

Adult Age Differences in the Temporal Characteristics of Category Free Recall

Arthur Wingfield, Kimberly C. Lindfield, and Michael J. Kahana
Brandeis University

Two experiments are reported that examined the temporal structure of recall for categorizable word lists by younger and older adults. All participants showed *response bursting*, in which recall order is clustered by semantic category, with longer interresponse times (IRTs) appearing between categories than within categories. Experiment 1 demonstrated that older adults, even when matched to younger adults in overall accuracy, differed in the rate of increase of between-category IRTs with output position, but not in within-category IRTs. Experiment 2 showed that this interaction is eliminated when the names of the response categories are provided to the participants. Results are interpreted in terms of combined effects of an age-compromised episodic memory system (between-category IRTs) accompanied by a comparatively preserved semantic system (within-category IRTs) in healthy aging.

Since the pioneering work of Tulving (Tulving & Pearlstone, 1966; Tulving & Psotka, 1971), it has been generally accepted that retrieval failure is the dominant cause of forgetting in studies of verbal learning and memory. It thus follows that to understand forgetting we must understand the mechanisms of retrieval. What are the relevant retrieval cues, and how does the rememberer use them under varying circumstances? In some memory tasks, such as recognition memory, fragment completion, and cued recall, retrieval cues are directly available. In free recall, however, the individual must initiate the formation of the retrieval cues that may facilitate access to the desired information. After some years of neglect, the cognitive literature has shown a resurgence of interest in the retrieval dynamics of free recall from episodic memory (Kahana, 1996; Kahana & Loftus, in press; Rohrer, Wixted, Salmon, & Butters, 1995; Wixted & Rohrer, 1996). The focus of these studies on episodic memory honors Tulving's (1972, 1983) distinction between episodic memory (memory for events, to include stimuli such as word lists learned in an experiment) and semantic memory (general knowledge not tied to a specific context).

The temporal output pattern observed in the free recall of verbal items such as digits or word lists is characterized by a

phenomenon known as *response bursting* (Bousfield, Cohen, & Whitmarsh, 1958; Bousfield & Puff, 1964; Bousfield & Sedgewick, 1944; Bousfield, Sedgewick, & Cohen, 1954). A clear case of this effect was described by Pollio, Richards, and Lucas (1969) who had participants learn lists of randomly arranged words drawn from several semantic categories. Under these conditions, participants tend to recall words clustered by category, with short interresponse times (IRTs) between items recalled from within a category and longer IRTs at the transitions between categories (Pollio, 1974; Pollio et al., 1969). Although IRTs generally increase exponentially as recall proceeds, Pollio et al. found that for categorized lists the increase in between-category IRTs was much larger than the increase in within-category IRTs. This increase in between-category IRTs, however, is eliminated when the names of the categories used in the study list are provided to participants either at the start of the study trials or at the time of recall (Patterson, Meltzer, & Mandler, 1971; Pollio & Gerow, 1968).

Patterson et al. (1971) postulated that between-category IRTs reflect the combined effects of the time needed to determine that no more items from a category can be recalled (category exit time) plus the time needed to gain access to the next category (category-access time). Once accessed, the participant may then use the category-retrieval cue to obtain the first word in the new category (word access). This characterization of IRTs is consistent with the view that in list learning, the learner spontaneously organizes the list into higher order units and then uses these higher order units to cue the recall of the individual words (Tulving, 1968; Wingfield & Byrnes, 1981, pp. 80–86).

In this regard, Tulving and colleagues (Tulving & Madigan, 1970; Tulving & Pearlstone, 1966) have drawn an important distinction between *cue-dependent* forgetting and *trace-dependent* forgetting. Cue-dependent forgetting refers to the failure to retrieve a higher order unit such as a category name that can serve as a retrieval cue. Trace-dependent forgetting refers to the failure to retrieve an elementary unit such as an individual word belonging to a given category. In the free recall of categorizable

Arthur Wingfield, Kimberly C. Lindfield, and Michael J. Kahana,
Department of Psychology and Volen National Center for Complex Systems,
Brandeis University.

We acknowledge support from National Institute on Aging Grants R37 AG15852 and T32 AG00204 and National Institute of Mental Health Grant R29 MH55687. We also acknowledge support from the W. M. Keck Foundation.

We thank Kelly Sullivan for help in running the participants in Experiment 2 and Anthony J. Velez for help with the latency measurements and technical support.

Correspondence concerning this article should be addressed to Arthur Wingfield, Volen National Center for Complex Systems, MS 013, Brandeis University, Waltham, Massachusetts 02254-9110. Electronic mail may be sent to wingfield@volen.brandeis.edu.

word lists, between-category IRTs could thus reflect the accessibility of the higher order memory units (e.g., category names), whereas the within-category IRTs could reflect the accessibility of the elementary memory units themselves.

These issues have special relevance in aging learning and memory. Although long-term semantic memory is relatively spared in normal aging, aging is invariably accompanied by significant declines in encoding efficiency, speed of processing, and memory for recent events (Burke & Light, 1981; Kausler, 1994; Salthouse, 1991; Smith, 1980; Tun, Wingfield, Stine, & Mecsas, 1992). Particularly vulnerable in adult aging are memory tasks described as self-initiated (Craig, 1986, 1994). These would include free-recall tasks unaided by environmental support from the presence of semantic context or other externally provided retrieval cues. By contrast, age differences are typically reduced or eliminated when environmental support is available from organizing cues present at the time of learning and recall (Craig & Jennings, 1992; Perry & Wingfield, 1994; Rankin & Firnhaber, 1986; West & Boatwright, 1983). Examining age differences in free recall of categorizable word lists, Hultsch (1975) argued that an age difference in the number of categories recalled is indicative of age-related changes in cue-dependent forgetting, whereas an age difference in the number of items recalled from within a category is indicative of age-related changes in trace-dependent forgetting. Evidence from learning rates implicated age differences in both cue- and trace-dependent forgetting.

What is the source of these age deficits in performance on self-initiated tasks? Because performance reflects both encoding and retrieval processes, age differences may be the result of either an encoding deficit, a retrieval deficit, or both. Prior research fails to distinguish these potential contributions to performance. Our aim was to examine age differences in retrieval under conditions where accuracy is equated for younger and older participants.

We conducted two experiments to examine age differences in between- and within-category IRTs in a categorized free-recall task. Our focus was not on age differences in rates of learning. From previous research one would expect older participants to require more learning trials to achieve an equivalent level of recall as the young (Hultsch, 1975; Kausler, 1994). We might also expect that older adults would be slower in beginning their recall of a category (McDowd, 1982). Rather, our focus was on potential age differences in participants' temporal output patterns and what this could tell us about retrieval factors that may underlie the well-established age differences in self-initiated recall. For example, if older adults show differentially longer IRTs than younger adults as they move from category to category in free recall, this would implicate a difficulty in accessing the appropriate retrieval cues. On the other hand, if there is a difference in the IRTs for items recalled from within a category, this would implicate an age-related memory decrement in the accessibility of individual memory traces.

