Neuropsychology
1997, Vol. 11, No. 3, 382–391

0894-4105/97/$3.00

Generation Effects and Source Memory in Healthy Older Adults and in Adults With Dementia of the Alzheimer Type

Kristi S. Multhaup and David A. Balota
Washington University

Recognition and source memory were explored in healthy older adults, adults diagnosed with very mild dementia of the Alzheimer type (DAT), and adults diagnosed with mild DAT. Two sentence-completion tasks were used. In Task 1, half of the sentences were completed (clozed) by the participant, and half by the experimenter. In Task 2, half were participant clozed, and half were participant read (already clozed). Recognition of the cloze words and accuracy of categorizing them as participant generated or experimenter generated (Task 1) and participant generated or participant read (Task 2) were measured (*source discrimination*). Contrary to previous reports, the DAT groups showed the generation effect, that is, better recognition for participant-generated words than experimenter-generated words (Task 1) or read words (Task 2). Source discrimination was disproportionately impaired in the DAT groups.

Memory impairment is the hallmark of dementia of the Alzheimer type (DAT; e.g., Bayles & Kaszniak, 1987). There are, however, other conditions that produce severe memory impairments. Individuals with damage to their frontal lobes or their temporal lobes and Korsakoff patients are examples of other populations that also show memory impairments (e.g., McCarthy & Warrington, 1990; Shallice, 1988). The precise nature of memory impairments varies across different populations. For example, Gabrieli et al. (1994) found that a DAT group showed impaired word stem completion priming whereas global amnesics did not. In addition, Gabrieli et al. reported that the DAT group showed intact priming on incomplete pictures (perceptual priming; for similar perceptual priming effects with words, see Balota & Duchek, 1991; Balota & Ferraro, 1996). On the basis of preserved perceptual priming and impaired nonperceptual priming in their DAT group, Gabrieli et al. argued that repetition priming is composed of perceptual and nonperceptual components that can be dissociated. Therefore, recent neuropsychological studies of memory indicate that thorough examination of different populations' memory performance can reveal important clues regarding the organization of human memory.

The goal of the present study is to provide information concerning the nature of the memory impairments associated with DAT. In pursuit of this goal, the present article explores three major issues. The first issue is whether early-stage DAT individuals show the *generation effect,* which occurs when people better remember information that they had to produce (e.g., producing associates to a word) compared with information that was given to them (e.g., reading words; Hirshman & Bjork, 1988; Slamecka & Graf, 1978). The second issue is whether DAT individuals can discriminate the source of information. As reviewed below, there is inconsistency in the existing literature concerning whether DAT individuals show generation effects and impaired source memory. The third issue is whether there are qualitative differences either in the generation effect or in source memory for individuals at differing degrees of dementia severity.

Generation Effect

There are several reports in the literature that suggest that individuals diagnosed with DAT do not show the generation effect (Dick, Kean, & Sands, 1989a, 1989b; Mitchell, Hunt, & Schmitt, 1986). For example, Mitchell et al. had participants read some sentences (e.g., "The horse jumped the fence") and generate the endings to other sentences (e.g., "The gentleman opened the ______"). Later, participants were given the subjects of the sentences (e.g., "horse"; "gentleman") and asked to recall the object that had been paired with the words in the previous sentences (e.g., "fence"; "door"). Healthy older adults showed higher recall for words that they had generated compared with

Kristi S. Multhaup and David A. Balota, Department of Psychology, Washington University.

Portions of these data were presented at the Second Annual Meeting of the Cognitive Neuroscience Society, March 1995, San Francisco, California. This research was supported by National Institute on Aging Grant AG00030.

Martha Storandt is gratefully acknowledged for allowing us to include our tasks in her neuropsychological test battery and for providing the psychometric data reported in this article. Special thanks go to Martha Storandt's research team, particularly Emily La Barge, for their help in the data collection. John C. Morris, Director, Washington University Alzheimer's Disease Research Center Clinical Core, is gratefully acknowledged for providing diagnostic and dementia severity ratings. We thank Debra A. Grosse-Fleischman for sharing her materials with us. Special thanks go to Greta Munger for her help with the final figures. Thanks also go to Mark E. Faust and to Marcia K. Johnson for helpful discussions of this project.

Correspondence concerning this article should be addressed to Kristi S. Multhaup, who is now at the Department of Psychology, Davidson College, Davidson, North Carolina 28036-1719. Electronic mail may be sent via Internet to krmulthaup@davidson.edu.

words that they had read (i.e., a significant generation effect). In contrast, individuals diagnosed with DAT did not show a difference in recall for words that they had generated and words that they had read. Similarly, Dick and colleagues (Dick et al., 1989a, 1989b) reported four experiments with the same group of 18 DAT individuals and found little evidence for a generation effect in word recall, word recognition, or recall of action instructions (e.g., "Break the toothpick").

In contrast to these studies, at least one report suggests that individuals diagnosed with DAT may show a generation effect, but not as large an effect as that shown by control participants (Fleischman et al., 1995). In that experiment, participants read some words and generated others on the basis of a definition and first-letter cue. Later, they were given a list of old and new words and asked to say yes if a word was one that they had read or generated from a definition and no if it was not. Fleischman et al. reported a Group (control vs. DAT) × Stimulus Type (read vs. generated) interaction indicating that the control group's generation effect (44%) was larger than the DAT group's (18%). The Fleischman et al. data suggest that further examination of generation effects in DAT individuals is warranted because of the practical importance of identifying a procedure than can benefit the memory performance of individuals diagnosed with DAT. Adding to this suggestion are data from a related paradigm in which self-generated cues lead to better recall performance than experimenter-provided cues, for both older adults who are healthy and older adults who have been diagnosed with DAT (Lipinska, Bäckman, Mäntylä, & Viitanen, 1994). Note that this paradigm is slightly different from the generation effect as defined above. In Lipinska et al. (1994), the *cues* were self-generated rather than the *targets* that were later tested. This slight difference noted, the Lipinska et al. data are clearly consistent with the idea that individuals diagnosed with DAT may show generation effects.

