Neuropsychology
2003, Vol. 17, No. 1, 82–92

0894-4105/03/$12.00 DOI: 10.1037/0894-4105.17.1.82

Discourse Analysis of Logical Memory Recall in Normal Aging and in Dementia of the Alzheimer Type

David K. Johnson, Martha Storandt, and David A. Balota
Washington University

This study examined the nature of errors in prose recall made in dementia compared with normal aging. Responses by 48 young adults, 47 nondemented older adults, and 70 people with very mild or mild Alzheimer's disease to the Logical Memory subtest of the Wechsler Memory Scale were examined in a propositional analysis. Compared with young adults, healthy older adults showed good immediate recall but deficits in retention over a delay. Demented individuals made errors of omission, not commission, at immediate recall. These errors probably reflect difficulty with attentional control rather than memory per se. In terms of clinical implications, veridical scoring of the Logical Memory subtest provides more sensitive detection of very mild dementia of the Alzheimer type than the current standard criteria for scoring.

Deficits in prose recall, as measured by the number of story units of the Logical Memory subtest of the Wechsler Memory Scale (WMS; Wechsler, 1945) recalled word for word, are prominent in the early stages of dementia of the Alzheimer type (DAT). The WMS Logical Memory subtest was the largest contributor to discrimination between healthy older adults and individuals with very mild dementia (Storandt & Hill, 1989); its beta weight was twice as large as that of any other predictor from a 1.5-hr battery of standard neuropsychological tests. It also was a predictor of future progression of healthy older adults to DAT (E. H. Rubin et al., 1998).

The nature of the errors that produce the declines in prose recall with dementia, however, has not been well studied. To our knowledge, only one study has conducted an error analysis of Logical Memory data of older adults with DAT. Haut, Demarest, Keefover, and Rankin (1994) reported that older adults with DAT were less able to recall thematically important ideas during recall and recognition. These findings parallel the psycholinguistic studies that have examined qualitative differences in the generation (not recall) of prose discourse by dementing older adults. Three studies reported semantic deficits experienced in the earliest stages of the disease, whereby dementing older adults were less able to generate meaningful test responses; this effect increased as the dementia became more severe (Ellis, 1996; Kemper et al., 1993; Lyons et al., 1994). Overall, the content of discourse generated by individuals with DAT is semantically impoverished and lacks organizational coherence; however, the grammatical quality of the discourse is relatively spared.

The original scoring instructions for the Logical Memory subtest did not specify criteria for successful completion of an idea unit. This failure allowed judges to use subjective criteria to evaluate recalled idea units and detracted from measurement reliability (Crosson, Hughes, Roth, & Monkowski, 1984; Larrabee, 1987). Limiting acceptable responses to word-for-word recall, as suggested by Russell (1975), increased both reliability and generalizability of results. Others have attempted to resolve the scoring ambiguity in other ways. For instance, some investigators (Haaland, Linn, Hunt, & Goodwin, 1983; Powers, Logue, McCarty, Rosenstiel, & Ziesat, 1979) scored gist recall with half points and verbatim recall with whole points. Still others (Schear, 1986; Sweet & Kolden, 1986) gave half points for incomplete verbatim responses (i.e., the respondent needed to recall only some of the verbatim idea unit to get credit). Abikoff and colleagues (1987) noted, however, that verbatim scoring is reliably better than gist in its sensitivity to clinical populations suffering memory problems.

Prose recall is rarely word for word, even in young adults. Recall is more typically filled with inexact interpretations of the passage, but it is difficult to define these errors of commission operationally. To better understand the semantic deficits of dementing adults as well as the underlying mechanisms that make the Logical Memory subtest a sensitive clinical predictor of DAT, we conducted an error analysis of archived Logical Memory recall data from the Washington University Alzheimer's Disease Research Center using a scoring protocol for the Logical Memory subtest based on propositional units.

David K. Johnson, Martha Storandt, and David A. Balota, Department of Psychology, Washington University.

These results were presented at the 108th Annual Convention of the American Psychological Association, Washington, DC, August 2000. This research was supported by National Institute on Aging Grants AG 00030, AG 03991, and AG 05681 and by the Richard and Mildred Poletsky Education Fund. Diagnosis and staging of dementia were provided by the Clinical Core of the Washington University Alzheimer's Disease Research Center, John C. Morris, Director.

Correspondence concerning this article should be addressed to Martha Storandt, Department of Psychology, Washington University, St. Louis, Missouri 63130. E-mail: mstorand@artsci.wustl.edu

Propositional Analysis

Propositional analysis is a well-developed method for decomposing text and prose into smaller units of language called propositions; the method retains the rich interconnections among the various aspects of meaning conveyed (Kintsch, 1994; Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975; Kintsch & van Dijk, 1978). A proposition is defined as the smallest unit of discourse that still retains meaning (Turner & Greene, 1977). Discourse theory asserts that propositional elements are concepts, rather than words from the text. Therefore, the exact content of any single proposition may vary depending on its use, but the lexical–semantic relations among its elements do not change (Turner & Greene, 1977). Thus, propositional decomposition of the text provides a set of definable relations among story elements that are the units of measurement in this error analysis. This approach to text analysis has been very productive in predicting reading time, comprehension measures, and recall performance (e.g., Kintsch & van Dijk, 1978; Kintsch & Vipond, 1979).

As an example of how propositional analysis captures the semantic interrelatedness of prose, consider the last sentence from the text of Story A of the WMS Logical Memory subtest: "The officers . . . made up a purse for her."[1] The descriptive clause "touched by her story" has been omitted to simplify this example. The propositional decomposition of this statement yielded five propositions, as shown in Table 1. Each proposition has two parts: the central concept (case relation) and its details (arguments). The central concept of the first proposition refers to *who* (i.e., the actor). The exact nature of the actor (i.e., the officers) is the argument, or detail, of this proposition. The central concept of the second proposition is the action taken by the actors: *made up* (i.e., created). This predicate needs three details to specify its particular use: who (Person 1) made up what (a gift) for whom (Person 2). Thus, the central concept of the first proposition is a detail of another proposition (Person 1 who took the action in Proposition 2).

Note that the propositional decomposition is distinct from a diagrammed sentence. Although concepts such as subject, verb, object, and indirect object are common to both, propositional decomposition specifies how each concept is related to other concepts within the context of the story as a whole. This is achieved through the redundancy of the detail or arguments throughout the entire set of story propositions. For example, the detail *gift* in Propositions 2, 3, and 4 in this example is redundant. The central concept in Proposition 3 specifies what object was made up (i.e., created) in Proposition 2, and the central concept of Proposition 4 is the purpose of the action: *for*.[2] The central concept in Proposition 5 specifies the recipient of the gift, in this case the pronoun *her*. This redundancy of detail, or arguments, among propositions delineates the exact relations between story propositions as well as specifies content.