Our strategy emphasizes the comparison of younger and older participants' temporal output patterns when differential amounts of practice were supplied, which allows us to place the two age groups on the same baseline of recall accuracy. In Experiment 1, we characterize the age-related retrieval deficit in self-initiated recall through an examination of within- and between-

category IRTs in categorized free recall. In Experiment 2 we attempt to modify these age differences through an experimental manipulation of retrieval cue availability.

Experiment 1

Experiment 1 was based on Pollio et al.'s (1969) original study in which they tested solely young adult participants after a single study trial. In our case participants were given multiple study-test trials in order to examine both the learning curves for older and younger participants, and the temporal structure of retrieval at a point where all items are successfully recalled.

Ideally, all participants would study the material until they reached a criterion of 100% recall. However, because some older participants had great difficulty mastering the list, training was limited to five study-test trials, and only those participants who attained 100% recall were included in the analyses. All of the younger participants tested were able to achieve this level within the five study-test trials, but twice as many older participants had to be tested to meet this criterion.

In the following experiment, we report data for 15 younger and 15 older participants who met our 100% criterion. (The characteristics of the older participants not meeting this criterion in Experiment 1 is described. In Experiment 2 all younger and older participants were trained to a 100% correct criterion, such that no data had to be excluded from that experiment.)

Method

Participants. The younger participants were 15 university undergraduates, 6 men and 9 women, with ages ranging from 17 to 23 years ($M = 19.4$, $SD = 1.7$). At time of testing the group had a mean of 13.3 years of formal education ($SD = 1.5$) and a mean Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) vocabulary score of 63 ($SD = 7.6$).

The older participants were 15 community-dwelling adults, 3 men and 12 women, with ages ranging from 64 to 80 years ($M = 71.4$ years, $SD = 5.5$). The older group had a mean of 14.9 years of formal education ($SD = 2.1$) and a mean WAIS-R vocabulary score of 68 ($SD = 8.9$). As a group, the older participants thus had an average of 1.6 more years of formal education at the time of testing, $t(28) = 2.47$, $p < .05$. The two participant groups did not differ on WAIS-R vocabulary $t(28) = 1.63$, *ns*. Both participant groups reported themselves to be in good health, and both groups were tested to insure they had no difficulty reading the words as they would be presented on the computer screen.

Stimuli. The stimulus list consisted of the original 25-item list used by Pollio et al. (1969, Experiment 2). Each of the 25 items in the list belonged to one of five categories. The five categories and their items were *animals* (bear, cat, cow, dog, horse), *gemstones* (diamond, emerald, ruby, pearl, sapphire), *trees* (birch, elm, maple, pine, oak), *transportation* (airplane, bus, car, train, truck), and *vegetables* (beans, carrots, corn, peas, potatoes). We found four of these five categories listed in Uyeda and Mandler's (1980) prototypicality norms. On a scale ranging from one (*very prototypical*) to five (*not very prototypical*), the average prototypicality ratings for the items in the categories of animals, trees, transportation, and vegetables were 2.4, 2.1, 2.2, and 2.2, respectively. Norms were not available for the gemstones category. The mean Francis and Kucera (1982) word-frequencies for the words in each of the five categories were, respectively, 92.4, 8.5, 5.0, 124.4, and 24.4 occurrences per 1 million words in print.

Procedure. Following Pollio et al.'s (1969) procedures, participants were shown each of the 25 items in the word list three times in random

order for a total list length of 75 words (25 words \times 3 repetitions). Each of the 75 words was presented one at a time in the center of a computer screen in upper-case letters. The presentation rate was one item per second. Items from the same semantic category were never presented in immediate succession and repetition of the items did not occur until all 25 words had been presented once. The end of the list presentation was signaled by a row of asterisks on the computer screen and an audible signal tone that were presented 1 s after the last word in the list had appeared. When this signal occurred, participants were to attempt to recall as many words as possible from the studied list in any order they wished. Recall was given aloud into a desk-mounted microphone. Participants were told that their responses would be scored for recall accuracy. No mention was made to the participants that we would also be examining the temporal output pattern of their responses, nor of the categorical nature of the list.

The procedure just described represented a single study-test trial. Participants who were unable to recall all 25 words in the study list were given an additional study-test trial with the words presented in a new random order. This procedure continued until the participant was either able to recall all 25 items or until five complete study-test trials had been conducted. For each trial participants were allowed as much time as they wished to recall the words.

The experiment was preceded by a brief practice session to familiarize participants with the experimental procedure. Two study lists of 10 words repeated three times in random order were used in this practice session. In order to conceal the nature of the experiment none of the items in the practice lists were drawn from the same category.

Results

Not surprisingly, the younger participants learned the list rapidly. Nine of the younger participants were able to recall all 25 words on the second trial, with the remaining participants mastering the list on their third trial. This was not the case for the older participants: Although 4 of the participants achieved 100% accuracy on the second trial, the remaining participants required three (5 participants), four (2 participants), or five (4 participants) trials to recall all 25 words.

Although the rate of learning was faster for the younger participants, participants in both age groups predominantly recalled the items clustered by categories. In some instances, participants produced words from three or four different categories and then went back to produce a new item from an already recalled category. However, this was rare. For example, on Trial 1 only 4 responses out of 317 given by the younger adults were produced out of category, and for the older adults only 7 out of 275 responses were produced out of category, $t(28) < 1$. On Trial 2 responses produced out of category were even rarer: one out of 363 responses for the younger participants and 4 out of 329 responses for the older participants $t(28) < 1$. On the participants' 100%-correct criterion trials, only 2 out of 375 responses given by the younger participants, and one out of 375 responses given by the older participants were out of category. These rare out-of-category responses were excluded from the IRT analyses. Intrusion errors (i.e., production of responses that were not in the list) were also exceedingly rare. On Trial 1 there was only one such occurrence for the younger participants and two occurrences for the older adults. On Trial 2 there were no intrusion errors for the younger participants and two for the older adults, whereas on the participants' 100%-correct criterion trials there were no such occurrences for either the younger or the older participants.¹

The two participant groups' accuracy and recall latencies are summarized in the six panels in Figure 1. The top and middle panels on the left show the first two learning and recall trials for the younger participants, and the top and middle panels on the right show these data for the older participants. The bottom two panels show the output latencies for the 100%-correct trials for the younger (lower left panel) and the older (lower right panel) participants.