Task difficulty may contribute to past failures to find reliable generation effects in DAT groups. Dementia of the Alzheimer type groups tend to show near-floor performance on recall tasks (see Dick et al., 1989a, 1989b; Mitchell et al., 1986), so it may be difficult to detect generation effects in DAT groups with recall tasks. However, Dick et al. (1989b) also failed to find a reliable generation effect with a recognition test, so task difficulty cannot completely explain the inconsistencies in the literature. Severity of the participants' disease may underlie the discrepant findings. The current project investigated whether generation effects are shown by participants diagnosed with very mild DAT or with mild DAT. Severity of dementia is more fully addressed in the Discussion section.

Source Memory

To our knowledge, source memory in DAT has been explored in only three published articles (Dick et al., 1989a, Experiment 2; Goldman, Winograd, Goldstein, O'Jile, & Green, 1994; Mitchell et al., 1986). The authors of these studies drew very different conclusions regarding the impairment or relative preservation of source memory performance in DAT. Consider first the Mitchell et al. experiment. After the cued-recall test described in the Generation Effect section, participants were given a list of subject–object pairs and asked to identify whether each pair was from a sentence they had read or from a sentence for which they had generated the ending. Dementia of the Alzheimer type participants' source discrimination was significantly worse than that of healthy older adults, and it was not different from chance. Note, however, that discriminating between words that were generated and words that were read can be a particularly difficult task (e.g., Rabinowitz, 1989; see also Johnson, Hashtroudi, & Lindsay, 1993). Consistent with the Mitchell et al. data are data from Dick et al. (1989a, Experiment 2). Dick et al. asked participants to indicate whether they had performed an act (e.g., sharpening a pencil) or whether the experimenter had done it. Young and healthy older adults were at ceiling on this source discrimination, whereas the DAT group showed performance roughly at chance level.

In contrast to Dick et al. (1989a) and Mitchell et al. (1986), Goldman et al. (1994) argued that source memory is preserved in early DAT. In their Experiment 1, questions with true answers (e.g., "What city did Elvis Presley live in?" Answer: "Memphis") and questions with made-up answers (e.g., "What type of dance does Carol Burnett like to do?" Answer: "tango") were asked one at a time. If participants did not know the answer to a question, it was provided. Questions with made-up answers were asked twice; the second occurrence was of interest. If participants remembered the answer to a question, they were asked to identify where they had learned that information: a person, TV or radio, print media, the experiment, or some other source. Healthy older adults did not make source errors on this task. The DAT group also made minimal errors (9% for made-up facts; none for true facts). Similar findings were reported in their Experiment 2. Goldman et al. concluded that source memory is not impaired in the early stages of DAT.

There are a number of methodological differences in these experiments that could account for the different conclusions of Goldman et al. (1994) compared with those of Dick et al. (1989a) and Mitchell et al. (1986). For example, the interval between study and test was several minutes in the Dick et al. and Mitchell et al. studies and roughly 30 s for the DAT group in the Goldman et al. study. The type of source discriminations also varied across the studies. It is also possible that the Dick et al. and Mitchell et al. DAT participants were more impaired than the Goldman et al. participants. In addition to the methodological issues, the interpretation of the reported data is complicated by the ceiling performance of the controls in the Goldman et al. study. To understand this second apparent discrepancy in the literature, a study capable of examining DAT participants at different severity levels with the same procedures is needed.

Summary

The present project allowed for the investigation of both generation effects and source memory on two tasks in healthy older adults, adults with very mild DAT, and adults with mild DAT. In the first task, participants generated the endings to some sentences, and the experimenter generated the endings to others. Later, participants were given a forced-choice recognition test for the words used to complete the sentences. After they chose the word that they thought was from the sentences, they were asked whether they or the experimenter had used that word to complete a sentence (source memory measure). The second task was very similar. Participants generated endings to some sentences and read some sentences that were already completed (the last word was underlined). Later, participants were given a forced-choice recognition test for the words that completed the sentences. After they chose the word that they thought was from the sentences, they were asked whether they had used the word to complete a sentence or if the word had already been filled into the blank at the end of a sentence (source memory measure).

It was expected that healthy older adults would show the generation effect, that is, better recognition for words that they had generated compared with words that the experimenter had generated (Task 1) or words that the participants had read (Task 2). Although generation effects in healthy older adults sometimes take several trials to appear (McFarland, Warren, & Crockard, 1985), the similarity of our procedures to those of Mitchell et al. (1986), who reported a generation effect in healthy older adults, led us to expect the effect in our sample. Of interest was whether the participants diagnosed with very mild DAT and those diagnosed with mild DAT would show the generation effect.

By using two different source discriminations, we were able to investigate whether the difficulty of a source discrimination would affect the size of group differences in source memory performance. On the basis of the source monitoring literature (for a review, see Johnson et al., 1993), it was expected that it would be more difficult for participants to discriminate words that they had generated from words that they had read (Task 2) than to discriminate words that they had generated from words that the experimenter had generated (Task 1). This is because the cognitive processes involved in generating a word and reading it probably overlap more than the cognitive processes involved in generating a word and listening to another person generate a word (e.g., Hashtroudi, Johnson, & Chrosniak, 1989). Of interest was whether the source memory impairment of the DAT groups would be disproportionate to their item recognition memory, as has been found for other patient populations (e.g., frontal patients; Janowsky, Shimamura, & Squire, 1989), and whether the degree of source memory impairment would vary with the type of source discrimination.