An important byproduct of the propositional decomposition of the text is that the number of units of measurement increases. For instance, the original decomposition of the Logical Memory passages (Wechsler, 1945) contained multiple concepts (i.e., propositions) per one unit of measurement. An example of an original scoring unit was *little children.* Does one need to recall the fact that the protagonist had children, the fact that those children were young, or both facts together?

Propositional decomposition eliminates much of the ambiguity about how strict or relaxed scoring criteria should be for any single unit. Further, recalled propositions do not need to be word-for-word repetitions of the whole phrase. Instead recall can contain veridical report of some relationships but err with respect to others. Once the propositional decomposition of the original text is determined by an expert, examiners can then compare subjects' recall of the passage to the standard decomposition in a straightforward and reliable manner. Scoring examples are provided in the Method section.

Table 1
Propositional Decomposition of a Phrase From the WMS Logical Memory Story A: "The Officers . . . Made Up a Purse for Her"

Proposition	Central concept	Details
1	Reference (specifies who)	Person 1 = officers
2	Predicate (made up/created)	Person 1
		Gift
		Person 2
3	Is a (specifies what was made)	Gift = purse
4	Purpose (what was for whom)	Gift
		Person 2
5	Reference (specifies recipient)	Person 2 = her

Note. WMS = Wechsler Memory Scale (Wechsler, 1945).

Hypotheses

Our major purpose was to examine the nature of the errors in prose recall made by people in the early stages of DAT. Given that previous analyses of prose production suggested that the grammatical quality of discourse is relatively spared in the early course of DAT (Ellis, 1996; Haut et al., 1994; Kemper et al., 1993; Lyons et al., 1994), we hypothesized that the grammatical relationships between items would be relatively maintained but that memory for the specific details of the stories would be diminished. We also wanted to know whether measuring such deficits by way of a propositional analysis would improve discrimination of healthy older adults from those in the very mild stages of the disease.

A second goal was to determine whether DAT-related deficits in prose recall (errors) occur for the same reason as age-related deficits in prose recall. If true, then similar cognitive processes might be implicated in the two processes—aging and Alzheimer's disease. That is, older and younger adults tend to recall similar items as well as make

[1] Reproduced by permission.

[2] Contrast this meaning with sentences in which *made up* is followed by *to* or *with.* The use of the word *for* connotes a gift rather than flattery (made up to) or reconciliation (made up with).

similar errors (Stine & Wingfield, 1988), but healthy older adults experience age-related decrements in their ability to remember facts from short stories and other prose passages (G. Cohen, 1979; Hultsch, Hertzog, Dixon, & Small, 1998; Moenster, 1972; Small, Dixon, Hultsch, & Hertzog, 1999; Taub, 1979). Will the same be true for people in the early stages of DAT, or will their pattern of errors be qualitatively different? Thus, we studied Logical Memory recall in four groups: young adults, healthy older adults, people with very mild DAT, and people with mild DAT.

Method

Participants

Data from three groups of older adults were identified and retrieved from the archival records (February 22, 1984, to June 16, 1991) of an ongoing longitudinal study of healthy aging and DAT. The procedures used in recruitment for this project have been described in detail elsewhere (Berg, Hughes, Danziger, Martin, & Knesevich, 1982). Data from these older adults have been included in numerous research reports from the project including those from Storandt and Hill (1989), Robinson-Whelen and Storandt (1992), and Chapman, White, and Storandt (1997). A fourth group of 48 young adults (16 men and 32 women) was recruited from a volunteer pool maintained in the psychology department at the same university. The young adults received $5 for participation. The Human Subject's Review Panel at the university approved all research participation.

All participants were selected on the condition that they had no known prior exposure to the Logical Memory stories and had transcripts for both immediate and delayed recall of the stories. In addition, older adult participants were evaluated for dementia within 2 months of test administration, and older adults suffering from reversible dementia or other medical or psychiatric conditions that might cause cognitive impairment other than DAT were excluded. This information is maintained as part of the clinical assessment for participation in the longitudinal study.

Experienced research physicians assessed each older adult for the presence and severity of dementia using the Clinical Dementia Rating (CDR; J. C. Morris, 1993) on the basis of a 90-min semistructured interview with the research participant and a knowledgeable collateral source as well as a neurological examination of the participant. The diagnosis of DAT was based on a history of gradual onset and progressive cognitive impairment and was comparable with that specified in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994). Diagnostic accuracy for Alzheimer's disease, as verified by postmortem examination in 207 individuals, is 93%, including the presence of histologic Alzheimer's disease in 17 or 18 people who died while very mildly demented (Berg et al., 1998).

The CDR describes the severity of dementia along six dimensions: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care. A global CDR is derived from a standard algorithm: A value of 0 indicates no dementia, and 0.5, 1, 2, and 3 indicate very mild, mild, moderate, and severe dementia, respectively. Only people with CDRs of 0, 0.5, and 1 were included in the present study. The CDR has demonstrated high interrater reliability with weighted kappa values ranging from .75 to .94 (Burke et al., 1988; McCulla et al., 1989). The diagnosis of dementia and the subsequent rating of its severity were made without knowledge of psychometric test scores.

The 47 nondemented (CDR 0) older adults included 13 men and 34 women who ranged in age from 54 to 92 years ($M = 73.45$, $SD = 11.13$). The 31 very mildly demented (CDR 0.5) people included 14 men and 17 women who ranged in age from 59 to 90 years ($M = 73.63$, $SD = 8.26$). The 39 mildly demented (CDR 1) individuals (13 men and 26 women) ranged in age from 56 to 84 years ($M = 71.60$, $SD = 7.89$). The three older adult groups did not differ significantly in age, $F(2, 114) = 0.54$, $p > .05$. The 48 young adults were between 18 and 24 years old ($M = 20.00$, $SD = 1.37$).

Although the nondemented older adults and the young adults were well matched for education (14 and 13 years, respectively), the nondemented older adults did have significantly more education than the very mildly and mildly demented older adults, $F(2, 114) = 8.36$, $p < .01$. The very mildly and mildly demented groups had roughly equivalent levels of education (12 and 11 years, respectively). Because of the difference in years of education between the nondemented and demented older adults, analyses of errors that compared these three groups were repeated using education as a covariate; the two discrepancies in results that were obtained are noted in the Results section.