The heights of the bars in Figure 1 represent the percentage of participants recalling an item in each output position. For example, the height of the seventh bar represents the likelihood of participants recalling a second item in the second category recalled. The numbered brackets above the responses indicate the category clusters in recall. It can be seen from the heights of the bars for Trials 1 and 2 in Figure 1 that participants' recall tended to fall off on the last category recalled, especially for the older participants. Indeed, we can see that many of the older participants in Trial 1 were unable to recall even the first item of the fifth category. (The orders of recalling any particular category, or order of items within a category, were quite variable between participants, and no particular preferred order of category recall, or items within a category, was seen either within or between age groups. Only responses that were given within-category were included in the mean, but as previously indicated, this criterion excluded very few responses.)

The positioning of the vertical bars along the x-axes in each of the panels represents the mean latencies (in seconds) from the signal to respond at the end of a presentation list to the onset of each of the responses in the participant's serial order of responses. Obtaining these measurements was accomplished by digitizing the real-time recordings of the participants' spoken recall responses and then using a computer generated visual display of the speech waveform to measure elapsed time from the recall signal at the end of the study list to the onset of each successive word recalled. (We chose to use onset-to-onset

¹ As indicated earlier, there were 15 older participants who did not achieve the 100% recall criterion by the fifth trial, a situation that was eliminated in Experiment 2. We refer to the 15 older participants who met the criterion and whose data are presented in Experiment 1 as the "good learners." The 15 older participants who did not meet the criterion we refer to as the "poor learners." The two groups did not differ significantly in age (71.4 years [good learners] vs. 74.1 years [poor learners], *ns*, nor WAIS-R vocabulary (67.7 [good learners] vs. 67.0 [poor learners], *ns*). The good learners were superior in both forward digit span (8.1 vs. 6.9, $p < .01$) and backward digit span (5.8 vs. 4.7, $p < .05$). Interestingly, the poor learners had completed an average of 1.5 more years of education than the good learners, $p < .05$. After the first list presentation, the poor learners had a lower baseline level of recall than the good learners. Furthermore, whereas the good learners improved steadily up to the criterion of 100% recall, the poor learners appeared to asymptote at about 80% recall with little improvement over the last three trials. The poor learners showed the basic pattern of response bursting but with longer interresponse times both between and within categories. Interestingly, this was the case even though the poor learners started their recall reports sooner than the good learners. In addition, the poor learners made more responses out of category than the good learners. Caution has to be observed with this small number of participants, but the issue of individual differences would appear to be a fruitful and important line for future inquiry.

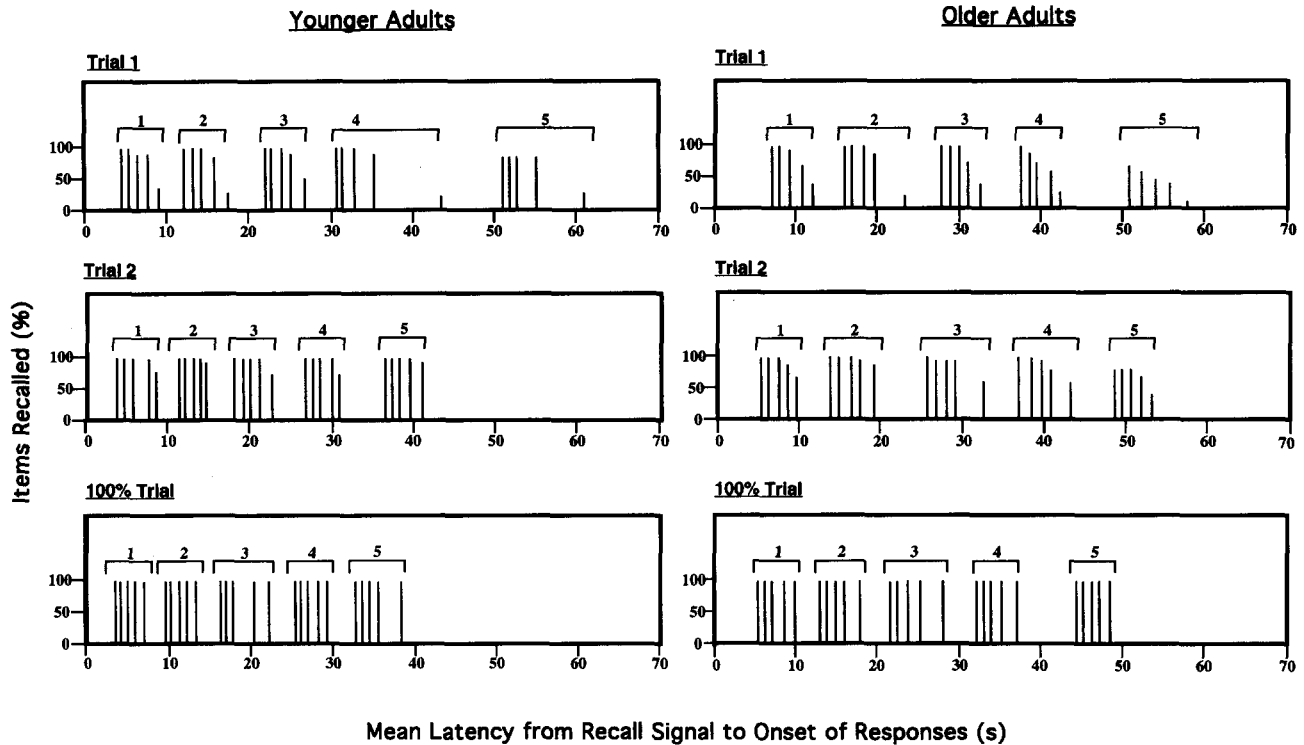


Figure 1. Recall accuracy and temporal output pattern for a categorized word list for a group of younger (left panels) and older (right panels) adults. Recall accuracy at each output position is indicated by the height of the bars. The position of each bar on the x-axis indicates the mean latency from the recall signal to the onset of each response. Numbered brackets above the bars denote category groupings.

latencies rather than latencies from the endings of recall items to the onset of the next because of participants' tendency to drag out word endings as they attempted to think of the next word.)

The top left and right panels in Figure 1 that show the younger and older participants' responses on their first learning and recall trial (Trial 1) demonstrate the classic pattern of *response bursting* in which categorically related items are recalled in rapid succession with longer interresponse times separating the category groups. An important characteristic of the temporal structure of free recall is the increase in IRTs with output position. This is seen primarily in the between-category IRTs, but also, to a lesser extent, in the within-category IRTs.

In Trial 2, the rate of responding tends to increase for both participant groups, but both groups continue to show the same pattern of category-determined response bursting. Both Trials 1 and 2 also show a tendency for the between-category IRTs to increase toward the end of the recall list. This increase is similar to that reported by Pollio et al. (1969), who had given younger adults just one learning trial. The final trial, in which all 25 items were recalled, is labeled as the *100% Trial* in Figure 1. The output pattern continues to show category-determined response bursting, although the effect remains stronger for the older participants than for the younger participants on this final trial.