Method

Participants

Forty-two healthy older adults, 23 very mild DAT adults, and 26 mild DAT adults were recruited from the Washington University Medical School Alzheimer's Disease Research Center, St. Louis, Missouri. All participants were seen by a physician and were screened for neurologic, psychiatric, or medical disorders with the potential to cause dementia. The inclusionary and exclusionary criteria for a diagnosis of DAT conform to those outlined in the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria (McKhann et al., 1984) and have been described in detail elsewhere (e.g., Morris, McKeel, Fulling, Torack, & Berg, 1988). Reported diagnostic accuracy for Alzheimer's disease (AD) has been high (e.g., 96%, AD confirmed in 102 of 106 consecutive autopsies in DAT individuals; Berg & Morris, 1994) when these criteria are used.

Dementia severity for each participant was staged in accordance with the Washington University Clinical Dementia Rating (CDR) Scale (Hughes, Berg, Danziger, Coben, & Martin, 1982; Morris, 1993). On the CDR Scale, a score of 0 indicates no cognitive impairment, a score of 0.5 indicates questionable or very mild dementia, a score of 1 indicates mild dementia, and a score of 2 indicates moderate dementia. At the Washington University Alzheimer's Disease Research Center, CDR scores of 0.5 have been found to accurately indicate the earliest stages of DAT (Morris et al., 1991).

The CDR is based on a 90-min interview with both the patient and his or her collateral source. The interviews are conducted by a board-certified physician, and each interview is videotaped and subsequently reviewed by a second physician for purposes of reliability. The interview is intended to assess cognitive functioning in areas of memory, orientation, judgment and problem solving, community affairs, hobbies, and personal care.

Four healthy older adults, 1 very mild DAT individual, and 1 mild DAT individual were excluded from the analyses because they showed recognition or source memory performance that was over 2.5 standard deviations below their respective group means. Ten healthy older adults and 2 very mild DAT individuals were excluded because of ceiling performance (100% recognition on both tasks or 100% on both the recognition and source memory for at least one task). Thus, data are presented for 73 participants: 28 healthy older adults (mean age = 77.0 years, $SD = 9.9$), 20 very mild DAT adults (mean age = 76.7 years, $SD = 7.6$), and 25 mild DAT adults (mean age = 76.1 years, $SD = 9.2$). The groups included in the analyses showed typical patterns of decreased memory abilities with increasing dementia (see Table 1). Note that when the participants that were outliers or at ceiling were included in analyses, the same basic patterns reported below emerged.

Participants were administered a 2-hr psychometric battery by trained psychometricians who were unaware of the individuals' CDR ratings. Similarly, the data from the psychometric battery were not available to the physicians who gave the CDR ratings. Selected psychometric test data for the groups are shown in Table 1. We assessed memory performance by means of the Wechsler Memory Scale (WMS; Wechsler & Stone, 1973) Logical Memory subscale, Forward Digit Span subscale, Backward Digit Span subscale, and Associate Learning subscale. We assessed lexical retrieval with the Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983) and the Word Fluency Test (Thurstone & Thurstone, 1949). The data were analyzed with a separate analysis of variance (ANOVA) for each measure, which included between-subjects variables of group (healthy old, very mild DAT, or mild DAT) and order (Task 1 first or Task 2 first; see the *Procedure* section). There were no effects of order, and order did not interact with group for any of the measures. The F and p values for the group effects are included in Table 1.

Materials

To optimize the sensitivity of the item memory measure, we used a recognition task rather than a recall task. Also, careful piloting

Table 1
Psychometric Data on Participants by Group

Test	Group			$F(2, 67)$[a]
	Healthy controls	Very mild DAT	Mild DAT	
WMS Logical Memory				
M	9.18	3.83	2.14	48.29*
SD	2.88	3.43	1.62	
WMS Forward Digit Span				
M	6.71	6.20	6.04	1.45†
SD	1.01	1.28	1.06	
WMS Backward Digit Span				
M	5.07	4.10	3.24	9.08*
SD	1.46	1.45	1.36	
WMS Associate Learning				
M	13.74	8.85	7.24	28.48*
SD	4.26	2.15	2.01	
Boston Naming				
M	55.82	44.55	35.08	23.11*
SD	4.51	11.37	14.55	
Word Fluency (S & P)				
M	33.21	23.90	18.80	16.59*
SD	8.40	9.25	7.96	

Note. DAT = dementia of the Alzheimer type; WMS = Wechsler Memory Scale.
[a]From group main effect.
$*p < .001$. $†p = .24$.

was conducted to identify the appropriate study list length, the appropriate distractor items for the recognition test, and the appropriate number of test items, so that ceiling effects in the healthy participants and floor effects in the mild DAT group would be minimized. Materials were a subset of the high-cloze sentences (high agreement across participants for completing a sentence with a particular word) used by Grosse, Wilson, and Fox (1990) and are listed in the Appendix. There were four sets of sentences (eight items per set) that were roughly equivalent in terms of mean cloze values, number of words in the sentences, number of letters in the target words (words used to cloze the sentences), and target-word categories (e.g., verbs or places). Sets were randomly assigned to be (a) clozed by the experimenter for the experimenter–participant task, (b) clozed by the participant for the experimenter–participant task, (c) clozed by the participant for the generate–read task (generated), or (d) read (already clozed) by the participant for the generate–read task. It would have been desirable to rotate each item set through each of the four item types, but the constraints of the battery in which we included our tasks did not allow for this. However, the fact that similar results were found on both tasks (see the Results section) can be considered a within-study replication that argues against idiosyncratic item effects driving our findings.

The distractor items were taken from the distractor list of Grosse et al. (1990). No distractor was semantically related to the target with which it was paired, but it was related to a different target in the set. For example, the distractor *mistake* was paired with the target *room* and was related to a different target, *accident*. This was done because pilot work showed that when distractors were unrelated to any target items in the set, even the DAT groups were at ceiling in their recognition performance.

For both tasks, the acquisition list consisted of 16 sentences, 8 of which were clozed by the participant. The test phase included 16 pairs of words (a target and a distractor). Acquisition lists were prepared by randomly ordering sentences and then adjusting them so that no more than two items from the same source (e.g., experimenter clozed) occurred in a row. On the recognition test, no more than three targets from the same source appeared in a row. In addition, half of any one target type (e.g., experimenter clozed) occurred in the first half of the test list, and half occurred in the second half of the test list. Finally, half of any one target type were the leftmost words in the target–distractor pairs, and half were the rightmost words.