Means and standard deviations for selected tests from the psychometric battery administered to the older adults at entry into the

Table 2
Means and Standard Deviations of Three Older Adult Groups on Selected Measures From a Psychometric Battery

Measure	No dementia		Very mild dementia		Mild dementia	
	M	*SD*	*M*	*SD*	*M*	*SD*
Logical Memory	8.32	2.71	3.69	2.42	1.49	1.58
Digit Span Forward	6.68	1.29	6.10	1.25	5.87	1.24
Digit Span Backward	4.81	1.34	4.00	1.46	2.72	1.68
Associate Learning	12.17	3.11	8.25	3.75	5.64	2.44
Information	20.70	4.06	13.68	5.62	8.67	4.89
Block Design	29.23	8.58	20.06	12.14	10.41	10.47
Digit Symbol	45.83	14.21	29.90	12.89	15.41	14.03
Trailmaking A (s)	51.47	29.27	73.23	38.30	120.41	50.61

Note. The first four measures are from the Wechsler Memory Scale (Wechsler, 1945). Verbatim scoring according to Russell (1975) was used for the Logical Memory subtest. The next three measures are from the Wechsler Adult Intelligence Scale (Wechsler, 1955). The final measure is Part A of the Trailmaking Test (Armitage, 1946).

longitudinal study are shown in Table 2 for descriptive purposes. Values from the standard verbatim scoring (Russell, 1975) of the WMS Logical Memory subtest are shown as well as values from those tests in the battery that typically were administered between immediate and delayed recall of the Logical Memory stories. As has been reported previously (e.g., Storandt & Hill, 1989), performance on all these measures declined significantly as the severity of DAT increased (Fs > 4.72, ps $< .02$). These group differences remained significant when education was controlled through analysis of covariance (ANCOVA).

Measures

The Logical Memory subtest of the WMS (Form I) is included in the standard psychometric assessment battery administered to all participants in the longitudinal study (Storandt, Botwinick, Danziger, Berg, & Hughes, 1984). Participants were read two thematically independent stories (A and B) and were asked to recall each story immediately after hearing it using as many of the same words of the original passage as they could remember, thus encouraging word-for-word recall. Recall was tape recorded and then transcribed word for word, including extraneous utterances unrelated to the task request. Thirty min after the initial presentation, participants were asked to recall all that they could from the first story and then the second story. Participants were unaware of the delayed recall task. As suggested in Russell's (1975) revision of the Logical Memory scoring procedure, all participants were given one cue indicating the topic of the story if the participant could not remember anything (e.g., the first story was about a scrubwoman). Although the number of intervening tasks between immediate and delayed recall varied somewhat, the order of task administration in the battery after the initial Logical Memory presentation was as follows: WMS Digit Span and WMS Associate Learning; two psychomotor speed tasks; and Information, Comprehension, Digit Symbol, and Block Design subtests of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955).

The Logical Memory subtest was administered to the younger adults according to the same procedure used for the older adults. Intervening tasks between the immediate and delayed recall trials included the WAIS Information subtest, a word pronunciation test, and another experimental protocol that required the young adults to answer questions about age biases. The young adults ($M = 20.79$) were well matched on the Information subtest with the nondemented older adults ($M = 20.70$), $t(93) = 0.12$, $p > .05$.

Scoring

The analyses in the present study were based on decomposition of the WMS Logical Memory stories so that, unlike the standard scoring, each unit of measurement contained a single proposition. Every proposition recalled was first scored on a dimension of correctness containing three categories: *veridical* reproduction of the original text, *gist* reproduction (i.e., reproduction of the original text that retained the passage's intended meaning but was not veridical), or *distortion* of the original text (Fletcher, 1994; Johnson-Laird, 1983; Kintsch, 1978, 1994; Zwaan & Radvansky, 1998). This technique expands on the scoring rules used in earlier reports of these data on the basis of Russell's (1975) study by examining what the participants get wrong as well as what they get right. Note that the veridical score is analogous to the verbatim scoring method of the Logical Memory subtest; however, veridical scoring in propositional analysis does allow the use of different verb tenses and does not necessarily have to be word for word, as the more common term verbatim would connote.

The scoring procedure captured two types of commission errors: gist and distortion. Gist responses are counted as correct for some units of the stories in the current version of the WMS (Wechsler, 1997). Further, two additional scores were noted for these errors. First, the location of the error (i.e., gist or distortion) was specified: Did it occur in the central concept, in the details, or in both? Second, if error occurred in the details, for how many details was the recall either gist or distorted? This value could range from 0 (if the error did not occur in the details) to 3 (the maximum number of details in any of the propositions in the stories). Scoring of three sample responses using the propositional analysis is shown in Table 3. If a proposition was not included in the response, the space next to it was left blank for that record.

A team of five raters was trained on propositional structure and scoring procedures for prose recall. Each participant's recall was matched to a corresponding set of propositions present in the literal passages (total = 68). The entire recall (both stories, immediate and delayed) for each participant was decomposed by a given rater and then checked by two other raters working independently. Team members brought all discrepant interpretations of recall to weekly review meetings, and these instances were scored by consensus. To calculate an interrater reliability statistic, a fifth rater rescored a subset of 20 protocols selected at random. The fifth rater's scores were compared with the consensus scores of the four previous raters. Cohen's kappa (a conservative statistic of interrater reliability) averaged .87, ranging from .60 to 1.00, and interrater agreement ranged from 88% to 100% of responses coded.

Results

Error Analysis

Immediate recall. The means and standard deviations of three types of responses at immediate recall across all groups are shown in Table 4. Number of responses, rather than percentages, were analyzed so that the locus of any differences could be identified; differences in percentages depend on the denominator as well as the numerator (J. Cohen & Cohen, 1983, pp. 73–76).

A mixed analysis of variance (ANOVA) with group (young, healthy old, very mildly demented, or mildly demented) as a between-subjects independent variable and response type (veridical, gist, or distortion) as the within-subjects independent variable revealed that the four groups differed significantly in overall production (i.e., total number of propositions attempted, the sum of veridical, gist, and distortion) as indicated by a main effect of group, $F(3, 161) = 101.91$, $p < .01$. There was also a significant main effect of response type, $F(2, 322) = 408.79$, $p < .01$. More important, there was the significant Group $\times$ Response Type interaction, $F(6, 322) = 58.87$, $p < .01$. Recall that the test of this interaction is statistically controlled for the main effects of group and response type. To determine the nature of the interaction, we conducted least significant difference ($\alpha = .05$) pairwise comparisons of the four groups for each of the three response types. The comparisons of interest were (a) of the young adults and the nondemented older adults (age effect) and (b) of the three older adult groups (dementia effect).