Figure 2 summarizes the IRT data for the 100% criterion

trial, in which both the younger and older participants recalled all 25 list items. The two upper curves show the younger and older participants' mean *between-category* IRTs (mean latencies from the onset of the last word in one category to the onset of the first word in the next category for the four between-category transitions). These data were submitted to a 2 (age: younger, older) \times 4 (between-category output position: 1, 2, 3, 4) mixed design analysis of variance (ANOVA) in which age was a between-subjects variable and position was a within-subjects variable. A significant main effect of age confirmed the appearance in Figure 2 that the older participants' between-category IRTs were generally longer than those of the younger participants, $F(1, 28) = 5.12$, $MSE = 12.71$, $p < .05$. A general pattern of increasing IRTs across the four between-category transitions was evidenced by a significant main effect of category output position, $F(3, 84) = 8.38$, $MSE = 4.15$, $p < .0001$. We can see from Figure 2, however, that this increase in between-category IRTs was dramatically larger for the older participants than for the younger participants, especially for the transition from the fourth to the fifth category in recall. This difference was reflected by a significant Age \times Output Position interaction, $F(3, 84) = 4.18$, $MSE = 4.15$, $p < .007$.

The two lower curves in Figure 2 summarize the younger and older participants' *within-category* IRTs averaged across all categories that were recalled on their 100%-correct trials. These data were analyzed with a separate 2 (age: younger, older) \times

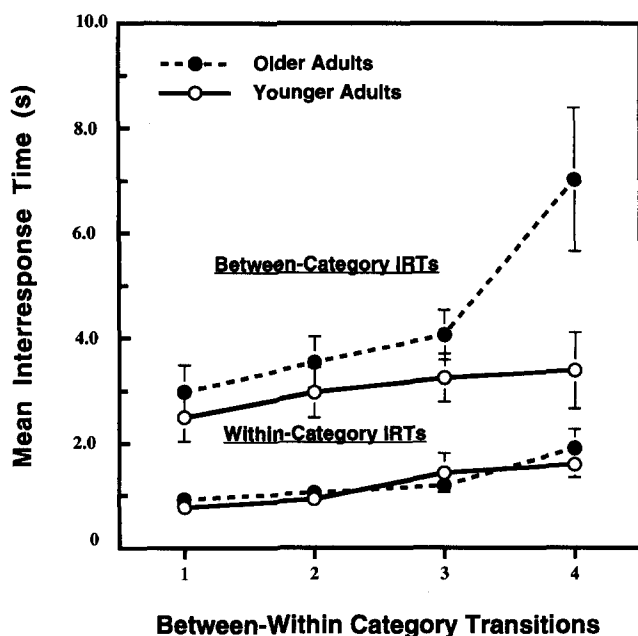


Figure 2. Top two curves show mean between-category interresponse times (IRTs) for younger and older adults over the four transitions between the five recalled categories. Bottom two curves show the mean within-category IRTs averaged across the five categories recalled for both younger and older adults. Error bars represent one standard error. (In some cases error bars are not visible because the standard errors were too small to plot.)

4 (within-category output position: 1, 2, 3, 4) mixed design ANOVA, with age as a between-subjects variable and position as a within-subjects variable. One can see a general trend for latencies to increase from item-to-item within the categories, as confirmed by a significant main effect of output position, $F(3, 84) = 14.33$, $MSE = 0.39$, $p < .0001$. In this case, however, there were no age differences, either in terms of a main effect of age, $F(1, 28) < 1$ ($MSE = 0.86$), or a significant Age \times Output Position interaction, $F(3, 84) < 1$ ($MSE = 0.39$). Thus, the general slowing shown by the older adults when moving from one category to another, and its amplification in the last between-category transition, did not appear in the within-category IRT data.

Experiment 2

In Experiment 1 we observed that the growth of between-category IRTs with output position was exaggerated for the older relative to the younger participants. To the extent that this effect is due to an age-related difficulty in retrieving the category cues, supplying category names to participants during learning and recall would be expected to attenuate this difference. By contrast, there should be no reason to believe that providing category cues would affect within-category IRTs. Because our focus would be on between-category IRTs, we increased the number of response categories from the five used by Pollio et al. (1969) to seven categories. In order to accommodate this increase while

still keeping the lists to within manageable lengths, we reduced the number of items per category from five to three.

In Experiment 1, as indicated, we wished to follow Pollio et al.'s (1969) procedures as closely as possible. We did this by using the same word list they employed and by using their procedure of presenting all of the words in the list three times before testing recall. As Pollio and colleagues argued, and then went on to demonstrate, it was not the case that the pattern of response bursting we have been describing was due to the particular word list that was employed (Pollio et al., 1969). We show this also in Experiment 2 by using three different word lists created for the experiment.

Two additional changes were made in Experiment 2. First, each learning trial consisted of just one presentation of each item rather than three. Second, we imposed no upper limit on the number of trials each participant would receive, such that no participant was excluded from the experiment solely on the basis of the number of trials he or she needed to reach a 100%-correct criterion.

Method

Participants. The younger participants were 24 university undergraduates, 5 men and 19 women, with ages ranging from 18 to 22 years ($M = 18.9$, $SD = 1.2$). At time of testing the group had a mean of 13.5 years of formal education ($SD = 1.0$), and a mean WAIS-R vocabulary score of 59 ($SD = 8.6$).

The older participants were 24 community-dwelling adults, 12 men and 12 women, with ages ranging from 64 to 81 years ($M = 72.9$ years, $SD = 5.3$). The older group had a mean of 17.1 years of formal education ($SD = 1.9$) and a mean WAIS-R vocabulary score of 69 ($SD = 9.4$). As a group, the older participants thus had an average of 3.6 more years of formal education at the time of testing, $t(46) = 8.08$, $p < .001$. The older participants were also superior on WAIS-R vocabulary $t(46) = 3.85$, $p < .001$. Both participant groups reported themselves to be in good health, and both groups were tested to insure they had no difficulty reading the words as they would be presented on the computer screen.

Stimuli. Three different 21-word lists were constructed for this experiment. Each list contained seven categories of items represented by 3 words per category. The category labels and words for each of the lists are given in the appendix. The mean prototypicality ratings for the category exemplars were similar for each of the three lists: 12.5, 12.7, and 13.0, on the basis of the Uyeda and Mandler (1980) norms. The words in the three learning lists were also similar in average word frequency, averaging 512, 517, 519 occurrences per million words in print (Francis & Kucera, 1982).