Procedure

Roughly half of the participants did the experimenter–participant task first ($n = 9$, 10, and 13 for CDR 0, 0.5, and 1, respectively), and the others did the generate–read task first ($n = 19$, 10, and 12 for CDR 0, 0.5, and 1, respectively). As noted in the Results section, there were no main effects of order, and order did not interact with any other factors. The two tasks were embedded in a 2-hr psychometric test battery (see the *Participants* section) and were separated by approximately 20 min of psychometric testing.

In the experimenter–participant task, participants were told that they would see sentences that had blanks at the end. They were told that for some of the sentences, they would be asked to read the sentence aloud and fill in the blank at the end with the first word that came to mind and that for other sentences, the experimenter would read the sentence aloud and fill in the blank. Participants were also told that later on they would be asked questions about the words that were used to fill in the blanks. For each sentence, the experimenter placed an index card containing a sentence with a blank at the end of it in front of the participant and while doing so indicated who should read and cloze (complete) the sentence. This was repeated until all 16 sentences were presented. After this acquisition phase, there was a 1-min distractor task. Participants were given a sheet of numbers and asked to circle the even numbers. Next was the test phase. The experimenter placed an index card that contained a pair of words in front of the participant and asked the participant to say which word he or she thought was from the sentences that were read earlier. After making a choice, the participant was asked who had used that word to fill in a blank, the experimenter or the participant.

The format of the generate–read task was a slight modification of the experimenter–participant task. In the acquisition phase, participants were told that for some of the sentences, they would be asked to read the sentence aloud and fill in the blank at the end with the first word that came to mind and that for other sentences, the sentence would be completed with an underlined word and they should just read those sentences aloud. Participants were also told that later on they would be asked questions about the underlined words and the words that they had used to fill in the blanks. After this acquisition phase, there was a 1-min distractor task. Participants were given a sheet of numbers and were asked to circle the odd numbers. Next was the test phase. The experimenter placed an index card that contained a pair of words in front of the participants and asked the participants to say which word they thought was from the sentences that were read earlier. After making a choice, they were asked whether that word was an underlined word or whether they had said it to fill in a blank.

Results

Unless otherwise noted, all analyses were mixed-factor ANOVAs. The analyses included the between-subjects factors of group (healthy old, very mild DAT, mild DAT) and order (experimenter–participant task first or generate–read task first) and the within-subjects factor of task (experimenter–participant task or generate–read task) or source (experimenter generated vs. participant generated or generated vs. read). There were no main effects of order, and order did not interact with any other factors.

Cloze Task

During the acquisition phase of each task, participants were asked to complete eight high-cloze sentences. The percentage of completions with target words was computed for each participant for each task. Analysis of these percentages revealed a main effect of group, $F(2, 67) = 5.34$, $MSE = 344.68$, $p < .01$, and a main effect of task, $F(1, 67) = 5.36$, $MSE = 93.57$, $p < .05$. A Newman–Keuls test on the group means revealed that healthy older adults completed a higher percentage of sentences (92%) than did mild DAT individuals (79%), $p < .01$, and that the very mild DAT individuals did not differ from the other groups (85%). Participants had a slightly higher completion percentage on the experimenter–participant task (88%) than on the generate–read task (84%). Items that were not clozed with the target word were excluded from calculations of recognition and source memory performance.

Overall Recognition

The recognition performance (percentage correct) of the groups on both tasks can be seen in Figure 1. As shown in Figure 1, recognition decreased with increased dementia, and recognition performance was very similar across tasks. An ANOVA confirmed these observations. The only significant effect was group, $F(2, 67) = 57.52$, $MSE = 143.14$, $p < .001$. A Newman–Keuls test on the group means revealed that healthy older adults (94%) showed higher recognition performance than very mild DAT individuals (77%; $p < .01$), who, in turn, showed higher recognition performance than mild DAT individuals (67%; $p < .01$). Note that the mild

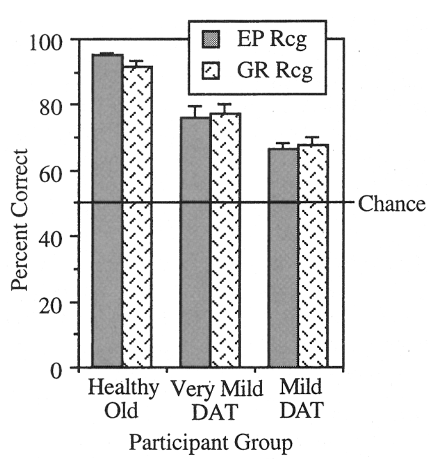

Figure 1. Recognition (percentage correct) by group on the experimenter–participant task (EP Rcg) and on the generate–read task (GR Rcg). DAT = dementia of the Alzheimer type.

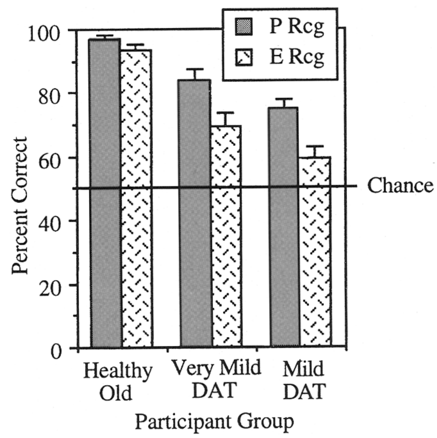

Figure 2. Recognition (percentage correct) by group on experimenter-generated words (E Rcg) and on participant-generated words (P Rcg). DAT = dementia of the Alzheimer type.