Young and healthy older adults did not differ significantly in the number of propositions recalled veridically (i.e., no age effect). There was, however, an effect of

Table 3
Scoring Examples for the Propositions Defined in Table 1

No.	Concept	Details	Response	Errors: Central	Errors: Details
		Sample response: "The police made up a purse for the woman"			
1	Reference	Person 1 = officers	Gist	No	1
2	Predicate (made up)	Person 1, gift, Person 2	Veridical		
3	Is a	Gift = purse	Veridical		
4	Purpose (for)	Person 2	Veridical		
5	Reference	Person 2 = her	Gist	No	1
		Sample response: "The police gave money to her"			
1	Reference	Person 1 = officers	Gist	No	1
2	Predicate (made up)	Person 1, gift, Person 2	Gist	Yes	0
3	Is a	Gift = purse	Gist	No	1
4	Purpose (for)	Gift, Person 1	Gist	Yes	0
5	Reference	Person 2 = her	Veridical		
		Sample response: "The police gave her food"			
1	Reference	Person 1 = officers	Gist	No	1
2	Predicate (made up)	Person 1, gift, Person 2	Gist	Yes	0
3	Is a	Gift = purse	Distortion	No	1
4	Purpose (for)	Gift, Person 2			
5	Reference	Person 2 = her	Veridical		

dementia. As severity of DAT increased, the number of veridical responses decreased (CDR 0 > 0.5 > 1). The same pattern was found for gist responses. There was no age effect, but gist responses decreased with increasing dementia severity (CDR 0 > 0.5 > 1).

To provide more information about the location of the gist response errors made by the demented individuals, we used a mixed ANOVA to compare gist responses made with case relations versus arguments (within-subjects independent variable) in the three older adult groups (between-subjects independent variable). Recall that the case relation of a proposition relates to its central concept, whereas the arguments provide detail. As already indicated, the number of gist errors decreased with dementia severity (the between-subjects group effect), $F(2, 114) = 34.13$, $p < .01$. As would be expected on the basis of the comparatively greater frequency of arguments (details), gist error responses involved arguments more often than case relations, $F(1, 114) = 68.38$, $p < .01$, although this effect was greatly reduced when education was included as a covariate, $F(1, 113) = 3.54$, $p < .07$. But the main effects were not the focus of this analysis. The hypothesized Group × Response Type interaction, however, was not significant, $F(2, 114) = 2.90$, $p < .07$. The decline in gist responses with increasing dementia severity was proportional for case relations and arguments.

The least significant difference pairwise comparisons revealed a different pattern of results for distortions. There was an age effect but no effect of dementia. Although the number of distortions was generally low, young adults pro-

Table 4
Means and Standard Deviations for Measures of Immediate Recall (Stories A and B Collapsed)

Recall	Young *M*	Young *SD*	Nondemented old *M*	Nondemented old *SD*	Very mild dementia *M*	Very mild dementia *SD*	Mild dementia *M*	Mild dementia *SD*
No. attempted	30.44	7.51	27.53	7.01	14.39	7.63	6.69	6.43
Veridical	20.98	6.48	19.57	6.43	8.81	5.47	3.26	3.62
Gist	6.04	2.55	5.68	2.67	2.90	2.34	1.69	2.07
Case relation	2.94	1.58	2.03	1.60	0.84	1.00	0.49	0.79
Argument	4.06	2.01	4.38	2.06	2.26	2.05	1.56	2.00
Distortion	3.42	2.18	2.28	1.64	2.68	2.23	1.74	2.10
Case relation	1.17	1.19	0.79	1.04	1.00	1.39	0.59	1.07
Argument	3.98	2.79	2.26	1.76	2.94	2.62	2.05	2.60

Note. Multiple arguments could be scored as gist or as distortion in a single proposition; therefore, case relation and argument errors may sum to more than the total scored as gist or distortion.

Table 5
Means and Standard Deviations for Measures of Delayed Recall (Stories A and B Collapsed)

Recall	Young		Nondemented old		Very mild dementia		Mild dementia	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
No. attempted	25.25	9.00	21.66	7.09	5.77	7.13	2.05	3.30
Veridical	14.94	7.16	11.83	5.49	2.42	4.83	0.54	1.05
Gist	6.94	2.93	7.04	2.86	2.10	2.56	0.62	1.11
Case relation	3.48	1.94	2.62	1.74	0.65	0.95	0.13	0.41
Argument	4.77	2.33	5.55	2.34	1.90	2.55	0.56	0.96
Distortion	3.38	2.04	2.79	2.12	1.26	1.55	0.90	1.59
Case relation	1.40	1.25	1.23	1.24	0.61	0.95	0.46	0.91
Argument	3.63	2.38	2.72	2.08	1.45	1.80	1.21	2.17

Note. Multiple arguments could be scored as gist or as distortion in a single proposition; therefore, case relation and argument errors may sum to more than the total scored as gist or distortion.

duced slightly more distortions than did healthy older adults (*M*s = 3.42 vs. 2.28, respectively). These distortions were of arguments ($p < .01$) but not of case relations ($p > .05$).

Thus, the omnibus Group × Response Type interaction probably had two sources. First, there was a much less severe drop in the gist responses with increasing dementia severity than in veridical responses. This occurred primarily because the number of gist responses was not large to begin with in the young and nondemented older group (about 6) as compared with the veridical responses (about 20). Second, there was a reversal of the age and dementia severity effects for one of the response types. There were no significant age effects for veridical and gist responses, but the young group produced more distortions than did the nondemented older group. Among the three older adult groups, there were significant effects of dementia severity on veridical and gist responses but none on distortions.

Delayed recall. Similar analyses were conducted for delayed recall (see Table 5 for means and standard deviations). Both the main effects of group, $F(3, 161) = 108.55$, $p < .01$, and response type, $F(2, 322) = 108.76$, $p < .01$, were statistically significant, as was their interaction, $F(6, 322) = 34.57$, $p < .01$. The effect of response type was eliminated when education was included as a covariate, $F(2, 320) = 0.96$, $p > .05$, although the main effect of group remained as did the significant Group × Response Type interaction.