Procedure. An equal number of individuals participated in the non-cued and cued conditions by random assignment. (The participants selected for the noncued and cued conditions did not differ significantly in years of education or WAIS-R vocabulary.) Within each of the conditions, participants were assigned to one of the three study lists. For both the noncued and the cued conditions, the words in the study list were presented one at a time in random order in the center of the computer screen. Presentation rate was again at one item per second. As in Experiment 1, the end of the list was signaled by an audible tone and a row of asterisks that appeared on the screen 1 s after the last word in the study list. Participants were asked to recall as many of the words from the study list as possible in any order they chose. Participants' spoken responses were recorded for later scoring. As in Experiment 1, no mention was made of our intention to measure temporal output patterns.

If the participant could not recall all 21 words on the list, the list was presented again following the same procedures, but with the words

presented in a new random order. This procedure was repeated for each participant until they could recall all of the list items.

Individuals who participated in the cued condition were told that the items in the list belonged to seven semantic categories. The names of these categories were displayed on an index card placed above the computer screen. The category names remained on display throughout both the study and recall phases of each trial. For each study list there were four orders in which the category names could appear on the card, and these orders were counterbalanced across participants. The list presentations in the noncued condition were the same except that no mention was made of the fact that the list items were drawn from a number of specific categories.

The experiment was preceded by a brief practice session to familiarize participants with the experimental procedure. A list of 10 common nouns not used in the main experiment made up the practice list. In order to conceal the nature of the experiment, none of the items in the practice list were drawn from the same category.

Results

As would be expected, the younger participants needed fewer trials to reach a 100% correct criterion than the older participants for both the noncued condition (mean trials to criterion: younger = 3.5; older = 5.9) and the cued condition (mean trials

to criterion: younger = 3.3; older = 5.1). Mann-Whitney *U* tests showed the age difference to be significant for both the noncued ($p < .002$) and the cued ($p < .02$) conditions. As can be seen, the younger participants did not differ in the number of trials needed to reach a 100% criterion in the two cue conditions. There was, however, a marginal difference for the older participants ($p = .06$). (One older participant in the cued condition required 12 trials to reach criterion, twice as long as the next slowest older participant. With this participant excluded, the mean number of trials to criterion in the cued condition reduces to 4.5 and raises the significance level of the difference between the two cue conditions for the older participants to $p < .02$.)

The older participants' need for fewer trials to reach criterion in the cued condition is consistent with the general literature on the value of supportive cues on older adults' learning and memory performance (Craik & Jennings, 1992; Smith, 1977). Our interest, however, is on the temporal output pattern of recall for the younger and older participants when both groups had mastered the list to 100% accuracy.

Figure 3 shows the mean between- and within-category IRTs for the noncued (left panel) and cued (right panel) conditions on the 100%-correct criterion trials for the younger and older

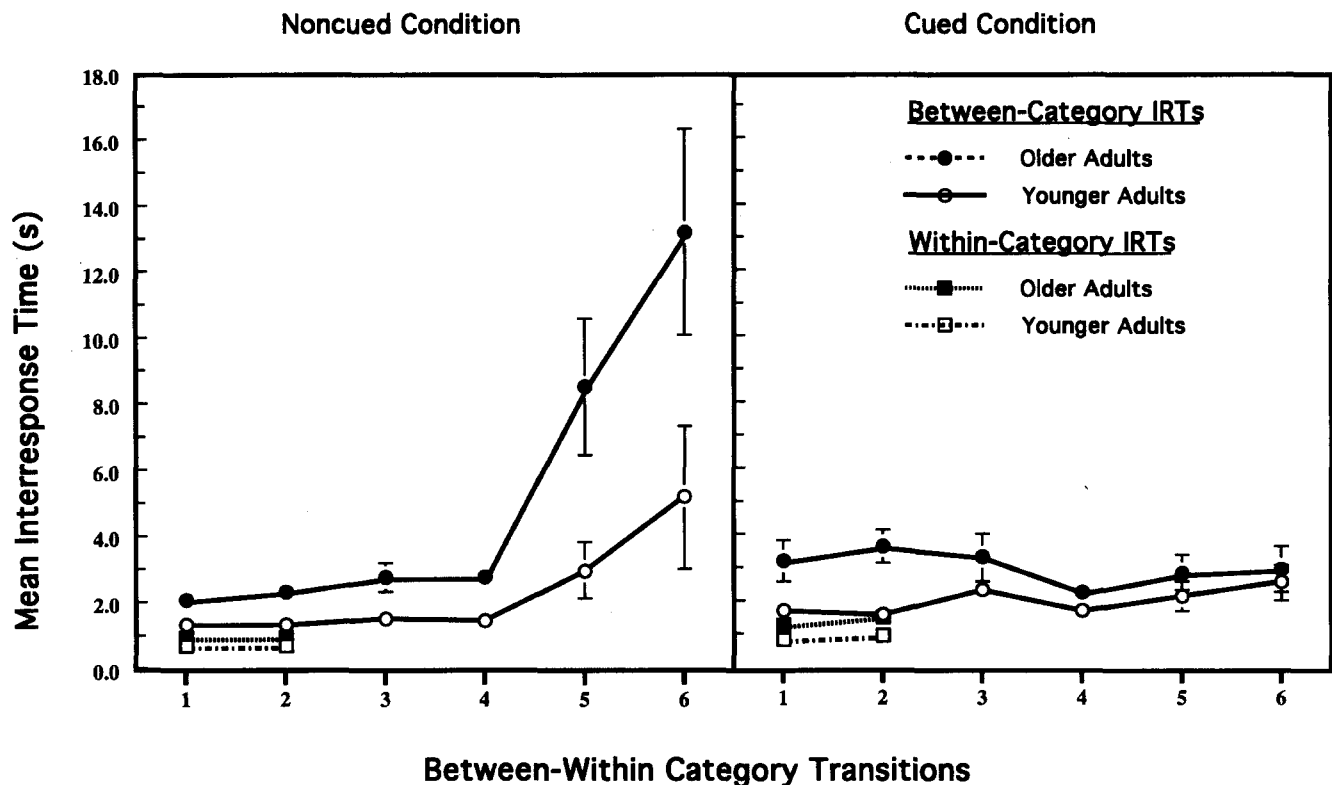


Figure 3. In each panel, the top two curves show between-category interresponse times (IRTs) for younger and older adults over the six transitions between the seven recalled categories. Bottom two curves show the two within-category IRTs averaged across the seven categories recalled for both younger and older adults. The left panel shows data obtained when category names were not supplied to participants (noncued condition), and the right panel shows data obtained when category names were supplied to participants (cued condition). Error bars represent one standard error. (In some cases error bars are not visible because the standard errors were too small to plot.)

participants. As in Experiment 1, items from the same category were generally recalled together. On the 100%-correct criterion trial for the noncued condition, only 13 out of 252 responses given by the younger participants and 16 out of 252 responses given by the older participants were produced out of category. In the cued condition, 10 out of the 252 responses given by the younger participants and 6 out of the 252 responses given by the older participants were produced out of category. These responses were excluded from the latency analysis. As in Experiment 1, intrusion errors were rare. On the 100%-correct criterion trials in the noncued condition, no intrusion errors were produced by either the younger or the older participants. In the cued condition there were no intrusion errors produced by the younger participants and only one intrusion produced by a single older participant.

