Bulletin of the Psychonomic Society
1993, 31 (1), 7-10

Age differences in free recall and clustering as a function of list length and trials

SUSAN BROWN-WHISTLER and JOEL S. FREUND
University of Arkansas, Fayetteville, Arkansas

To further investigate the relationship between age and multitrial free recall of categorized lists, 144 subjects learned a categorized word list for five trials. The basic design was a 2×3 factorial combination of list length (32 or 64 words) and age (young, middle-aged, and elderly). Both list lengths were composed of eight categories, with either four or eight instances per category. Lists were presented for five alternate study-test trials, and after a 20-min filler task a final recall trial was administered. Overall, young subjects recalled more words than did either middle-aged or elderly subjects, but the two older groups did not differ. Follow-up analyses of the significant age × list length interaction indicated that there were significant age differences only on the long list. While there were systematic age-related differences in recall, there were no parallel differences in clustering, indicating that although elderly subjects do use category-organization strategies during encoding, they do not benefit as much from those strategies as do younger subjects.

The present study provides additional empirical data on multitrial free-recall learning for various age groups.

That advancing age is associated with impaired memory function on various laboratory tasks has been well established (see, e.g., Bowles & Poon, 1985; Craik & McDowd, 1987; Duchek, 1984). A substantial body of research supports the conclusion that at least some portion of this observed age difference can be attributed to processing differences that occur during encoding (Craik & Rabinowitz, 1984; Perlmutter, 1979). Evidence for this conclusion comes from studies examining subject-imposed and experimenter-imposed organization (Hasher & Zacks, 1988).

Subject-Imposed Organization

The studies in this area utilized unrelated items and various tasks to assess effectiveness of organizational strategy, including overt rehearsal (Jackson & Schneider, 1985), reorganization tasks (Zivian & Darjes, 1983), cued recall (Rabinowitz, 1986), and strategy choice (Jackson & Schneider, 1982; Perlmutter, 1978). In general, these studies show that elderly adults, relative to younger adults, are less likely to spontaneously use organizational strategies that contribute to high levels of recall (Burke & Light, 1981; Reder, Wible, & Martin, 1986; Witte, Freund, & Sebby, 1990). They are also more likely to use less effective encoding strategies (Hultsch, 1969; Jackson & Schneider, 1985; Perlmutter, 1978; Sanders, Murphy, Schmitt, & Walsh, 1980) or are not as skillful at making relational associations (Rabinowitz, 1986). However, when older subjects are instructed to organize, or are prompted by an orienting task, age differences in recall are diminished (Hulicka & Grossman, 1967).

This experiment was conducted as part of the first author's master's thesis. Financial support for the study was provided by a grant from the Marie Wilson Howells Fund. Requests for reprints should be sent to J. S. Freund, Department of Psychology, 216 Memorial Hall, University of Arkansas, Fayetteville, AR 72701-1201 (e-mail: Bitnet: jsfreund@uafsysb; Internet: jsfreund@uafsysb.uark.edu).

Given the presumed causal relationship between organization and recall (Mandler, 1979; Sternberg & Tulving, 1977), one would expect parallel effects of age on recall and measures of organization. However, studies that have provided both measures have not consistently demonstrated that relationship.

Some studies (Hultsch, 1971; Rabinowitz, Craik, & Ackerman, 1982) have reported that elderly subjects show the same levels of organization, even though differences in recall persist, whereas others (Smith, 1980; Witte, Freund, & Brown-Whistler, 1990) have provided support for the organizational deficit hypothesis. However, methodological differences among these studies may partly account for the inconsistent findings. These differences include list length, presentation order, and method of recall. Witte, Freund, and Brown-Whistler (1990) used a different random order on each trial, whereas Hultsch (1971) used the same order on every trial. Smith (1980), using written recall, showed age-related differences, whereas Jackson and Schneider (1982), using oral recall, did not find age-related organizational differences.

Experimenter-Imposed Organization

Studies in this group utilized categorized word lists and compared age groups on a variety of measures of clustering. The pattern of results from these studies is similar to that derived from experiments with unstructured lists. In general, if the inherent organization of the list is made obvious, elderly subjects benefit more in recall than do younger subjects. Thus, in each of the following circumstances, the recall of elderly subjects benefited more than did the recall of younger subjects: (1) providing cues versus no cues (Rabbitt, 1968); (2) providing instructions on how to organize versus providing no instructions (Sanders et al., 1980); (3) highly related words relative to unrelated words (Laurence & Trotter, 1971); (4) providing category superordinates at presentation or recall versus not providing the superordinates (Davis & Fried-

rich, 1982-1983); and (5) blocked versus random order of presentation of instances during learning (Bower, Clark, Lesgold, & Winzenz, 1969). Although these manipulations led to consistent age differences in recall, they typically did not result in consistent age differences in clustering, using any of the traditional measures.

