DEVELOPMENTAL NEUROPSYCHOLOGY, 1986, *2*(2), 77–87

Unit Analysis of Prose Memory in Clinical and Elderly Populations

Kathryn A. Schultz
Durham VA Medical Center

Frederick A. Schmitt
Sanders-Brown Research Center on Aging, University of Kentucky Medical Center, and Lexington VA Medical Center

Patrick E. Logue
Duke University Medical Center

David C. Rubin
Duke University

Interpretation of clinical memory tests generally emphasizes the quantitative aspects of recall. This study presents an additional unit analysis of the Logical Memory subtest of Russell's revision of the Wechsler Memory Scale for a variety of older adult groups. Patients' neuropsychological test data were reviewed, and the paragraphs from the Logical Memory subtest were analyzed using unit analysis (Rubin, 1978). The older adults consisted of a healthy group as well as groups whose diagnoses included Alzheimer's and multi-infarct dementias, head trauma, and metabolic and affective disorders. Quantitative analyses of recall revealed group differences. Qualitative analysis of which memory units were recalled, however, showed similarities in memory processing among these groups.

Memory impairment is part of the *DSM-III* (American Psychiatric Association, 1980) definition of dementia and appears to be the first and most disabling of the cognitive deficits associated with organic brain syndromes. As a result, much of the research to date has focused on episodic memory

---

Requests for reprints should be sent to Frederick A. Schmitt, Department of Neurology (MS-129), University of Kentucky Medical Center, Lexington, KY 40536-0084, or to David C. Rubin, Department of Psychology, Duke University, Durham, NC 27706.

abilities in the diagnosis of dementia (e.g., Rosen, 1983). A number of studies have begun to characterize the nature of this episodic memory dysfunction (e.g., Corkin, 1982; Miller, 1981; Weingartner, Graffman, Boutelle, Kaye, & Martin, 1983; Weingartner et al., 1981). Some investigators have studied the episodic memory differences in dementia; others have begun to look at the effects of dementia on semantic memory. Studies of language function in dementia suggest that some aspects of semantic memory are impaired (Bayles, 1982; Clark, 1980). Other data, however, indicate that semantic memory may remain intact for patients with Alzheimer's disease (AD; Nebes, Martin, & Horn, 1984) and Korsakoff's disease (Kinsbourne & Wood, 1982).

Although the memory deficit in dementia appears to be clear, many of the tasks used to investigate memory in demented patients require a great deal of attentional capacity. Because attentional processes might be affected by the dementing process, the performance demands of a particular memory task may overestimate the severity of the memory disturbance. Much of what is known about the memory changes associated with normal and pathological aging is based on tasks using traditional comparisons between different diagnostic groups. For example, the Wechsler Memory Scale (WMS) and Russell's revision of the WMS (RWMS) are often used as an index of memory impairment in neuropsychological test batteries (e.g., Rosen, 1983). Further, as demonstrated by Logue and Wyrick (1979), differences in how many memory units are recalled on the RWMS is often indicative of memory dysfunctions associated with dementia. Although the standard approach to memory assessment allows us to demonstrate differences in the amount of material that is remembered, the differences in that total amount remembered for a given memory task may be influenced by factors such as attention, motivation, and affect. Memory decline may also reflect a general disruption of "fluid" intellectual abilities secondary to organic processes even when more static measures such as IQ scores are relatively intact.

Although examining how much information is recalled clearly allows for statements regarding memory dysfunction, our knowledge of what actually occurs in terms of memory processes associated with various clinical conditions is limited. If we can make statements regarding memory processing (e.g., rehearsal or use of semantic structure), we might not only better describe the memory changes associated with pathological conditions, but we might also improve the diagnostic accuracy of our clinical measures. One alternative method of analysis is to look quantitatively at which units (as opposed to how many units) of the to-be-remembered information are recalled (e.g., Rubin, 1977). Rubin, Olson, Richter, and Butters (1981) demonstrated that memory for prose in both Korsakoff and schizophrenic populations appears less severely impaired when what is recalled rather than how much is recalled was assessed. In this study both schizophrenic and Korsakoff patients recalled fewer memory units but recalled essentially the same items from the

Logical Memory subtest of the WMS. Therefore, even though the schizophrenics were clearly experiencing a thought disorder, and the Korsakoff patients suffered from amnesia, both groups tended to process the simple prose passages in ways similar to those used by the control group. These data then indicate regularities in the processing of memory material, which may be determined by the structure inherent in the memory information at both semantic and syntactic levels.