Noncued condition. The upper two curves in the left panel show the mean between-category IRTs for the younger and older participants in the noncued condition for the 7 three-item categories represented in the lists. For both age groups the IRTs between recall of the initial categories are relatively short, increasing markedly toward the ends of the lists. This effect was clearly exaggerated for the older participants. These observations were confirmed by a two-way, mixed-design ANOVA conducted on the noncued between-category IRTs. There were significant main effects of age, $F(1, 22) = 12.76$, $MSE = 25.36$, $p < .0005$, of output position, $F(5, 110) = 12.16$, $MSE = 19.62$, $p < .0001$, and a significant Age \times Output Position interaction, $F(5, 110) = 3.02$, $MSE = 19.62$, $p < .015$.

Having only three items per category in this experiment, one can plot only two points for each of the within-category IRTs. These are shown for the younger and older participants in the noncued condition as the lower two curves in the left panel of Figure 3. A two-way ANOVA conducted on these data failed to show a significant effect of age, $F(1, 22) = 2.89$, $MSE = .34$, *ns*, nor was there a significant main effect of output position, $F(1, 22) < 1$ ($MSE = 0.01$), or an Age \times Position interaction, $F(1, 22) < 1$ ($MSE = 0.01$).

Cued condition. The right panel of Figure 3 shows the data for the cued condition in which the category names were available to the participants. Although the older participants continued to have longer between-category IRTs than the younger participants, supported by a significant main effect of age, $F(1, 22) = 10.03$, $MSE = 4.11$, $p < .005$, there was no longer a significant main effect of output position, $F(5, 110) = 1.14$, *ns*, $MSE = 2.13$, nor was there a significant Age \times Output Position interaction, $F(5, 110) = 1.03$, *ns*, $MSE = 2.13$.

For the within-category IRTs, there was a main effect of age, $F(1, 22) = 4.97$, $MSE = 0.59$, $p < .04$. Although there appears to be a trend for the second within-category IRT to be slightly longer than the first, this difference was not significant, $F(1, 22) = 2.67$, *ns*, $MSE = 0.15$. There was also no Age \times Output Position interaction, $F(1, 22) < 1$, $MSE = 0.15$.

As might be expected from this picture of main effects and interactions, an omnibus ANOVA including the factors of age, cue condition and output position revealed a significant three-way interaction for the between-category IRTs, $F(5, 220) = 3.39$, $MSE = 10.87$, $p < .005$. A similar analysis for the within-category IRTs, however, did not show a significant effect, $F(1, 44) < 1$, $MSE = 0.08$.

Discussion

The primary aim of these experiments was to use the temporal output pattern of free recall to examine age-differences in retrieval from episodic memory. The categorized free-recall task was chosen because it requires self-initiated recall of lists containing both highly associated items and unrelated categories. Older participants are known to have great difficulty in self-initiated recall tasks (Craik, 1986, 1994), but they are also aided by the presence of powerful interitem associations (Craik & Jennings, 1992; Perry & Wingfield, 1994).

Experiments 1 and 2 both demonstrated that the classic pattern of response bursting in categorized free recall previously demonstrated for younger adults (Patterson et al., 1971; Pollio et al., 1969) also holds for older adults. That is, both younger and older participants recalled items grouped predominantly by category, with temporal output patterns characterized by fast within-category IRTs and slow between-category IRTs.

The literature on category clustering in free recall in younger and older participants is an extensive one (e.g., see the review by Kausler, 1994, pp. 230–238). As Kausler notes, many of these studies reflect a confounding of the occurrence of clustering in recall and the number of items recalled. The present experiments have the important feature of examining output latencies for younger and older participants when both groups had reached a criterion of 100% recall. In this way we were able to examine the aging interaction with IRT without the confound of differences in degree of original learning.

The observed finding of long between-category IRTs and short within-category IRTs is consistent with the notion of a content-addressable memory system in which individual category items are accessed using the superordinate category as a retrieval cue (Tulving & Psotka, 1971). Experiments 1 and 2 examined these temporal output patterns both at the early stages of learning and after the participants had achieved total list recall. Experiment 1 and the noncued condition of Experiment 2 showed that even after both groups had reached a 100% correct criterion, a significant age difference appeared in the between-category IRTs. Further, the older participants showed a steeper rate of increase in between-category IRTs as recall proceeded from category to category. However, once a category was found, the older participants could recall within-category items as rapidly (Experiment 1) or almost as rapidly (Experiment 2) as the younger participants. It would thus appear that the older participants' retrieval difficulty is focused primarily on accessing the appropriate semantic category, but once that category has been accessed, retrieval is relatively unimpaired.

It might be suggested that the rapid within-category IRTs resulted from a strategy of participants simply generating highly prototypical category members as each category name was activated, rather than reflecting an actual recall response. Our analysis of error responses suggests that this was unlikely to have been the case. That is, such a strategy would be expected to produce not only responses that happened to be correct but also intrusions of highly prototypical members of the studied categories that had not been in the list (Deese, 1959; Roediger & McDermott, 1995). Under our experimental conditions, however, such intrusions were extremely rare, suggesting that the participants were not simply generating highly prototypical cate-

gory items. The rarity of extralist intrusions even among older participants is especially interesting in the context of arguments favoring an inhibition deficit in normal aging (Hasher & Zacks, 1988). Were such an inhibition deficit operating, one might have expected to see a high incidence of extralist intrusions in the responses of the older adults. As we saw, however, this was not the case.

A finding of increasing IRTs with output position, whether between categories or within categories, can be accounted for in several different ways. One of these is the *random search with replacement retrieval model* (McGill, 1963; Murdock & Okada, 1970; Wixted & Rohrer, 1996). According to this account, participants search through memory until they detect a recently activated item (with the recency of the activation signaling that the item was a member of the learning set), but not an item that had already been recalled. Assuming that it takes time to sample each item, as the ratio of recalled to nonrecalled items increases in the memory set, the time to recall the remaining items increases as a consequence of the resampling and rejection of items already recalled. Pollio et al.'s (1969) interpretation of their data on between-category effects was closely related to the random search with replacement view.