Least significant difference ($\alpha = .05$) pairwise comparisons revealed greater veridical recall of propositions 30 min later in young adults than in healthy older adults. This effect remained significant when the number of veridical responses at immediate recall was controlled through an ANCOVA.[3] With respect to dementia severity, veridical recall of propositions was significantly less in both demented groups than in the healthy older adults after the delay interval. The effect of dementia severity was not statistically significant ($p > .05$) when immediate veridical recall was controlled through an ANCOVA.[4]

Age (young vs. healthy old) did not influence either gist or distortion errors at delayed recall. This lack of an age effect persisted when errors at immediate recall were controlled through an ANCOVA. Effects of dementia severity, however, were observed for both types of errors at delayed recall. As dementia severity increased, the numbers of gist and distorted responses decreased (CDR 0 > 0.5 > 1), although the difference between the very mildly and mildly demented groups was not statistically significant for distortions (1.26 vs. 0.90). This same pattern of results was observed when the number of gist and the number of distortion errors at immediate recall were controlled through an ANCOVA.[5]

As shown in Table 5, gist and distortion errors at delayed recall were divided into those made for case relations (central concepts) and those made for arguments (details). There were significant differences among the three older adult groups on all four measures (*p*s < .01), even after statistically controlling for the number of corresponding errors made at immediate recall (*p*s < .01). In all cases, the healthy older group made more such errors than did either of the two demented groups; the very mildly and mildly demented groups were significantly different from each other only in terms of the number of gist errors made for

[3] This result should be interpreted cautiously. The parallelism assumption was not met. The correlation between immediate and delayed veridical recall was .84 in the young adults and .70 in the healthy older adults.

[4] Again, the parallelism assumption was not met. The correlation between immediate and delayed veridical recall was, as already reported, .70 in the healthy older adults. It was .66 and .55 in the very mildly and mildly demented groups, respectively.

[5] The parallelism assumption was not met in the ANCOVA of gist responses in the three older adult groups. The correlations between immediate and delayed gist responses were .43, .64, and .32 for the healthy, very mildly, and mildly demented groups, respectively. The reduced correlation in the healthy group probably reflects increased gist responses at delayed recall by some participants; note that the mean number of gist responses for this group was higher at delayed recall ($M = 7.04$) than it was at immediate recall ($M = 5.68$). The reduced correlation in the mildly demented group probably reflects a floor effect at delayed recall ($M = 0.62$).

arguments (1.90 vs. 0.56). There appeared to be a floor effect in the mildly demented group for both types of gist errors as well as for distortions of case relations.

Unit Analyses

We conducted three sets of unit analyses (cf. D. C. Rubin, 1978, 1986) to examine the influence of story properties beyond case relations and arguments on the deficits in prose recall related to dementia severity. A unit analysis is a correlational analysis wherein the unit of measurement is the proposition, summing recall across subjects. This method of analysis allowed us to ask questions about how the story properties affected veridical propositional recall. Three different methods were used to define weighting algorithms for propositions within each story. Each weighting method codes for a specific story property (e.g., serial position or word frequency). These weights were used to correlate the hypothesized probability of propositional recall of a given proposition with its observed frequency of veridical recall within each group of older adults: nondemented, very mildly demented, and mildly demented. To increase the sample size (i.e., number of propositions) for these correlational analyses, we combined data across Stories A and B.

To facilitate the application of computer software for the three weighting systems, David K. Johnson and a consultant collapsed six propositions in Story A and two propositions in Story B to create multiple word propositions. By result, the current analyses used a slightly smaller subset of propositions than the original (29 out of the original 35 for Story A and 31 out of the original 33 for Story B). For example in Story A, "Anna Thompson" was originally scored as two distinct propositions; in the unit analyses, it was combined into a single unit that required veridical recall of both elements. Veridical data from the original propositional decomposition were collapsed so that both of the original propositions had to be achieved in full to receive veridical credit in the second propositional decomposition.

The first unit analysis used a set of unit weights derived from the construction integration model (Kintsch, 1978, 1988, 1998). On the basis of a theory of spreading activation within a propositional network, the model creates hypothetical associations between propositions and predicts the probability of recall for each proposition in relation to all other propositions in the story. These probabilities are called *strength values.* They describe a set of logical relationships within the story on the basis of the repetition of a proposition throughout a story. That is, a proposition may be either a discrete recall unit or used as an argument of another proposition. According to the model, the latter increases the likelihood that the proposition will be recalled. As shown in Table 6, the correlations between the strength values of the propositions and their observed recall were relatively strong (.38 to .45) for the young, nondemented, very mildly demented, and mildly demented groups, respectively.

We also used the word frequency norms of Kučera and Francis (1967) to determine whether this well-established list-learning variable might be related to observed prose recall. To be conservative, we assigned a proposition containing multiple words the value of the word in the propositional set with the lowest frequency. As also shown in Table 6, proposition frequencies of occurrence in the English language were not strongly correlated with observed recall (*r*s ranged from −.12 to −.28). The negative values indicate that, as is usually found, less frequent words were more likely to be recalled.

Table 6
Correlations Between Story Propositions and Frequency of Propositional Recall

Group	Construction integration model[a]	Serial position	Word frequency[b]
Young	.43**	.13	−.12
Nondemented	.45**	.15	−.23
Very mildly demented	.42**	.29*	−.28
Mildly demented	.38**	.22	−.24

[a] Kintsch (1988). [b] Kučera and Francis (1967).
* $p < .05$. ** $p < .01$.

Also as shown in Table 6, we examined the effect of serial position, an important determiner of recall in list learning. The serial positions of the propositions in each story were converted so each could be linearly related to observed frequency of recall. The story's middle most proposition was given a value of 1, with propositions preceding it ranked in order (i.e., 2, 3, 4 . . .) ending with $(n + 1)/2$ for the first proposition in the story. Similarly, the propositions following the middle most proposition were ranked in order (i.e., 2, 3, 4 . . .) ending with $(n + 1)/2$ for the last proposition in the story. None of the correlations between serial position and observed recall were impressive (*r*s ranged from .13 to .29). The one correlation (.29) that was significantly different from 0 was primarily due to better recall by the very mildly demented group of the last two propositions in Story B.

Discriminant Analyses

Finally, a discriminant function analysis was conducted to determine if the revised scoring procedure would improve the ability of the subtest to differentiate between healthy aging and very mild Alzheimer's disease (CDR 0 vs. 0.5). Veridical recall of propositions correctly classified 86% of the healthy and very mildly demented older adults (Table 7) with 84% sensitivity and 87% specificity. This new scoring procedure increased correct classification of very mildly dementing from healthy older adults slightly (4%) over the Russell scoring procedure used by Storandt and Hill (1989). Although sensitivity was the same for both projects (84%), specificity increased by 6%.