DAT individuals did recognize old words at a level above chance (50%) in both the experimenter–participant task (66%), $t(24) = 7.22$, $p < .001$, and the generate–read task (67%), $t(24) = 6.96$, $p < .001$.

Generation Effects

Experimenter–participant task. The recognition performance (percentage correct) of the groups for items clozed by the experimenter and items clozed by the participant can be seen in Figure 2. As indicated in the overall analysis, recognition decreased as dementia increased. Most important, it is clear from Figure 2 that all groups showed the generation effect, namely, better recognition for words they had generated (shaded bars) compared with words the experimenter had generated (speckled bars).

An ANOVA confirmed these observations. There was a main effect of group, $F(2, 67) = 44.01$, $MSE = 221.50$, $p < .001$, and a main effect of source, $F(1, 67) = 26.93$, $MSE = 161.65$, $p < .001$. A Newman–Keuls test on the group means revealed that healthy older adults (95%) showed higher recognition performance than very mild DAT individuals (77%; $p < .01$), who, in turn, showed higher recognition performance than mild DAT individuals (67%; $p < .01$). Participants better recognized words they had generated themselves (86%) compared with words the experimenter had generated (75%). There was also a marginal Group × Source interaction, $F(2, 67) = 2.97$, $MSE = 161.65$, $p < .06$, which reflects the somewhat reduced generation effect of the healthy older adults compared with the other groups, most likely because they were approaching ceiling performance.

Participants showed the generation effect at an individual level as well as at the group level. Each participant could show a generation effect (better recognition for items he or she had generated compared with items the experimenter had generated), a reverse-generation effect (better recognition for items the experimenter had generated compared with items he or she had generated), or no effect. Of the 17 healthy older adults who were not at ceiling performance in recognition, 11 (65%) showed the generation effect. Of the 19 very mild DAT individuals not at ceiling performance, 15 (79%) showed the generation effect. Of the 25 mild DAT individuals, 17 (68%) showed the generation effect.

Generate–read task. The recognition performance (percentage correct) of the groups for items clozed by the participant and read by the participant (already clozed) can be seen in Figure 3. In general, the data are very consistent with the experimenter–participant task data. As indicated in the overall analysis, recognition decreased as dementia increased. Most important, it is clear from Figure 3 that all groups showed the generation effect.

An ANOVA confirmed these observations. There was a main effect of group, $F(2, 67) = 23.37$, $MSE = 250.99$, $p < .001$, and a main effect of source, $F(1, 67) = 50.03$, $MSE = 164.59$, $p < .001$. A Newman–Keuls test on the group means revealed that healthy older adults (92%) showed higher recognition performance than very mild DAT individuals (78%; $p < .01$), who, in turn, showed higher recognition performance than mild DAT individuals (69%; $p < .05$). Participants better recognized words they had generated themselves (88%) compared with words they had read (73%). The Group × Source interaction did not approach significance ($F < 1$).

As on the experimenter–participant task, participants showed the generation effect at an individual level as well as at the group level. Again, each participant could show a generation effect, a reverse generation effect, or no effect. Of the 21 healthy older adults who were not at ceiling performance in recognition, 17 (81%) showed the generation effect. Of the 19 very mild DAT individuals not at ceiling performance in recognition, 14 (74%) showed the generation effect. Of the 25 mild DAT individuals, 21 (84%) showed the generation effect.

Source Memory

After participants selected which of two words had been in the sentence-completion task (recognition), they were

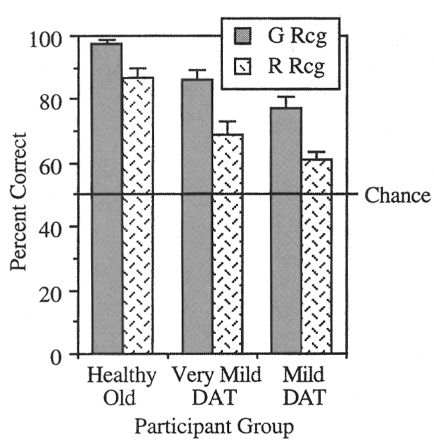

Figure 3. Recognition (percentage correct) by group on participant-generated words (G Rcg) and on read (already clozed) words (R Rcg). DAT = dementia of the Alzheimer type.

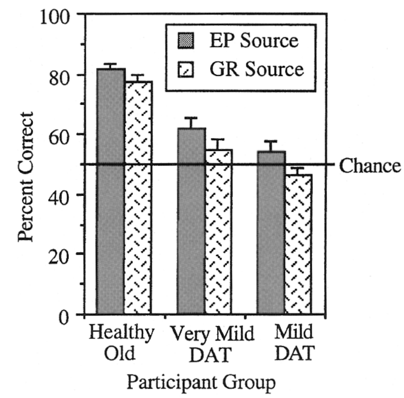

Figure 4. Source memory (percentage correct) by group on the experimenter–participant task (EP Source) and on the generate–read task (GR Source). DAT = dementia of the Alzheimer type.

asked to indicate the source of that word, namely, that it was either experimenter generated or participant generated (for the experimenter–participant task) or that it was participant generated or participant read (generate–read task). In scoring the data, source memory performance was evaluated for items that participants correctly recognized (i.e., source memory was conditional on item memory).

The source memory performance (percentage correct) of the groups on both tasks can be seen in Figure 4. Source memory decreased with increasing dementia. Also, source memory performance was better on the experimenter–participant task (shaded bars) compared with the generate–read task (speckled bars), as predicted by the source monitoring framework (Johnson et al., 1993). An ANOVA confirmed these observations. There was a main effect of group, $F(2, 67) = 39.60$, $MSE = 268.90$, $p < .001$, and a main effect of task, $F(1, 67) = 8.89$, $MSE = 173.23$, $p < .01$. A Newman–Keuls test on the group means revealed that healthy older adults (79%) showed higher source memory performance than very mild DAT individuals (58%; $p < .01$) and mild DAT individuals (50%; $p < .01$). Source memory performance was better on the experimenter–participant task (67%) than on the generate–read task (60%). The Group × Task interaction did not approach significance ($F < 1$). The very mild DAT group showed source memory performance above chance (50%) for the experimenter–participant task (62%), $t(19) = 3.32$, $p < .01$, but their source memory performance was at chance level on the generate–read task (55%), $t(19) = 1.29$, $p > .05$. Note that this pattern is consistent with the prediction that it would be easier to discriminate between the experimenter and participant sources than between the read and generate sources. The mild DAT group was at chance level in their source memory performance for both the experimenter–participant task (54%), $t(24) = 1.10$, $p > .05$, and the generate–read task (46%), $t(24) = -1.34$, $p > .05$.