Analyses of the syntactic and semantic structure in certain prose recall tasks might therefore shed additional light on the types of memory failures seen in older clinical populations. For example, Weingartner et al. (1983) argued that semantic memory was intact and accessible in amnesic patients as indexed by their normal performance on tasks such as category member generation. By contrast, demented patients were unable to access and utilize semantic information. Both groups did, however, show severe episodic memory dysfunction. Data such as these support the notion of a semantic deficit in certain older adult clinical populations. Studies of language in dementia (Bayles, 1982) have suggested, however, that syntactic and phonological use remains fairly intact. Therefore, although there appear to be marked decrements and perhaps wide differences between clinical populations in the nature of memory performance, episodic and semantic memory systems may be differentially disrupted by different clinical entities.

If we look at which specific "bits" of information are remembered rather than what amount is remembered, an analysis of episodic and semantic structure in memory is possible. Further, analyzing what is remembered, as well as how much is remembered, allows for inferences about memory processing and elimination of various extraneous subject variables such as mood and attention (Rubin, 1978; Rubin et al., 1981). Using a unit analysis of prose recall, we can generate information on what may actually occur in the memory processes associated with various clinical conditions.

Because the WMS is one of the most widely used memory tasks for estimating memory impairment, the present study expands on the traditional use of the WMS and the RWMS by providing an analysis of both the qualitative (what is remembered) and quantitative (how much is remembered) aspects of prose recall. Although the WMS has been criticized (e.g., Erickson & Scott, 1977), it is still a useful tool for the assessment of immediate and long-term recall (RWMS) of verbal and visual materials.

The current study reports a unit analysis for a wide variety of older adult populations similar to the analysis reported by Rubin et al. (1981). Of interest were normal older adults as well as older adults whose diagnoses included probable AD; closed head injury (CHI), and metabolic, affective, and vascular disorders. An analysis of RWMS prose recall among these groups not only provided statements regarding how much was remembered by each group of individuals but also provided an analysis of how these groups recalled and what they were recalling. Therefore, the traditional comparisons of recall performance between these elderly groups were available for analysis (e.g., Logue & Wyrick, 1979). More important, the differences in which units were recalled by each group provided data regarding semantic memory functioning as well.

METHOD

Subjects

Six groups of older adults took part in the investigation: 69 clinically diagnosed AD patients (*DSM-III* diagnosis), 13 multi-infarct dementia (MID) cases, 14 CHI cases, 15 metabolic disorder cases, 14 affective disorder cases, and 37 older adult normals (a control group). The characteristics of these groups are presented in Table 1. Test results for each person were taken from their neuropsychological test batteries obtained from the files maintained by the Psychodiagnostic Laboratory at Duke University Medical Center. The control group consisted of older adults who had been given the RWMS as part of other experimental studies at the Duke University Center for the Study of Aging and Human Development.

Materials

The Logical Memory subtest of Form 1 of the WMS (Wechsler, 1945) was used. This form of the WMS had not been administered previously to any of the subjects. The subtest consisted of two passages of approximately 50 words each. The WMS stories originally were presented using the RWMS (Russell, 1975) and included an update scoring scheme providing partial credit for approximations of verbatim recall as derived by Power, Logue, McCarty, Rosenstiel, and Ziesat (1979) and Prigatano (1978).

TABLE 1
Group Characteristics

	Age		*Education*		*Sex*	
Group	*M*	*SD*	*M*	*SD*	*Male*	*Female*
AD	62.81	6.08	12.81	3.86	31	38
MID	58.77	11.25	10.92	1.98	7	6
CHI	40.29	19.29	12.36	2.21	10	4
Metabolic	51.43	12.79	12.67	3.17	5	10
Affective	50.29	7.77	13.46	2.33	4	10
Control	66.51	11.47	11.73	3.67	20	17

Procedure

All subjects were tested individually as part of a clinical neuropsychological assessment, or, in the case of the normal controls, as part of studies of memory and aging. The normal group consisted of volunteers from the Duke University Center for the Study of Aging and Human Development subject pool. The clinical groups were patients tested in the Psychodiagnostic Laboratory at Duke University Medical Center. In general, the clinical patients were given a full RWMS following the procedures described by Russell (1975), including a delayed 30-min recall trial for both the Logical Memory and Visual Reproduction subtests. The clinical groups were diagnosed medically, psychiatrically, and psychologically using appropriate diagnostic procedures.

The Logical Memory and Visual Reproduction subtests of the RWMS were administered to all subjects. Patients who were involved in clinical evaluations generally completed the Wechsler Adult Intelligence Scale (WAIS) before the RWMS was administered. The 30-min interval for delayed recall was then usually used to complete the Halstead-Reitan Trail Making, Finger Tapping, and Sensory Screening tests. Normal older adults generally completed the RWMS with the delay interval being filled with WAIS subtests (e.g., Vocabulary, Digit Symbol), brief mental status evaluations, or questionnaires.