The inclusion of gist recall did not improve classification. Overall correct classification using combined veridical and gist recall was 85%, with 77% sensitivity and 89% specificity. Thus, inclusion of gist recall actually makes it more difficult to identify very mildly demented people. To further examine the role that errors may play in the classification of individuals with dementia, we rescored this older adult

Table 7
Classification Analysis for Veridical Recall

	Predicted			
	No dementia		Very mild dementia	
Actual	*n*	%	*n*	%
No dementia	41	87	6	13
Very mild dementia	5	16	26	84

Note. Eighty-six percent of the total sample was correctly classified.

sample using the original Wechsler Memory Scale (1945) idea units, each scored along the dimension of correctness as prescribed in this article. The inclusion of gist recall in the discrimination of healthy and very mildly dementing older adults resulted in a slight drop in correct classifications (80% versus 82% using veridical recall only).

The discriminant function analysis on the basis of veridical recall was replicated using 98 healthy and 82 very mildly demented individuals from Chapman et al.'s (1997) Sample 2, which was recruited at the same center between 1991 and 1995. The sample size is slightly reduced because transcripts of the Logical Memory responses were unavailable for some people, and therefore, the subtest could not be rescored according to the propositional decomposition criteria. A 79% correct classification rate was achieved in this replication (77% sensitivity, 81% specificity). Although slightly less than the values obtained for the sample recruited from 1984 to 1991, the results using the propositional decomposition scoring were very similar to those obtained when the Logical Memory stories were scored according to the Russell procedure (78% overall correct classification, 75% sensitivity, and 81% specificity).

Discussion

Effects of Age

Contrary to expectation, there was no age-related deficit in immediate prose recall. It was not present in either the veridical or gist responses. The one age effect that did occur at immediate recall involved slightly more distortions of the details (arguments) of the propositions by the younger adults. The lack of age differences observed may depend, in part, on the type of healthy older adults who were included in the sample; they were carefully screened for any sign of cognitive decline. Previous research may have used samples of older adults that included people who were in the very earliest stage of dementia (J. C. Morris, 1999; J. C. Morris et al., 1996).

What about delayed memory? Here, the expected (e.g., Albert, 1988) age-related difference in veridical recall was obtained. Although healthy older adults were as able as younger adults to retain veridical representations in immediate recall, those representations deteriorated more in older people during 30 min of interpolated activities than they did in younger adults. (There were no age effects in gist recall or distortions at delayed recall.) This pattern of data is quite intriguing with respect to age-related changes in acquisition and retention, which has been a relatively difficult issue to address in past aging studies. As Kausler (1994) noted, one difficulty in this area of research is that younger and older adults are not typically equivalent on immediate recall, and so the retention function is difficult to establish. Moreover, by equating young and older adults on initial acquisition by providing older adults more study time, one does not take into account the effect of additional learning during the acquisition test (see Crowder, 1976, for a discussion of this issue). The present study provides some intriguing data in this regard. Specifically, the statistically equivalent performance in the younger and older adults in immediate veridical recall (suggesting that the information was clearly encoded and could be retrieved) and the age-related deficit in delayed surface-level recall provide support for the idea that there is a breakdown in retention processes with age. Of course, we are not suggesting that this is the only locus of age-related change in memory performance, but rather we are noting that the present fortuitous pattern of data provides some support for a role of age-related changes in retention.

Effect of Dementia

In contrast to what was seen in normal aging, the effect of dementia was primarily one of dampening of immediate prose recall with little additional effect after the delay interval. Indeed, many of the demented individuals performed at floor at delayed recall. This result was consistent with the analyses of these data reported previously using the Russell scoring procedure (e.g., Chapman et al., 1997; Robinson-Whelen & Storandt, 1992; Storandt & Hill, 1989) and also by others investigating logical memory performance in individuals with dementia (e.g., G. Cohen, 1979; Spilich, 1983).

The effect of dementia on immediate prose recall was most dramatic for veridical recall of the propositions. Even though veridical recall, which presumably provides support for gist recall, was severely attenuated, demented individuals were still able to extract some gist-related information (i.e., mean gist recall was above floor), but gist recall did not increase as a function of dementia severity nor did demented individuals show increased numbers of distortions. The relatively low number of overall gist responses also probably contributed to our failure to detect a differential effect of gist errors in the details of the propositions as compared with their central elements. In sum, the errors in prose recall made by individuals with dementia tended to be errors of omission rather than commission.

The dramatic decline in veridical recall with increasing dementia severity combined with minimal change in the level of gist or distorted recall produced a decrease in total production. Therefore, examination of the relative production of the three types of responses would lead one to conclude, for example, that one of the effects of dementia is increased distortions. Such a conclusion would be misleading unless one emphasized its relative nature and the reason (i.e., the decrease in veridical production).

Clinical Implications

The veridical scoring procedure for the propositional decompositions of the Logical Memory stories provides a useful method of differentiating nondemented and very mildly demented older adults. This project has not previously been able to achieve quite as high correct classification of very mildly demented people with a single brief psychometric subtest. Although this scoring method needs to be studied in a clinical setting with a more heterogeneous sample, this study indicates that prose recall may be an extremely sensitive method to detect the presence of dementia in its very earliest stages. Although the two methods produce comparable results, the propositional decomposition used in the present study has a stronger theoretical basis.

Russell (1975), Abikoff et al. (1987), and the current study have shown that inclusion of gist-based recall in the scoring of prose memory recall actually decreases the clinician's ability to detect dementia. Allowing gist-based recall in the 1997 version of the Logical Memory subtest dilutes the subtest's sensitivity to decrements associated with dementia. Clinicians may wish to modify their scoring practices for this test.

The veridical scoring procedure was reported by the raters to be easy to learn and apply. Indeed, training was rapid and interrater reliability was high. The stories are simple to administer and well accepted by patients. Clearly, the most difficult part of the procedure used in this project was the initial propositional decomposition of the two story texts that were read to participants; it required consultation with an expert in propositional analysis. Scoring of individual participant responses, however, merely required comparison of the transcribed participant recall to the standard developed by the expert in the same fashion that examiners refer to scoring manuals for a wide variety of procedures. Subsequent routine implementation of this scoring procedure in the Washington University Alzheimer's Disease Research Center has been quite successful.

Results from the unit analyses indicate that recall across age and dementia severity groups tended to favor thematically important propositions more than a proposition's serial position or associated word frequency. These data indicate that the application of standard list-learning techniques to analyze prose recall data may not capture relevant clinical information specific to prose recall.