The random search with replacement account of IRTs in free recall can be seen as a specific instantiation of the general notion of *output interference*: that the act of recalling list items impairs access to other list items (Kausler, 1994, pp. 267–268; Smith, 1975; Tulving & Arbuckle, 1963). Such interference could be due to the resampling of the recalled items or to a direct effect of the recalled list items on the accessibility of the not-yet-recalled memories. In either way, output interference and random search with replacement proposals can be contrasted with a strength-based account for the growth of IRTs with output position. In a strength-based model, items with the strongest representations are recalled first and fastest (Hogan, 1975). The remaining items, being necessarily weaker, take longer to recall, thus producing the accelerating interresponse times with output position. Note that this view does not assume that recalling some items has an effect on recall of subsequent items.

It is not surprising to see a slowing of responses in the older participants relative to the younger participants, as slowing is a virtually ubiquitous finding in aging research (Cerella, 1994; Salthouse, 1991). It is important to note, however, that in our experiments we found a clear age dissociation in the difference between the rate of growth of between-category IRTs versus within-category IRTs. These results support those who have suggested that a single slowing factor would be an unlikely account for the full range of age-related performance differences one observes in the literature (Fisher & Glaser, 1996; Fisk & Fisher, 1994). One suggestion has been that slowing does not occur in tasks that involve lexical processes to the extent that it is observed in nonlexical task domains (Hale & Myerson, 1996). This particular distinction, however, does not capture the age-difference in growth rates for between- versus within-category IRTs we observed within a single task domain.

Our finding of an age interaction in the rate of growth of between-category IRTs but not within-category IRTs in both Experiment 1 and the noncued condition of Experiment 2 is not consistent with a deficit of older adults in some basic process of output interference, memory strength, or generalized slowing.

Rather, it is our view that these temporal dissociations are a direct reflection of the operation of episodic and semantic factors in retrieval in younger and older participants. The first part of this distinction lies in the likelihood that within-category items are part of a semantic network representing associative connections that already existed prior to the participant entering the experiment. The learning task in this case involves only temporally based activation (i.e., activation by recency) of preexisting associations. For example, suppose a participant encounters the words *diamond* and *ruby* in the study list. These words are already bound together as part of the gemstone concept. One need only register the recency of their activation, as distinct from the nonactivation of other gemstone names that were not in the list.

By contrast, no preexisting binding existed for the particular five or seven category names used in the two experiments reported. That is, although *diamond* and *ruby* as category exemplars have a preexisting relationship (for example, one can cue the other in recall), the different categories represented in the experiment have no preexisting connections between them. In this case, the participant must learn *animals*, *gemstones*, *trees*, *transportation*, and *vegetables* as associations for the first time. It mirrors, in fact, the typical paired-associate learning task (albeit with the need to associate five or seven arbitrary categories), a task on which older adults are known to have far more difficulty than younger adults (Arenberg, 1967; Kausler, 1994, pp. 66–77). Thus, the within-category responses are more resistant to aging because learning of these items builds on preexisting relations within the semantic system. That is, within-category items are related through a network of semantic associations, such that performance in free recall is supported by knowledge structures learned long before the participant entered the laboratory.

If the age deficit lies in the process of accessing the studied category, then supplying participants with the category names should eliminate the age differences in between-category IRTs. This is not to say that a main effect of age would necessarily be eliminated; older participants might still be slower in producing their responses. Rather, one should see the elimination of age differences in the rate of growth of the between-category IRTs with output position. This interpretation is supported by the finding in Experiment 2 that the age difference in the growth of between-category IRTs was eliminated when participants were provided with category labels as retrieval cues. This finding is another demonstration of the role of environmental support in adult aging as a moderator of age differences (e.g., Craik & Jennings, 1992).

The finding that healthy older adults require more trials to learn a list of items than the younger participants is consistent with an age-related learning deficit. Examining the temporal dynamics of retrieval after younger and older participants have all attained a common criterion of 100% correct recall, however, allowed us to look at retrieval factors beyond age differences in original learning. As we saw, even when the younger and older participants achieved mastery of a list, attaining a criterion of 100% correct, significant age differences in the rate of increase in between-category IRTs were still observed. This finding supports the view that older participants also have a specific impairment in retrieval within episodic memory.

The suggestion from our data of an age-related dissociation between episodic and semantic memory is not intended to argue that these are independent subsystems (McKoon, Ratcliff, & Dell, 1986; Tulving, 1986). Rather, we use these terms descriptively to distinguish between experimentally determined versus preexperimentally determined associations or, in a broader context, between personal episodes and semantic knowledge (Tulving, 1972, 1983). Indeed, to say that older adults show an impairment in episodic memory serves more to frame the question than to supply an answer. Our goal must thus be twofold: first, to specify those operations that translate experienced events into episodic representations and, second, to specify the locus of the age-related deficit in this process.

Memory for the presence of a word in a word list requires that participants form an association between the unit of information that is the word and some representation of the time-bound event that is the word list. Failure to recall that a word was in a learned list does not imply that the word or its meaning were lost; both remain firmly fixed in semantic memory. What has been lost is the knowledge of whether that word had occurred in a just-experienced word list. That is, a critical aspect of episodic memory, as distinct from semantic memory, is a representation of when the information was encountered. This may be referred to as the *temporal context* of an item. It is the integration of temporal context with semantic information that makes the memory of the event unique. We refer to this process as *trace individuation*.

Within this framework, trace individuation underlies the encoding specificity principle (Tulving & Thomson, 1973). Because studied information is learned in a temporal context, cues provided at retrieval are effective to the extent that they match conditions present during study and thereby reinstate the learning context. In the absence of semantic connections that lead to trace individuation, or in the absence of semantic cueing through environmental support (Craig & Jennings, 1992), the trace must be identified solely by its temporal context. It is here that we believe the aging deficit is most pronounced.

How might temporal context be represented? Just as memories have perceptual and semantic features, memories also have temporal features (Underwood, 1969). In the context of our experiments, these temporal features enable participants to know that two unrelated categories, such as *gemstones* and *animals*, were both present in the studied list. Trace individuation requires both the coding of temporal context and the binding of this temporal information with the semantic and perceptual features of the event. An age-related deficit in either of these processes would thus severely impair explicit memory performance. In the laboratory, such deficits would reveal themselves in impaired judgments of temporal ordering (Naveh-Benjamin, 1990), higher levels of false memory (Norman & Schacter, 1997; Tun, Wingfield, Rosen, & Blanchard, 1998), and a differential impairment in category access relative to exemplar retrieval as we observed in our experiments.