Scoring of the RWMS Logical Memory subtest followed the original procedures of Russell (1975) with the following exceptions: Raw scores on the Logical Memory subtest were obtained using Prigatano's (1978) method providing partial credit for approximations of verbatim recall. To simplify the analysis of the raw data, a conversion factor (multiplying by two) was applied separately to each individual item in each story. This had the overall effect of doubling the total recall score. Following Russell's (1975) procedure, raw scores for the stories were not divided by two.

RESULTS

Analysis of the Amount Recalled

The mean recall score per unit is displayed in Table 2. As expected, the various subject populations recalled different amounts from the stories, $F(5, 151) = 34.59$, $p < .001$, and the immediate retention interval produced greater recall than the delayed retention interval (.499 vs. .346), $F(1, 151) = 147.10$, $p < .001$. In addition, the first story of the RWMS Logical Memory subtest produced a higher level recall than did the second (.474 vs. .371), $F(1, 151) = 51.41$, $p < .001$. In the administration of the RWMS, the stories are

TABLE 2
Mean Recall Score Per Unit

Group	*n*	*Story*	*Immediate (I)*	*Delayed (D)*	*I − D*	*(I − D) / I*
AD	64	1	.222	.094	.128	.577
		2	.200	.082	.118	.590
MID	13	1	.596	.330	.266	.446
		2	.388	.206	.182	.469
CHI	14	1	.643	.449	.194	.302
		2	.549	.412	.137	.250
Metabolic	15	1	.683	.503	.180	.264
		2	.527	.391	.136	.258
Affective	14	1	.735	.550	.185	.252
		2	.513	.383	.130	.253
Control	37	1	.976	.795	.181	.185
		2	.771	.608	.163	.211

Note. Maximum score = 2.00.

always presented in the same order, confounding the effects of primacy and differences in the stories themselves.

Two interactions were significant. Story 1 lost more over the retention interval than did Story 2 (a decrease of .168 vs. .138), $F(1, 151) = 4.84$, $p < .05$. This interaction may be due to the higher initial level of recall of Story 1. There was also a Group × Story interaction, $F(5, 151) = 5.64$, $p < .001$, which again is hard to interpret because of the confounding of story order and story content. There were no significant interactions of Group × Retention Interval, $F(5, 151) = 1.33$, $p = .25$, or Group × Retention Interval × Story, $F(5, 151) = 0.41$, $p = .84$. The two interactions involving group and retention interval may be misleading, however, because of the very different levels of initial recall. All groups dropped about .15 between immediate and delayed recall. For the control subjects, this was a drop from .874 to .702, or 20%; for the AD group, the change was from .211 to .088, or 58%. The last two columns of Table 2 provide two ways of looking at the drop in recall level with retention interval. The first of these columns indicates the absolute drop used in the analysis of variance (ANOVA). The second of these columns indicates the percentage drop. For the six groups studied here, the lower the initial level of recall of a group, the greater the percentage drop over retention interval.

Analysis of Which Units Are Recalled

The previous analysis examined the amount of recall. The raw data for that analysis were the scores subjects received on each unit of two RWMS stories.

The data were collapsed by summing over units to provide a value for each subject for each of the two stories. These two values for each subject were used to calculate Table 2 and the resulting ANOVA. In the analysis of which units are recalled, the identical raw data were used, but they are collapsed differently. Instead of summing over units to provide a value for each subject for each story, the sum is made over all the subjects in a particular group to provide a value for each unit in each story. This allows us to answer the question of whether different groups recall or fail to recall the same units. Because correlations are used, the absolute level of recall does not influence the results unless it is so low that it produces floor effects. In interpreting the correlations obtained, it is assumed that differential cognitive deficits among groups would result in different units being recalled. Thus, for example, if one clinical group failed to understand a story and its main points, or if another group were subject to marked interference, these groups would recall units different than those that would be called by the control subjects.

Table 3 presents the reliabilities of the sets of recall scores for the units of the stories. The values of Cronbach's alpha (Cronbach, 1951) listed in the table can be interpreted as the correlation obtained between the actual data and data from an equal number of subjects selected from the same population. The more subjects in a group, the higher the reliability should be, other things being equal (Walker & Lev, 1953).