Theoretical Implications

What goes awry in immediate prose recall in dementia? Although the idea that there are deficits in the control aspects of working memory in DAT is not new (e.g., Collette, Van der Linden, & Salmon, 1999; R. G. Morris, 1994), it is more often studied in span, sentence comprehension, or divided attention tasks (e.g., Grober & Bang, 1995; Spieler, Balota, & Faust, 1996). Baddeley (2000) recently added an episodic buffer to his model of working memory because he found it difficult to explain prose recall in terms of the central executive and its two slave systems (the phonological loop and the visuospatial sketch pad) of his original model (Baddeley & Hitch, 1974). The episodic buffer is "a limited-capacity temporary storage system that is capable of integrating information from a variety of sources" (Baddeley, 2000, p. 421), including long-term memory. It has a larger capacity (15 to 20 units) than does the phonological loop (approximately 7 units). The episodic buffer can store these units temporarily. The central executive, which controls the episodic buffer, does not have storage capacity; its primary function is attentional control. We propose that the memory deficit seen in immediate prose recall in the early stages of DAT occurs in the interplay of the central executive processes and the episodic buffer.

Of course, it is possible that DAT produces only a capacity limitation in the episodic buffer or somehow interferes with its integration function. Neither of these mechanisms seems to be the major difficulty, however. Recall that Kintsch's (1998) construction integration model was equally effective across the four groups examined in this study. The more central a proposition was to the prose passage, the more likely it was to be recalled. This suggests relatively intact comprehension, retention, and integration of the relations among the elements of the story in individuals in the early stages of dementia. In the context of these brief stories, some of the elements that were integrated were at the beginning of the story, some in the middle, and some at the end. Therefore, it does not appear that the episodic buffer's capacity is reduced in the early stages of dementia; it does not seem to "fill up" and become unable to accept more incoming information. For example, there was no evidence of a primacy effect. The very modest correlation between serial position of a proposition and probability of recall seen in the very mildly demented group was due to slightly better recall of the last two propositions in the second story—a recency effect.

Instead, we hypothesize that information fades from the episodic buffer more rapidly in DAT because the central executive may be forced to engage in other attention-demanding control tasks such as verbally outputting the integrated information (see Baddeley et al., 1986, for a similar argument). Recall that Baddeley (2000) proposed that the episodic buffer is a *temporary* store "accessed by the central executive through the medium of conscious awareness" (p. 421). The central executive influences the content of the episodic store by attending to a given source of information. Although clearance of the episodic buffer is not specifically discussed, if information can be placed in this temporary store by focusing attention on it, it may be cleared from the store by removing attention from it. Therefore, the difficulty seen in immediate prose recall in the early stages of DAT may not be a memory problem at all. It may represent a deficit in the control of attention (see Balota & Faust, 2002).

This interpretation is consistent with recent evidence accumulating from a number of distinct but related paradigms. First, consider the integrity of preexisting structures in early stage DAT and the availability of such structures for the integration processes. The evidence from semantic priming studies supports the notion that the automatic activation of related information is relatively intact in early

stage DAT (see Balota et al., 1999; Nebes, 1989; Ober & Shenaut, 1995). There is also evidence from the Deese–Roediger–McDermott paradigm (after Deese, 1959; Roediger & McDermott, 1995) that converges on this notion. In this paradigm, the person sees or hears a list of words that are all related to a critical word that is not on the list. A substantial number of people will report, falsely, that the critical word was on the list. There is considerable stability in this false recall in both healthy aging and early stage dementia, although memory for the words actually on the list decreases. Recall of the critical nonpresented word presumably reflects the extraction of the relations among the related words (e.g., Balota et al., 1999; Budson, Daffner, Desikan, & Schacter, 2000; Norman & Schacter, 1997; Watson, Balota, & Sergent-Marshall, 2001). This pattern parallels the present results in which surface memory decreases precipitously in DAT but the deficit in gist-related information is less dramatic.

References

Abikoff, H., Alvir, J., Hong, G., Sukoff, R., Orazio, J., Solomon, S., & Saravay, S. (1987). Logical Memory subtest of the Wechsler Memory Scale: Age and education norms and alternate-form reliability of two scoring systems. *Journal of Clinical and Experimental Neuropsychology, 9,* 435–448.

Albert, M. S. (1988). General issues in geriatric neuropsychology. In M. S. Albert & M. B. Moss (Eds.), *Geriatric neuropsychology* (pp. 3–10). New York: Guilford Press.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Armitage, S. G. (1946). An analysis of certain psychological tests used for the evaluation of brain injury. *Psychological Monographs, 60*(1, Serial No. 277).

Baddeley, A. D. (2000). The episodic buffer: A new concept of working memory? *Trends in Cognitive Science, 4,* 417–423.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

Baddeley, A. D., Logie, R., Bressi, S., Della Sala, S., & Spinnler, H. (1986). Dementia and working memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 38*(A), 603–618.

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

Balota, D. A., & Faust, M. E. (2002). Attentional control in dementia of the Alzheimer's type. In F. Boller & S. Cappa (Eds.), *Handbook of neuropsychology* (Vol. 6, pp. 51–80). Amsterdam: Elsevier Science.

Berg, L., Hughes, C. P., Danziger, W. L., Martin, R. L., & Knesevich, J. (1982). Mild senile dementia of Alzheimer type (SDAT): Research diagnostic criteria, recruitment, and description of a study population. *Journal of Neurology, Neurosurgery, and Psychiatry, 45,* 962–968.

Berg, L., McKeel, D. W., Jr., Miller, J. P., Storandt, M., Rubin, E. H., Morris, J. C., et al. (1998). Clinicopathologic studies in cognitively healthy aging and Alzheimer disease. *Archives of Neurology, 55,* 326–335.

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

Burke, W. J., Miller, J. P., Rubin, E. H., Morris, J. C., Coben, L. A., Duchek, J., et al. (1988). Reliability of the Washington University Clinical Dementia Rating (CDR). *Archives of Neurology, 45,* 31–32.

Chapman, L. L., White, D. A., & Storandt, M. (1997). Prose recall in dementia: A comparison of delay intervals. *Archives of Neurology, 54,* 1501–1504.

Cohen, G. (1979). Language comprehension in old age. *Cognitive Psychology, 11,* 412–429.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed., pp. 73–76). Hillsdale, NJ: Erlbaum.

Collette, F., Van der Linden, M., & Salmon, E. (1999). Executive dysfunction in Alzheimer's disease. *Cortex, 35,* 57–72.

Crosson, B., Hughes, C. W., Roth, D. L., & Monkowski, P. G. (1984). Review of Russell's (1975) norms for the Logical Memory and Visual Reproduction subtests of the Wechsler Memory Scale. *Journal of Consulting and Clinical Psychology, 52,* 635–641.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

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

Ellis, D. G. (1996). Coherence patterns in Alzheimer's discourse. *Communication Research, 23,* 472–495.