The present study extended previous work by including a middle-aged group and by varying list length. Thus, it compared young, middle-aged, and elderly adults on multitrial free recall of either a 32- or a 64-item list.

METHOD

Design and Subjects

The design for this experiment was a 2×3×5 factorial with list length and age of subjects manipulated as between-subject factors and trials as a within-subject factor. Forty-eight subjects from each age group were tested. Twenty-four subjects from each group saw the short list, and 24 saw the long list.

Older subjects were recruited through citizens' organizations and at a retirement community in southern Missouri; the organizations were paid $5 for each subject's participation. Demographic data are presented in Table 1.

Recruitment at the retirement community included a few middle-aged subjects who were tested with the elderly volunteers. The remainder of the middle-aged subjects were recruited from introductory psychology classes and through the returning students' association at the University of Arkansas, Fayetteville. The subjects from this group who were not enrolled in a psychology class were paid $5 each for their participation; the students participated for partial fulfillment of course requirements. All of the young subjects were students in introductory psychology classes and participated for partial fulfillment of course requirements.

An analysis of variance (ANOVA) revealed a reliable difference in level of education among the age groups [$F(2,141) = 10.50$, $MS_e = 5.80$]. Multiple comparisons showed that the middle-aged group differed from the two extreme age groups, but that the difference in educational level between young and elderly subjects was not significant.

An ANOVA indicated that elderly, middle-aged, and young subjects did not differ on self-reported health ratings [$F(2,164) = 1.8$, $MS_e = 1.40$]. Also, subjects who were currently taking medication or who had had a medical procedure or illness that they believed had adversely affected their memories were replaced.

Materials

Words were chosen from the Battig and Montague (1969) word norms as clear representatives of their respective categories. Categories were chosen from among those showing the highest response frequencies ($M = 7.34$, $SD = 1.13$). The eight words with the highest response frequency ($M = 244.55$, $SD = 104.92$) from each of the categories were used for the lists. No two-word instances were included.

Order of presentation for each trial was determined by dividing both the 32- and the 64-item lists into eight blocks with either four or eight positions per block. Words were then randomly assigned to list positions with two restrictions: (1) 1 word from each category was included in each block, and (2) no 2 words from the same category appeared in adjacent positions between blocks. The words were typed on matte acetate and cut and mounted as slides, with one word per slide. The words were presented for 3 sec each, using a Kodak Carousel projector and associated timer.

Test booklets consisted of 7 pages. The cover page contained instructions for the experiment and an informed consent form that subjects were asked to read and sign. The second page contained questions concerning the subject's age, health, and educational level. The remaining pages were recall forms that had three columns of lines. The subjects were told to complete one column before using the next column. They were also informed that not all the lines would be needed for complete recall.

Table 1
Demographic Measures for Young, Middle-Aged, and Elderly Groups

Measure	Age Group		
	Young	Middle-Aged	Elderly
Mean age (years)	20.51 (2.44)	47.16 (4.58)	68.98 (6.52)
Age range (years)	18.58-29.92	40.00-57.00	60.33-89.00
Gender (F/M)	13/35	29/19	27/21
Education (years)	13.26 (.96)	14.83 (3.04)	12.68 (2.68)
Health	5.64	5.65	5.26

Note—Standard deviations are in parentheses. Health is rated on a scale with 1 = *poor* and 7 = *very good*.

Procedure

The subjects were run in 21 sessions, with 3-10 subjects per session. Following general instructions, the subjects completed the cover sheets and questionnaires and then turned to the first recall sheet and prepared to view the presentation of words. The final slide on each trial read "BEGIN RECALL NOW." The subjects were instructed to keep their pencils down until they saw the final slide, at which time they were to write down as many words as they could remember from the slide presentation.

To adjust for differences in list length, the subjects who received the 32-word list were given 3 min for recall; those who received the 64-word list were allowed 5 min. Following each recall trial, the subjects were asked to turn to the next blank sheet in their test booklets, put their pencils down, and prepare for another presentation. After the fifth trial, the subjects were asked to complete a Memory in Aging questionnaire (Dixon & Hultsch, 1983). The 108-item instrument took about 20 min. When all subjects had completed the questionnaire, they were asked once more to recall as many words as they could from the list that had been presented earlier. A separate recall sheet was distributed just prior to this trial to ensure that the subjects would not be expecting the final recall.

