. *Psychology and Aging, 9,* 229–235.

Goodglass, H., & Kaplan, E. (1983). *Boston Naming Test.* Philadelphia, PA: 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.

Hashtroudi, S., Johnson, M.K., & Chrosniak, L.D. (1990). Aging and qualitative characteristics of memories for perceived and imagined complex events. *Psychology and Aging, 5,* 119–126.

Henkel, L.A., Johnson, M.K., & De Leonardis, D.M. (1998). Aging and source monitoring: Cognitive processes and neuropsychological correlates. *Journal of Experimental Psychology: General, 127,* 251–269.

Jacoby, L.L. (this issue). Deceiving the elderly: Effects of accessibility bias in cued recall performance. *Cognitive Neuropsychology, 16,* 417–436.

Jacoby, L.L. (in press). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology, Learning, Memory and Cognition.*

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

Kanne, S.M., Balota, D.A., Storandt, M., McKeel, D.W., & Morris, J.C. (1998). Relating anatomy to function in Alzheimer's disease: Neuropsychological profiles predict regional neuropathology five years later. *Neurology, 5,* 979–985.

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

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. (1997). Priming on perceptual implicit memory tests can be achieved through presentation of associates. *Psychonomic Bulletin and Review, 4,* 582–586.

McEvoy, C.L., & Nelson, D.L. (1998). *False memories in young and older adults: The view from an associative model of memory.* Paper presented at the 1998 Cognitive Aging Conference, Atlanta, Georgia.

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.

Morris, J.C., Storandt, M., McKeel, D.W. Jr, Rubin, E.H., Price, J.L., Grant, E.A., & Berg, L. (1996). Cerebral amyloid deposition and diffuse plaques in "normal" aging: Evidence for presymptomatic and very mild Alzheimer's Disease. *Neurology, 46,* 707–719.

Moscovitch, M., & Winocur, G. (1995). Frontal lobes, memory, and aging. In J. Grafman, K.J. Holyoak, & F. Boller (Eds.), *Structure and functions of the human prefrontal cortex* (Annals of the New York Academy of Sciences, Vol. 769, pp. 119–150). New York: New York Academy of Sciences.

Multhaup, K.S. (1995). Aging, source, and decision criteria: When false fame errors do and do not occur. *Psychology and Aging, 10,* 492–497.

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. (1994). 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: Lawrence Erlbaum Associates Inc.

Norman, K.A., & Schacter, D.L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory and 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 difference in word and language processing* (pp. 247–271). Amsterdam: Elsevier.

Rabinowitz, J.C., Craik, F.I.M., & Ackerman, B.P. (1982). A processing resource account of age differences in recall. *Canadian Journal of Psychology, 36,* 325–344.

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

Reder, L.M., Wible, C., & Martin, J. (1986). Different memory changes with age: Exact retrieval versus plausible inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 72–81.

Robinson, K.J. (1998). *The automatic mechanisms of false memory.* Unpublished Masters thesis.

Roediger, H.L., & McDermott, K.B. (1995). Creating false memories: Remembering words not presented in list. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803–814.

Schacter, D.L., Curran, T., Galliccio, L., Milberg, W.P., & Bates, J.F. (1996a). False recognition and the right frontal lobe. *Neuropsychologia, 34,* 793–808.

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

Schacter, D.L., Reiman, E., Curran, T., Yun, L.S., Bandy, D., McDermott, K.B., & Roediger, H.L. III (1996b). Neuroanatomical correlates of veridical and illusory recognition memory revealed by PET. *Neuron, 17,* 267–274.

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

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.

Shimamura, A.P., & Jurica, P.J. (1994). Memory interference effects and aging: Findings from a test of frontal lobe function. *Neuropsychology, 8,* 408–412.

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.

Stadler, M.A., McDermott, K.B., Roediger, H.L., Miller, E.C., & Cowan, N. (1999). *Effects of attention on false recognition.* Unpublished data.

Stadler, M.A., Roediger, H.L., & McDermott, K.B. (in press). Normed materials for creating false memories: Remembering words not presented in lists. *Memory and Cognition.*

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.

Thapar, A., McDermott, K.B., & Fong, C.T. (1997). *Effects of level of processing and retention interval on false recall.* Poster presented at the 38th Psychonomics Society Conference in Philadelphia, PA.

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

Toglia, M.P., Neuschatz, J.S., & Goodwin, K.A. (1999). Recall accuracy and illusory memories: When more is less. *Memory, 7,* 233–256.

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.

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

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

West, R.L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin, 120,* 272–292.

APPENDIX A

Lists Used During Study and Free Recall

List 1: door, glass, pane, shade, ledge, sill, house, open, curtain, frame, view, breeze (critical word WINDOW)
List 2: bed, rest, awake, tired, dream, wake, snooze, blanket, doze, slumber, snore, nap (critical word SLEEP)
List 3: nose, breath, sniff, aroma, hear, see, nostril, whiff, scent, reek, stench, fragrance (critical word SMELL)
List 4: nurse, sick, lawyer, medicine, health, hospital, dentist, physician, ill, patient, office, stethoscope (critical word DOCTOR)
List 5: sour, candy, sugar, bitter, good, taste, tooth, nice, honey, soda, chocolate, heart (critical word SWEET)
List 6: table, sit, legs, seat, couch, desk, recliner, sofa, wood, cushion, swivel, stool (critical word CHAIR)

Words Used for Recognition

house, smoke, lawyer, window, trash, waste, rough, curtain, taste, refuse, wood, smell, soft, aroma, gravel, needle, soda, chocolate, pillow, snooze, billow, eye, cushion, chair, rugged, anger, blanket, point, pollution, cotton, thimble, fear, swivel, snore, sweet, feather, rage, sweep, ill, see, frame, office, wrath, sand, doctor, cigar, sleep, fragrance