Table 4 presents the correlations calculated from the mean recall scores for each unit of the six groups of subjects. Although some variation among the correlations exists, the correlations are all quite high when compared to those of Table 3. In fact, the simple average of all Table 4 correlations corrected for attenuation is .987. Thus, it appears that the different groups tended to recall or fail to recall the same units.

Table 5 presents the correlations between the data from the two retention intervals. Again, there appears to be little effect on which units are recalled at different times, though there are significant differences in how many units are recalled. The simple average of all Table 5 correlations corrected for attenuation is .868.

TABLE 3
Reliabilities as Measured by Cronbach's Alpha

	Story 1		*Story 2*	
Group	*Immediate*	*Delayed*	*Immediate*	*Delayed*
AD	.893	.743	.874	.753
MID	.642	.635	.831	.815
CHI	.828	.757	.698	.724
Metabolic	.643	.658	.811	.769
Affective	.806	.763	.783	.844
Control	.915	.918	.929	.909

TABLE 4
Mean Unit Recall Score Correlations

Group	*AD*	*MID*	*CHI*	*Metabolic*	*Affective*	*Control*
Stories 1 & 2 combined (immediate)						
AD	–					
MID	.741	–				
CHI	.790	.770	–			
Metabolic	.768	.834	.808	–		
Affective	.678	.764	.824	.697	–	
Control	.739	.819	.846	.785	.799	–
Alpha	.871	.761	.788	.745	.801	.925
Stories 1 & 2 combined (delayed)						
AD	–					
MID	.778	–				
CHI	.729	.793	–			
Metabolic	.704	.757	.743	–		
Affective	.604	.736	.793	.732	–	
Control	.762	.866	.791	.835	.803	–
Alpha	.732	.713	.704	.713	.810	.912
RI-D[a]	.870	.723	.908	.824	.856	.913
Stories 1 & 2 combined (immediate) – corrected for attenuation						
AD	–					
MID	.910	–				
CHI	.954	.994	–			
Metabolic	.954	1.11	1.06	–		
Affective	.812	.979	1.04	.902	–	
Control	.823	.976	.991	.946	.928	–
Stories 1 & 2 combined (delayed) – corrected for attenuation						
AD	–					
MID	1.08	–				
CHI	1.02	1.12	–			
Metabolic	.836	1.06	1.05	–		
Affective	.784	.968	1.05	.963	–	
Control	.933	1.07	.987	1.04	.934	–
RI-D[a]	1.04	.772	1.02	.861	.846	.926

[a]RI-D is the correlation between immediate and delayed recall.

TABLE 5
Correlations Between Immediate and Delayed Recall

Group	*Story 1*	*Story 2*	*Stories 1 & 2*
AD	.8511	.8927	.8701
MID	.6554	.7519	.7230
CHI	.8961	.9310	.9081
Metabolic	.8001	.8355	.8238
Affective	.8580	.8365	.8555
Control	.8596	.9739	.9128

DISCUSSION

It is well established in the current clinical and research neuropsychological literature that type of lesion is an important differentiating variable as it relates to behavioral/clinical effects. The current study shows just such an effect in terms of the quantified absolute number of memory "bits" recalled. Further, the obtained ordinal rankings of the groups show roughly the expected relationships among the groups—with normal groups and affective disorder groups at one extreme of performance and MID and AD at the other—as far as the amount of material remembered is concerned.

It is equally well established that memory functions are sensitive to a variety of neurological and psychiatric conditions. All our clinical groups can be differentiated from the normal control group on that basis. These findings are expected and have been reported many times (e.g., Logue & Wyrick, 1979). The extreme drop-off effect of delay in the clinical (especially AD) group, but not in the normal group or the affective disorder group, also fits with the existing literature (e.g., Russell, 1981).

It is in the second stage of analysis, however, at the level of relative probability of a given unit's recall, that some elaboration and differentiation of existing thinking might be considered. The simple conclusion derived from the data is that the clinical groups cannot be differentiated from one another or from the normal control group on the basis of which particular units are recalled. The AD patient recalls much less than the normal older individual but recalls essentially the same major ideas. Whether this relationship somehow fits into the semantic–episodic dichotomy or might be better explained as an example of the greater existence of overpracticed fixed structures and abilities that underly semantic memory/language awaits further studies and analyses.

Rubin (1985), Rubin et al. (1981), and the current study suggest that the effect noted holds over highly diverse clinical groups; that is, the oridinal probability of a given unit being remembered remains constant regardless of which group is tested (MID, normal, or AD).