Fletcher, C. R. (1994). Levels of representation in memory discourse. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 589–607). San Diego, CA: Academic Press.

Grober, E., & Bang, S. (1995). Sentence comprehension in Alzheimer's disease. *Developmental Neuropsychology, 11,* 95–107.

Haaland, K. Y., Linn, R. T., Hunt, W., & Goodwin, J. S. (1983). A normative study of Russell's variant of the Wechsler Memory Scale in a healthy elderly population. *Journal of Consulting and Clinical Psychology, 51,* 878–881.

Haut, M. W., Demarest, D., Keefover, R. W., & Rankin, E. D. (1994). Semantic sensitivity for prose in patients with probable Alzheimer's disease. *Aging and Cognition, 1,* 238–246.

Hultsch, D. F., Hertzog, C., Dixon, R. A., & Small, B. J. (1998). *Memory change in the aged.* New York: Cambridge University Press.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, MA: Harvard University Press.

Kausler, D. H. (1994). *Learning and memory in normal aging.* San Diego, CA: Academic Press.

Kemper, S., LaBarge, E., Ferraro, R. F., Cheung, H., Cheung, H., & Storandt, M. (1993). On the preservation of syntax in Alzheimer's disease. *Archives of Neurology, 50,* 81–86.

Kintsch, W. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363–394.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95,* 163–182.

Kintsch, W. (1994). Text comprehension, memory and learning. *American Psychologist, 49,* 294–303.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* New York: Cambridge University Press.

Kintsch, W., Kozminsky, E., Streby, W. J., McKoon, G., &

Keenan, J. M. (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior, 14,* 196–214.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363–394.

Kintsch, W., & Vipond, D. (1979). Reading comprehension and readability in educational practice and psychological theory. In L. G. Nilson (Ed.), *Perspectives of memory research* (pp. 325–366). Hillsdale, NJ: Erlbaum.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Larrabee, G. J. (1987). Further cautions in interpretation of comparisons between the WAIS–R and the Wechsler Memory Scale. *Journal of Clinical and Experimental Neuropsychology, 9,* 456–460.

Lyons, K., Kemper, S., LaBarge, E., Ferraro, R. F., Balota, D., & Storandt, M. (1994). Oral language and Alzheimer's disease: A reduction in syntactic complexity. *Aging and Cognition, 1,* 271–281.

McCulla, M. M., Coats, M., Van Fleet, N., Duchek, J., Grant, E., & Morris, J. C. (1989). Reliability of nurse specialists in staging of dementia. *Archives of Neurology, 46,* 1210–1211.

Moenster, P. A. (1972). Learning and memory with relation to age. *Journal of Gerontology, 27,* 361–363.

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology, 43,* 2412–2414.

Morris, J. C. (1999). Is Alzheimer's disease inevitable with age? *Journal of Clinical Investigation, 104,* 1171–1173.

Morris, J. C., Storandt, M., McKeel, D. W., 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.

Morris, R. G. (1994). Working memory in Alzheimer-type dementia. *Neuropsychology, 8,* 544–554.

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

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

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

Powers, D. G., Logue, P. E., McCarty, S. M., Rosenstiel, A. K., & Ziesat, H. A. (1979). Inter-rater reliability of the Russell version of the Wechsler Memory Scale: An attempt to clarify ambiguities in scoring. *Journal of Clinical Neuropsychology, 1,* 343–345.

Robinson-Whelen, S., & Storandt, M. (1992). Immediate and delayed prose recall among normal and demented adults. *Archives of Neurology, 49,* 32–34.

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

Rubin, D. C. (1978). A unit analysis of prose. *Journal of Verbal Learning and Verbal Learning Behavior, 17,* 599–620.

Rubin, D. C. (1986). Unit analysis of prose memory in clinical and elderly populations. *Developmental Neuropsychology, 2,* 77–87.

Rubin, E. H., Storandt, M., Miller, P., Kinscherf, D. A., Grant, E. A., Morris, J. C., & Berg, L. (1998). A prospective study of cognitive function and onset of dementia in cognitively healthy elders. *Archives of Neurology, 55,* 395–401.

Russell, E. W. (1975). A multiple scoring method for the assessment of complex memory functions. *Journal of Consulting and Clinical Psychology, 43,* 800–809.

Schear, J. M. (1986). Utility of half-credit scoring of Russell's revision of the Wechsler Memory Scale. *Journal of Clinical Psychology, 42,* 783–787.

Small, B. J., Dixon, R. A., Hultsch, D. F., & Hertzog, C. (1999). Longitudinal changes in quantitative and qualitative indicators of word and story recall in young-old and old-old adults. *Journal of Gerontology: Psychological Sciences, 54B,* P107–P115.

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

Spilich, G. (1983). Life-span components of text processing: Structural and procedural differences. *Journal of Verbal Learning and Verbal Behavior, 22,* 231–244.

Stine, E. A. L., & Wingfield, A. (1988). Memorability as an indicator of qualitative age differences in text recall. *Psychology and Aging, 3,* 179–183.

Storandt, M., Botwinick, J., Danziger, W., Berg, L., & Hughes, C. (1984). Psychometric differentiation of mild senile dementia of the Alzheimer's type. *Archives of Neurology, 41,* 497–499.

Storandt, M., & Hill, R. D. (1989). Very mild senile dementia of the Alzheimer's type: II. Psychometric test performance. *Archives of Neurology, 46,* 383–386.

Sweet, J. J., & Kolden, G. G. (1986). Effectiveness of Wechsler Memory Scale scoring guidelines for improving scorer reliability: A followup study. *International Journal of Clinical Neuropsychology, 8,* 38–40.

Taub, H. A. (1979). Comprehension and memory of prose by young and old adults. *Experimental Aging Research, 5,* 3–13.

Turner, A., & Greene, E. (1977). *The construction and use of a propositional text base* (Tech. Rep. No. 63). Boulder: University of Colorado, Institute for the Study of Intellectual Behavior.

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

Wechsler, D. (1945). A standardized memory scale for clinical use. *Journal of Psychology, 19,* 87–95.

Wechsler, D. (1955). *Manual: Wechsler Adult Intelligence Scale.* New York: Psychological Corporation.

Wechsler, D. (1997). *Wechsler Memory Scale III: Administration and scoring manual.* Orlando, FL: Psychological Corporation.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123,* 162–185.

Received October 25, 2001
Revision received May 13, 2002
Accepted June 29, 2002 ■