As stated in the introduction, one of the interesting questions about source memory performance in individuals diagnosed with DAT is whether any source memory impairments that they show are greater than one would expect based on their item memory impairments. Given that the groups differed in their item memory (recognition scores), one way to examine whether individuals diagnosed with DAT show disproportionate source memory impairments, given their item memory impairments, is to look at median splits of the groups based on participants' recognition scores (item memory). This was done separately for the experimenter–participant and generate–read tasks (see Figures 5 and 6, respectively). Fortunately, these median splits produced groups that were comparable on overall recognition memory but varied in level of dementia. The comparisons of interest were (a) the bottom half of the healthy older group with the top half of the very mild DAT group and (b) the bottom half of the very mild DAT group with the top half of the mild DAT group.

Consider first the experimenter–participant task (see Figure 5). A Newman–Keuls test on group item recognition scores revealed that all comparisons were significant ($p < .05$), except for the comparison between the bottom half of the healthy older adults and the top half of the very mild DAT group (92% and 87%, respectively). Thus, the bottom half of the healthy older adults and the top half of the very mild DAT group showed equivalent item recognition and allowed investigation of whether DAT is associated with disproportionate disruption of source memory. The Newman–Keuls test on group source memory scores revealed that the source memory of the top very mild DAT recognizers (65%) was lower than the source memory of the bottom healthy older adult recognizers (82%; $p < .05$). Moreover, a second noteworthy comparison was between the top half of the mild DAT group and the bottom half of the very mild DAT group. Here, the top mild DAT individuals (76%) showed better item recognition than the bottom very mild DAT individuals (66%), and yet source memory was equivalent for the two groups (56% and 58%, respectively).

The data from the generate–read task were very similar (see Figure 6). A Newman–Keuls test on group item recognition revealed that all comparisons were significant ($p < .01$), except for the comparison between the bottom half of the healthy older adults and the top half of the very mild DAT group (both were 87%). The Newman–Keuls test on source memory revealed that the source memory of the top very mild DAT recognizers (62%) was lower than the source memory of the bottom healthy older adult recognizers (75%; $p < .05$). As in the experimenter–participant task data, the top mild DAT individuals showed better item recognition (78%) than the bottom very mild DAT individuals (67%), and yet source memory was equivalent for the two groups (46% and 47%, respectively).

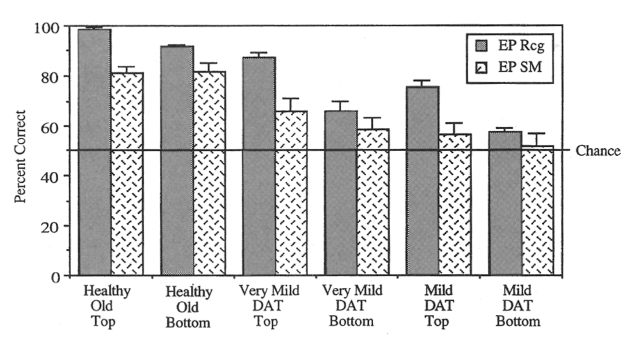

Figure 5. Percentage correct on recognition (EP Rcg) and source memory (EP SM) by group on the experimenter–participant task. Top = top half of the median split for that group; bottom = bottom half of the median split for that group. DAT = dementia of the Alzheimer type.

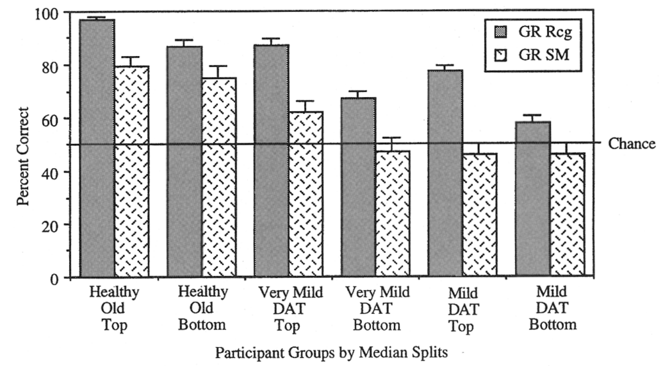

Figure 6. Percentage correct on recognition (GR Rcg) and source memory (GR SM) by group on the generate–read task. Top = top half of the median split for that group; Bottom = bottom half of the median split for that group. DAT = dementia of the Alzheimer type.

Thus, although the top very mild DAT group showed similar item recognition to the bottom healthy older adult group, the top very mild DAT group showed worse source memory than the bottom healthy older adult group (see Figures 5 and 6). This pattern of data is consistent with the idea that DAT is associated with source memory impairment that is disproportionate to the level of item recognition impairment. In the other comparison of interest, although the top mild DAT group showed better item recognition than the bottom very mild DAT group, the groups were similar in their source memory performance (see Figures 5 and 6). The near-floor performance of these two groups on source memory scores does indicate that these data must be interpreted with caution. Even so, it is worth noting that this pattern of data is also consistent with the idea that the progression of DAT is associated with a source memory impairment that is disproportionate to the level of item recognition impairment.