RESULTS

Immediate Recall

All tests of significance were evaluated using a 5% level of significance, and multiple comparisons utilized Tukey's procedure. The proportion recalled as a function of trials, list length, and age is presented in Figure 1. There was a significant effect of age on total recall [$F(2,138) = 12.09$, $MS_e = .11$]. Follow-up Tukey tests indicated that young subjects recalled more words than did either middle-aged or elderly subjects, but that middle-aged and elderly subjects did not differ. The age × list interaction was significant [$F(2,138) = 3.08$, $MS_e = .11$]. Separate analyses of the long and short lists indicated a reliable age effect only for the long list [$F(2,69) = 11.75$, $MS_e = .02$]. On the short list, the age effect only approached significance [$F(2,69) = 2.87$, $MS_e = .02$, $p = .08$].

As expected, recall increased across trials [$F(4,552) = 502.36$, $MS_e = .01$]. There was also a reliable effect of list length [$F(2,138) = 12.09$, $MS_e = .11$], with the short list showing the greatest benefit of trials. A significant age × list interaction [$F(2,138) = 3.08$, $MS_e = .11$] revealed that performance on the two list lengths was equivalent for young subjects, but that middle-aged and elderly subjects showed a greater proportion of recall on the short list compared with the long list.

The number of categories recalled and the number of instances per category recalled are reported in Table 2. Analysis of the number of categories recalled revealed a main effect of age [$F(2,138) = 6.65$, $MS_e = 1.19$]. The mean number of categories recalled was 7.71, 7.54, and 7.35 for the young, middle-aged, and elderly subjects, respectively. Multiple comparisons indicated that only the young and elderly subjects differed significantly. In

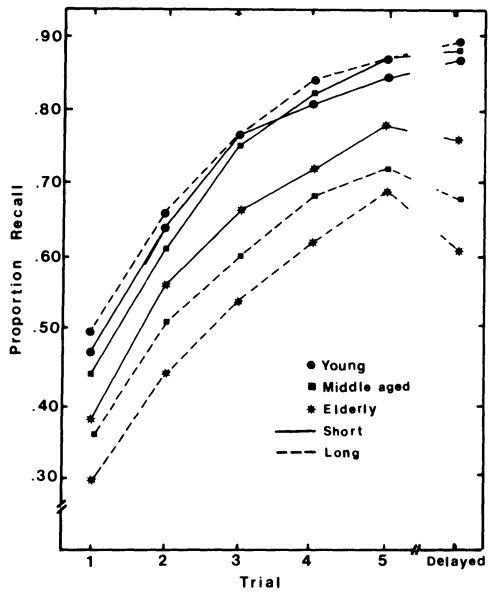

Figure 1. Proportion recalled as a function of age and list length.

addition, there was the expected effect of trials [$F(4,552) = 130.23$, $MS_e = .34$]. The significant age × trials interaction [$F(8,552) = 6.92$, $MS_e = .37$] probably reflects the fact that the young and middle-aged subjects were near ceiling performance on number of categories recalled after Trials 2 and 4, respectively.

An ANOVA on the number of items per category revealed a main effect of age [$F(2,138) = 9.81$, $MS_e = 3.64$]. On the short list, the mean number of items per category recalled was 2.94, 2.91, and 2.65 for the young, middle-aged, and elderly subjects, respectively. On the long list, the means were 5.87, 4.89, and 4.65. Multiple comparisons revealed significant differences between the young and middle-aged subjects and between the young and elderly subjects. The middle-aged and elderly subjects did not differ.

There were significant main effects of trials [$F(4,552) = 429.06$, $MS_e = .22$] and list length [$F(1,138) = 261.88$, $MS_e = 3.64$]. In addition, age interacted with list length [$F(2,138) = 4.80$, $MS_e = 3.64$]. More items per category were recalled from the long list than from the short list, and this difference was greater for the young subjects than for the middle-aged or elderly subjects. The difference in items per category recalled from short and long lists did not differ for the middle-aged and elderly subjects. The list length × trials interaction [$F(4,552) = 46.78$, $MS_e = .22$] was probably due to a ceiling effect.

Delayed Recall

An ANOVA on final delayed recall revealed a reliable main effect of age [$F(2,138) = 9.13$, $MS_e = .04$], with the young subjects showing the highest proportion of recall, followed by the middle-aged and, finally, the elderly subjects. There was also a main effect of list length [$F(1,138) = 5.37$, $MS_e = .04$], with a significantly larger proportion of words being recalled on the short list than on the long list. Results of final recall are included in Figure 1. No age × list length interaction was present.