Another way of looking at this reliable effect, however, might be through examination of the characteristics of the linguistic items remembered versus those likely to be forgotten in all groups. If those characteristics (affective vs. neutral, verb vs. noun, etc.) could be manipulated, it might be possible to control, to some limited extent, the absolute number of items remembered by all groups, including the AD group. Speculation on whether this would have rehabilitative significance is premature, but at least demonstration of a measure of control memory efficacy could be a heuristic first step. Work by Davidoff et al. (1984) suggests that this might be a fruitful approach. Whether this control can be demonstrated and whether it would be on the basis of manipulation semantic versus syntactic representation remain to be seen.

ACKNOWLEDGMENTS

We thank Dr. Heyman, Patricia Neems, Elizabeth Calkins, and the staff of the Psychodiagnostic Laboratory at Duke University Medical Center for their assistance.

Partial support for this paper was provided by National Institute on Aging Grants R01-AG04278 and 1P50-AG05144.

This study reflects equal contributions by the authors.

REFERENCES

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

Bayles, K. A. (1982). Language function in senile dementia. *Brain and Language, 16,* 265–280.

Clark, E. O. (1980). Semantic and episodic memory impairment in normal and cognitively impaired elderly adults. In L. K. Obler & M. L. Albert (Eds.), *Language and communication in the elderly* (pp. 47–57). Lexington, MA: Lexington.

Corkin, S. (1982). Some relationships between global amnesias and the memory impairments in Alzheimer's disease. In S. Corkin, K. L. Davis, J. H. Growdin, E. Usdin, & R. J. Wurtman (Eds.), *Alzheimer's disease: A report of progress in research* (pp. 149–164). New York: Raven.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297–334.

Davidoff, D. A., Butters, N., Gerstman, L. J., Zurif, E., Paul, I. H., & Mattis, S. (1984). Affective/motivational factors in the recall of prose passages by alcoholic Korsakoff patients. *Alcohol, 1,* 63–69.

Erickson, R. C., & Scott, M. I. (1977). Clinical memory testing: A review. *Psychological Bulletin, 84,* 1130–1149.

Kinsbourne, M., & Wood, F. (1982). Theoretical considerations regarding the episodic–semantic distinction. In L. Cermak (Ed.), *Human memory and amnesia* (pp. 195–215). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Logue, P., & Wyrick, L. (1979). Initial validation of Russell's revised Wechsler Memory Scale: A comparison of normal aging versus dementia. *Journal of Consulting and Clinical Psychology, 47,* 176–178.

Miller, E. (1981). The nature of the cognitive deficit in senile dementia. In N. E. Miller & G. D. Cohen (Eds.), *Clinical aspects of Alzheimer's disease and senile dementia* (pp. 103–120). New York: Raven.

Nebes, R. D., Martin, D. C., & Horn, E. C. (1984). Sparing of semantic memory in Alzheimer's disease. *Journal of Abnormal Psychology, 93,* 321–330.

Power, D. G., Logue, P. E., McCarty, S. M., Rosenstiel, A. K., & Ziesant, H. A. (1979). *Journal of Clinical Neuropsychology, 1,* 343–354.

Prigatano, G. P. (1978). Wechsler Memory Scale: A selective review of the literature. *Journal of Clinical Psychology, 34,* 816–832.

Rosen, W. G. (1983). Clinical and neuropsychological assessment of Alzheimer's disease. In R. Mayeux & W. G. Rosen (Eds.), *Advances in neurology: Vol. 38. The dementias* (pp. 51–64). New York: Raven.

Rubin, D. C. (1977). Very long-term memory for prose and verse. *Journal of Verbal Learning and Verbal Behavior, 16,* 611–621.

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

Rubin, D. C. (1985). Memorability as a measure of processing: A unit analysis of prose and list-learning. *Journal of Experimental Psychology: General, 114,* 213–238.

Rubin, D. C., Olson, E. H., Richter, M., & Butters, N. (1981). Memory for prose in Korsakoff and schizophrenic populations. *International Journal of Neuroscience, 13,* 81–85.

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

Russell, E. W. (1981). The pathology and clinical examination of memory. In S. B. Filskov & T. J. Boll (Eds.), *Handbook of clinical neuropsychology* (pp. 287–319). New York: Wiley.

Walker, H. M., & Lev, J. (1953). *Statistical inference.* New York: Holt, Rinehart & Winston.

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

Weingartner, H., Graffman, J., Boutelle, W., Kaye, W., & Martin, P. R. (1983). Forms of memory failure. *Science, 221,* 380–382.

Weingartner, H., Kaye, W., Smallberg, S. A., Ebert, E. H., Gillian, J. C., & Sitaram, N. (1981). Memory failures in progressive ideopathic dementia. *Journal of Abnormal Psychology, 90,* 187–196.