Discussion

We now return to the issues raised in the introduction. The first issue addressed was whether DAT individuals produce generation effects. The present data indicate that at least within the present experimental paradigm, DAT individuals produced rather large generation effects. This finding is consistent with Fleischman et al. (1995) and Lipinska et al. (1994) but does contrast with the conclusions of Dick et al.

(1989a, 1989b) and Mitchell et al. (1986). An important factor that may account for the discrepancy is severity of DAT. Dick et al. (1989b, Experiment 2) noted that when their DAT group was split on the basis of Mini-Mental State Examination (MMSE) scores, only the high-MMSE group showed a significant generation effect in word recognition, the magnitude of which was at least as large as the present DAT groups'. It is possible that the Dick et al. (1989b) low-MMSE group had progressed even further than our mild DAT group and that the ability to benefit from generation decreases later in the disease. This may also apply to the Mitchell et al. DAT data that were collapsed across individuals diagnosed with mild and moderate DAT; classification was made, in part, with the same dementia scale used in the present work (Hughes et al., 1982). It is possible that the Mitchell et al. DAT group was composed primarily of individuals with moderate dementia who, like the Dick et al. low-MMSE group, may have progressed in DAT past the stages in which they could benefit from generation. In contrast, the current project investigated both very mild and mild DAT, and Fleischman et al. (1995) and Lipinska et al. (1994) studied mild DAT. Thus, it appears that individuals diagnosed with DAT can benefit from generation through the very mild and mild stages of the disease but that this benefit may disappear in more severe stages of the disease. However, note that the present DAT individuals were consistently lower on overall recognition performance and on the psychometric tests, compared with the control individuals. Thus, it is not the case that we simply have a group of unimpaired DAT individuals.

The interesting, and potentially practical, point is that individuals in the early stages of DAT (very mild and mild DAT) can benefit from generating information. Although it has been reported that DAT groups do not benefit from other encoding manipulations (e.g., depth of processing, imagery, or verbal elaboration) that affect memory performance in healthy adults (Bayles & Kaszniak, 1987, but see Bird & Luszcz, 1993, and Nebes, 1992), generating information does affect the memory performance of individuals in the early stages of DAT. Given the daily frustrations that DAT individuals must face, any type of mnemonic that could increase memory performance and reduce frustration would be helpful. The present data show generation effects for words (not unlike short shopping lists) and suggest that investigation involving other materials may be promising.

The second issue addressed was whether individuals diagnosed with DAT can discriminate information that they generated themselves from information that they received from some other source. The data clearly show that individuals diagnosed with DAT show source memory impairments, consistent with the reports of Dick et al. (1989a) and Mitchell et al. (1986). Although Goldman et al. (1994) argued that source memory is spared in the early stages of DAT, we found impairment in our very mild DAT group as well as in our mild DAT group. Goldman et al.'s conclusion was based on the few source memory errors that their DAT group made. We believe that their pattern of data was likely due in large part to their very brief retention interval (one intervening question) and that for retention intervals of at least several minutes, as in the present study, even individuals in the very early stages of DAT show source memory impairments.

The present data also suggest that source memory is disproportionately disrupted in DAT. For example, a subset of the very mild DAT participants showed recognition levels that were comparable to a subset of the healthy older adults, yet the very mild DAT group showed significantly lower source memory performance than the healthy older adults. The fact that source memory is impaired in DAT is not surprising in and of itself, but the fact that the source memory impairments in DAT are disproportionate to item recognition impairments is of interest. Such a pattern has also been reported for frontal patients (e.g., Janowsky et al., 1989). The general similarity of the source memory disruption in the present early-stage DAT groups and frontal patients contrasts with the more traditional view that the frontal lobes are relatively spared in the early stages of DAT (e.g., Martin, 1990) but is quite consistent with the growing evidence of frontal involvement in DAT (e.g., Morris et al., 1996; Morrison et al., 1986; Parasuraman & Haxby, 1993; Spieler, Balota, & Faust, 1996).

The final issue addressed in this study was whether generation effects and source memory differ for very mild DAT individuals and those for whom the disease has progressed further. Regarding generation effects, the very mild and mild DAT groups showed similar generation effects on both the experimenter–participant and generate–read tasks. Regarding source memory, both groups showed disproportionately impaired source memory performance. Thus, although the mild DAT group showed poorer memory performance than the very mild DAT group, the general patterns were very consistent across the two levels of DAT tested in this study.

In summary, the present data suggest that the early stages of DAT are associated with source memory impairments that are disproportionate to item memory impairments. In contrast, the benefits of generating information are preserved in the very mild and mild stages of DAT. Thus, the present data point to a potentially useful mnemonic for those individuals in the early stages of DAT and, moreover, indicate that encoding processes are not completely disrupted in the early stages of DAT (see also Bird & Luszcz, 1993).

References

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

Balota, D. A., & Ferraro, F. R. (1996). Lexical, sublexical, and implicit memory processes in healthy young and healthy older adults and in individuals with dementia of the Alzheimer type. *Neuropsychology, 10,* 82–95.

Bayles, K. A., & Kaszniak, A. W. (1987). *Communication and cognition in normal aging and dementia.* Austin, TX: pro·ed.

Berg, L., & Morris, J. C. (1994). Diagnosis. In R. D. Terry, R. Katzman, & K. Bick (Eds.), *Alzheimer's disease* (pp. 9–25). New York: Raven Press.

Bird, M., & Luszcz, M. (1993). Enhancing memory performance in Alzheimer's disease: Acquisition assistance and cue effectiveness. *Journal of Clinical and Experimental Neuropsychology, 15,* 921–932.

Dick, M. B., Kean, M.-L., & Sands, D. (1989a). Memory for action events in Alzheimer-type dementia: Further evidence of an encoding failure. *Brain and Cognition, 9,* 71–87.

Dick, M. B., Kean, M.-L., & Sands, D. (1989b). Memory for internally generated words in Alzheimer-type dementia: Breakdown in encoding and semantic memory. *Brain and Cognition, 9,* 88–108.