Clustering

The results of three measures of clustering, ratio of repetition, modified ratio of repetition, and adjusted ratio of clustering, are presented in Table 3. Since all measures yielded essentially the same pattern of results, only the Fs for the adjusted ratio of clustering will be presented here. There was no reliable effect of age [$F(2,136) = .99$, $MS_e = .28$]. As would be expected, clustering increased across successive trials [$F(4,544) = 32.19$, $MS_e = .04$]. There was also a significant effect of list length [$F(1,136) = 38.70$, $MS_e = .28$], with the longer list ($M = .75$) showing greater levels of clustering than the short list ($M = .50$). This difference was constant across all age groups. Additionally, only the adjusted ratio of clustering revealed a list × trials interaction [$F(4,544) = 2.65$, $MS_e = .04$], indicating that clustering on the long list increased at a faster rate on the first two trials than clustering on the short list. The final three trials did not differ in rate of clustering between the two lists.

DISCUSSION

The pattern of age-related differences typically reported in multitrial free-recall tasks was observed in the present study. The younger subjects achieved higher proportions of recall than the older subjects. Age-related differences were greater on the long list than on the short list. Performance of the young subjects was virtually identical for the two lists, whereas the recall of both the middle-aged and the elderly subjects was significantly impaired by the longer list. The middle-aged subjects actually showed the greatest drop in performance between the two list lengths. The middle-aged subjects recalled significantly more words

Table 2
Mean Number of Categories and Words per Category Recalled

	Trials									
	Short List					Long List				
Age	1	2	3	4	5	1	2	3	4	5
	Number of Categories Recalled									
Young	7.25	7.54	7.88	7.79	7.83	7.50	7.67	7.83	7.92	7.88
Middle-Aged	6.79	7.58	7.79	7.92	7.96	6.96	7.38	7.58	7.75	7.67
Elderly	6.25	7.33	7.62	7.67	7.88	6.17	7.42	7.62	7.79	7.71
	Number of Items per Category Recalled									
Young	2.05	2.75	3.09	3.34	3.48	4.14	6.32	6.16	6.74	6.97
Middle-Aged	2.07	2.59	3.07	3.32	3.51	3.33	4.42	5.02	5.63	6.03
Elderly	1.93	2.40	2.74	3.01	3.16	3.30	3.95	4.71	5.39	5.90

Table 3
Clustering as a Function of Age, List Length, and Trial

	Trials									
	Short List					Long List				
Age	1	2	3	4	5	1	2	3	4	5
	Ratio of Repetition									
Young	.25	.33	.44	.45	.45	.59	.67	.71	.75	.78
Middle-Aged	.30	.39	.42	.49	.56	.48	.63	.68	.67	.72
Elderly	.23	.34	.38	.43	.46	.53	.57	.61	.65	.68
	Modified Ratio of Repetition									
Young	.45	.50	.61	.62	.67	.76	.81	.84	.87	.90
Middle-Aged	.55	.62	.60	.68	.76	.66	.79	.83	.80	.84
Elderly	.43	.57	.59	.62	.64	.73	.75	.78	.78	.80
	Adjusted Ratio of Clustering									
Young	.32	.40	.54	.55	.61	.71	.78	.81	.85	.88
Middle-Aged	.42	.54	.53	.62	.72	.57	.75	.80	.76	.80
Elderly	.21	.48	.50	.54	.58	.64	.67	.74	.74	.77

on the short list than did the elderly subjects, but the performance of the two older groups was equivalent on the long list.

In the present study, the three measures of clustering did not differentiate among the age groups, indicating that elderly subjects are able to detect inherent category structure of a word list and to spontaneously employ that information during recall. However, the age-related differences in recall remained, thus leaving the question of the source of the age-related deficit still unanswered.

The fact that there were no age-related differences in clustering while there were differences in recall indicates that although elderly subjects use category-organization strategies during encoding, they do not benefit as much as do younger subjects from those strategies.

REFERENCES

BATTIG, W. F., & MONTAGUE, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monographs*, **80**(3, Pt. 2).

BOWER, G. H., CLARK, M. C., LESGOLD, A. M., & WINZENZ, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning & Verbal Behavior*, **8**, 323-343.

BOWLES, N. L., & POON, L. W. (1985). Aging and retrieval of words in semantic memory. *Journal of Gerontology*, **40**, 71-77.