Fleischman, D. A., Gabrieli, J. D. E., Reminger, S., Rinaldi, J., Morrell, F., & Wilson, R. (1995). Conceptual priming in perceptual identification for patients with Alzheimer's disease and a patient with right occipital lobectomy. *Neuropsychology, 9,* 187–197.

Gabrieli, J. D. E., Keane, M. M., Stanger, B. Z., Kjelgaard, M. M., Corkin, S., & Growdon, J. H. (1994). Dissociations among structural–perceptual, lexical–semantic, and event–fact memory systems in Alzheimer, amnesic, and normal subjects. *Cortex, 30,* 75–103.

Goldman, W. P., Winograd, E., Goldstein, F. C., O'Jile, J., & Green, R. C. (1994). Source memory in mild to moderate Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 16,* 105–116.

Grosse, D. A., Wilson, R. S., & Fox, J. H. (1990). Preserved word-stem-completion priming of semantically encoded information in Alzheimer's disease. *Psychology and Aging, 5,* 304–306.

Hashtroudi, S., Johnson, M. K., & Chrosniak, L. D. (1989). Aging and source monitoring. *Psychology and Aging, 4,* 106–112.

Hirshman, E., & Bjork, R. A. (1988). The generation effect: Support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 484–494.

Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry, 140,* 566–572.

Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia, 27,* 1043–1056.

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

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

Lipinska, B., Bäckman, L., Mäntylä, T., & Viitanen, M. (1994). Effectiveness of self-generated cues in early Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 16,* 809–819.

Martin, A. (1990). Neuropsychology of Alzheimer's disease: The case for subgroups. In M. F. Schwartz (Ed.), *Modular deficits in Alzheimer-type dementia* (pp. 143–175). Cambridge, MA: MIT Press.

McCarthy, R. A., & Warrington, E. K. (1990). *Cognitive neuropsychology: A clinical introduction.* New York: Academic Press.

McFarland, C. E., Jr., Warren, L. R., & Crockard, J. (1985). Memory for self-generated stimuli in young and older adults. *Journal of Gerontology, 40,* 205–207.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, 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 Service Task Force on Alzheimer's Disease. *Neurology, 34,* 39–44.

Mitchell, D. B., Hunt, R. R., & Schmitt, F. A. (1986). The generation effect and reality monitoring: Evidence from dementia and normal aging. *Journal of Gerontology, 41,* 79–84.

Morris, J. C. (1993). Clinical dementia rating. *Neurology, 43,* 2412–2414.

Morris, J. C., McKeel, D. W., Fulling, K., Torack, R. M., & Berg, L. (1988). Validation of clinical diagnostic criteria for Alzheimer's disease. *Annals of Neurology, 24,* 17–22.

Morris, J. C., McKeel, D. W., Jr., Storandt, M., Rubin, E. H., Price, J. L., Grant, E. A., Ball, M. J., & Berg, L. (1991). Very mild Alzheimer's disease: Informant-based clinical, psychometric, and pathologic distinction from normal aging. *Neurology, 41,* 469–478.

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.

Morrison, J. H., Scherr, S., Lewis, D. A., Campbell, M. J., Bloom, F. E., Rogers, J., & Benoit, R. (1986). The laminar and regional distribution of neocortical somatostatin and neuritic plaques: Implications for Alzheimer's disease as a global neocortical disconnection syndrome. In A. B. Scheibel, A. F. Wechsler, & M. A. B. Brazier (Eds.), *The biological substrates of Alzheimer's disease* (pp. 115–131). New York: Academic Press.

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.

Parasuraman, R., & Haxby, J. V. (1993). Attention and brain function in Alzheimer's disease: A review. *Neuropsychology, 7,* 242–272.

Rabinowitz, J. C. (1989). Judgments of origin and generation effects: Comparisons between young and elderly adults. *Psychology and Aging, 4,* 259–268.

Shallice, T. (1988). *From neuropsychology to mental structure.* New York: Cambridge University Press.

Slamecka, N., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 592–604.

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

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

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

Appendix

Stimulus Materials

Sentence	Cloze value
Experimenter–participant task	
Experimenter clozed	
Dick wrote a chapter in the *book.*	.86
The bad boy was sent to his *room.*	.78
The old milk tasted very *sour.*	.80
They went as far as they *could.*	.96
Three people were killed in a terrible highway *accident.*	.84
Bill jumped in the lake and made a big *splash.*	.99
The child was born with a rare *disease.*	.88
The politician spoke out for law and *order.*	.78
Participant clozed	
Water and sunshine help plants *grow.*	.99
It's unlucky to walk under a *ladder.*	.78
The parents pleaded with their daughter to come *home.*	.88
The new store had a grand *opening.*	.78
The man who didn't eat all day was very *hungry.*	.78
His job was to keep the sidewalk *clean.*	.96
The baby weighed six pounds at *birth.*	.78
He scraped the cold food from his *plate.*	.87
Generate–read task	
Generate (cloze)	
John swept the floor with a *broom.*	.96
The children enjoyed the three ring *circus.*	.78
He has trouble adding and subtracting large *numbers.*	.78
When the power went out the house became *dark.*	.90
The set was so loud he couldn't hear himself *think.*	.79
She called her husband at his *office.*	.85
At first the woman refused but she changed her *mind.*	.99
The orchestra played very pretty *music.*	.78
Read (already clozed)	
His boss refused to give him a *raise.*	.96
Joan boiled the eggs in *water.*	.88
He campaigned so he would win the *election.*	.78
The movie was so jammed they couldn't find a single *seat.*	.99
The basketball players were all very *tall.*	.78
The cows moved from the sun into the *shade.*	.82
After a long wait, the package finally *arrived.*	.78
He can't hear you because he is *deaf.*	.78

Note. These materials are a subset of those used by Grosse, Wilson, and Fox (1990).

Received June 26, 1996
Revision received December 20, 1996
Accepted December 21, 1996 ■