BURKE, D. M., & LIGHT, L. L. (1981). Memory and aging: The role of retrieval processes. *Psychological Bulletin*, **90**, 513-546.

CRAIK, F. I. M., & MCDOWD, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 474-479.

CRAIK, F. I. M., & RABINOWITZ, J. C. (1984). Age differences in the acquisition and use of verbal information: A tutorial review. In H. Bouma & D. G. Boushuis (Eds.), *Attention and performance: X. Control of language processes* (pp. 471-499). Hillsdale, NJ: Erlbaum.

DAVIS, D. D., & FRIEDRICH, D. D. (1982-1983). Memory performance of young, middle, and old adults. *International Journal of Aging & Human Development*, **16**, 311-323.

DIXON, R. A., & HULTSCH, D. F. (1983). Structure and development of metamemory in adulthood. *Journal of Gerontology*, **38**, 682-688.

DUCHEK, J. M. (1984). Encoding and retrieval differences between young and old: The impact of attentional capacity usage. *Developmental Psychology*, **20**, 1173-1180.

HASHER, L., & ZACKS, R. T. (1988). Working memory, comprehension and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193-225). Academic Press: New York.

HULICKA, I. M., & GROSSMAN, J. L. (1967). Age group comparisons for the use of mediators in paired-associate learning. *Journal of Gerontology*, **22**, 46-51.

HULTSCH, D. V. (1969). Adult age differences in the organization of free recall. *Developmental Psychology*, **6**, 673-678.

HULTSCH, D. F. (1971). Adult age differences in free classification and free recall. *Developmental Psychology*, **4**, 338-342.

JACKSON, D. K., & SCHNEIDER, H. G. (1982). Age differences in organization and recall: An analysis of rehearsal processes. *Psychological Reports*, **50**, 919-924.

JACKSON, D. K., & SCHNEIDER, H. G. (1985). Age, organization, and memory: Effects of presentation rate and rehearsal strategy. *Psychological Reports*, **56**, 471-479.

LAURENCE, M. W., & TROTTER, M. (1971). Effect of acoustic factors and list organization in multi-trial free recall learning of college age and elderly adults. *Developmental Psychology*, **5**, 202-210.

MANDLER, J. M. (1979). Category and schematic organization in memory. In C. R. Puff (Ed.), *Memory organization and structure* (pp. 303-319). New York: Academic Press.

PERLMUTTER, M. (1978). What is memory aging the aging of? *Developmental Psychology*, **14**, 330-345.

PERLMUTTER, M. (1979). Age differences in adults' free recall, cued recall and recognition. *Journal of Gerontology*, **34**, 533-539.

RABBITT, P. (1968). Age and the use of structure in transmitted information. In G. A. Talland (Ed.), *Human aging and behavior* (pp. 75-92). New York: Academic Press.

RABINOWITZ, J. C. (1986). Priming in episodic memory. *Journal of Gerontology*, **41**, 204-213.

RABINOWITZ, J. C., CRAIK, F. I. M., & ACKERMAN, B. P. (1982). A procession resource account of age differences in recall. *Canadian Journal of Psychology*, **36**, 325-344.

REDER, L. M., WIBLE, C., & MARTIN, J. (1986). Differential memory changes with age: Exact retrieval versus plausible inference. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 72-81.

SANDERS, R. E., MURPHY, M. D., SCHMITT, F. A., & WALSH, K. K. (1980). Age differences in free recall rehearsal strategies. *Journal of Gerontology*, **35**, 550-558.

SMITH, A. D. (1980). Age differences in encoding, storage, and retrieval. In L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg, & L. W. Thompson (Eds.), *New directions in memory and aging* (pp. 23-46). Hillsdale, NJ: Erlbaum.

STERNBERG, R. J., & TULVING, E. (1977). The measurement of subjective organization in free recall. *Psychological Bulletin*, **84**, 539-556.

WITTE, K. L., FREUND, J. S., & BROWN-WHISTLER, S. (1990, March). *Adult age differences in free recall and category clustering*. Paper presented at the Cognitive Aging Conference, Atlanta.

WITTE, K. L., FREUND, J. S., & SEBBY, R. A. (1990). Age differences in free recall and subjective organization. *Psychology & Aging*, **5**, 307-309.

ZIVIAN, M. T., & DARJES, R. W. (1983). Free recall by in-school and out-of-school adults: Performance and metamemory. *Developmental Psychology*, **19**, 513-520.

(Manuscript received June 24, 1992.)