An impairment in the process of temporal coding has implications for both encoding and retrieval. During encoding, participants may be less likely to effectively bind the temporal context with the learned information. During retrieval, participants may not effectively use temporal context to guide self-initiated recall. In everyday memory, as in laboratory experiments, retrieval is

supported by participant-generated and in some cases externally provided cues. Self-initiated retrieval (e.g., free recall) depends critically on the ability to form a retrieval cue that includes a representation of the temporal context in which the desired information was encoded. By contrast, retrieval that is supported substantially by experimenter-provided cues (e.g., recognition or cued recall) does not depend as critically on participants' ability to form a representation of temporal context. Thus, an age-related deficit in the ability to use temporal context in retrieval would result in severe impairments on those tasks in which externally provided cues are minimal (e.g., free recall), but relative sparing of performance in tasks that provide external cues and contextual support (e.g., recognition or cued recall). Consequently, in the case of tasks that make no demands on participants to consciously direct retrieval to a studied event (e.g., implicit memory), one would expect to find nearly equivalent performance for younger and older participants (Mitchell, 1989; Nyberg, Bäckman, Erngrund, Olofsson, & Nilsson, 1996).

We suggest that it is not so much that older adults have difficulty forming associations between items as it is that they have a difficulty in the encoding and use of temporal information. This specifies a possible locus for the age-related deficit in episodic memory. That is, an age-related impairment in temporal coding would give rise to both learning and retrieval deficits. During learning, more trials would be needed to ensure the effective binding of temporal context and content information. During retrieval, even after effective encoding has been achieved, successful retrieval would depend on effective use of temporal information. An age-related impairment in this regard would give rise to the progressive slowing of between-category IRTs with output position as observed in our older participants.

The view that age-related changes in memory reflect both storage and retrieval deficits is captured in the Mishnaic adage, "What a child learns is comparable to ink written on a new sheet; but what an old person learns is comparable to ink written on a parched sheet." It is hard both to read a parched sheet and to write on it.² Our finding that the age-related retrieval deficit is greater in between-category IRTs, as opposed to within-category IRTs, illustrates the importance of preexisting memory structures in the acquisition and use of new information. The structure of semantic memory was formed before the parchment became coarse. As such, when new information can be framed within these knowledge structures, the retrieval deficits associated with aging are greatly attenuated. When new information cannot be framed in terms of existing knowledge structures, participants must rely on the encoding and use of temporal information. It is under these conditions that older adults show the greatest memory difficulties.

² The Mishna is a compendium of ancient Hebrew law and wisdom compiled between the 6th and 1st centuries B.C. This saying is taken from Tractate Avot (chap. 4, Mishna 20) and is attributed to the Tanna Elisha ben Abuya. We are grateful to Moshe Naveh-Benjamin (personal communication, April 12, 1997) for calling our attention to this adage and its implications for aging memory.

References

- Arenberg, D. (1967). Age differences in retroaction. *Journal of Gerontology*, 22, 88-91.
- Bousfield, W. A., Cohen, B. H., & Whitmarsh, G. A. (1958). Associative clustering in the recall of words of different taxonomic frequencies of occurrence. *Psychological Reports*, 4, 39-44.
- Bousfield, W. A., & Puff, C. R. (1964). Clustering as a function of response dominance. *Journal of Experimental Psychology*, 67, 76-79.
- Bousfield, W. A., & Sedgewick, C. H. (1944). An analysis of sequences of restricted associated responses. *Journal of General Psychology*, 30, 149-165.
- Bousfield, W. A., Sedgewick, C. H., & Cohen, B. H. (1954). Certain temporal characteristics of the recall of verbal associates. *American Journal of Psychology*, 57, 111-118.
- Burke, D. M., & Light, L. L. (1981). Memory and aging: The role of retrieval processes. *Psychological Bulletin*, 90, 513-546.
- Cerella, J. (1994). Generalized slowing and Brinley plots. *Journal of Gerontology: Psychological Sciences*, 49, P65-P71.
- Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities, mechanisms, and performances* (pp. 409-422). Amsterdam: Elsevier.
- Craik, F. I. M. (1994). Memory changes in normal aging. *Current Directions in Psychological Science*, 3, 155-158.
- Craik, F. I. M., & Jennings, J. M. (1992). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of aging and cognition* (pp. 51-110). 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.
- Fisher, D. L., & Glaser, R. A. (1996). Molar and latent models of cognitive slowing: Implications for aging, dementia, depression, development, and intelligence. *Psychonomic Bulletin and Review*, 3, 458-480.
- Fisk, A. D., & Fisher, D. L. (1994). Brinley plots and theories of aging: The explicit, muddled, and implicit debates. *Journal of Gerontology: Psychological Sciences*, 49, P81-P89.
- Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin.
- Hale, S., & Myerson, J. (1996). Experimental evidence for differential slowing in the lexical and nonlexical domains. *Aging, Neuropsychology, and Cognition*, 3, 154-165.
- 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). San Diego, CA: Academic Press.
- Hogan, R. M. (1975). Interitem coding and directed search in free recall. *Memory & Cognition*, 3, 197-209.
- Hultsch, D. F. (1975). Adult age differences in retrieval: Trace-dependent and cue-dependent forgetting. *Developmental Psychology*, 11, 197-201.
- Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, 24, 103-109.
- Kahana, M. J., & Loftus, G. (in press). Accuracy and latency in human cognition. In R. Sternberg (Ed.), *The concept of cognition*. Cambridge, MA: MIT Press.
- Kausler, D. M. (1994). *Learning and memory in normal aging*. San Diego, CA: Academic Press.
- McDowd, J. M. (1982). *Attention switching in young and old adults*. Unpublished master's thesis, University of Toronto, Toronto, Ontario, Canada.
- McGill, W. J. (1963). Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 309-360). New York: Wiley.
- McKoon, G., Ratcliff, R., & Dell, G. S. (1986). A critical evaluation of the semantic-episodic distinction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 295-306.
- Mitchell, D. B. (1989). How many memory systems? Evidence from aging. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 31-49.
- Murdock, B. B., & Okada, R. (1970). Interresponse times in single-trial free recall. *Journal of Verbal Learning and Verbal Behavior*, 86, 263-267.
- Naveh-Benjamin, M. (1990). Coding of temporal order information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 117-126.
- 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.
- Nyberg, L., Bäckman, L., Erngrund, K., Olofsson, U., & Nilsson, L.-G. (1996). Age differences in episodic memory, semantic memory, and priming: Relationships to demographic, intellectual, and biological factors. *Journal of Gerontology: Psychological Sciences*, 51, P234-P240.
- Patterson, K. E., Meltzer, R., & Mandler, G. (1971). Inter-response times in categorized free recall. *Journal of Verbal Learning and Verbal Behavior*, 10, 417-426.
- Perry, A. R., & Wingfield, A. (1994). Contextual encoding by young and elderly adults as revealed by cued and free recall. *Aging and Cognition*, 1, 120-139.
- Pollio, H. R. (1974). *The psychology of symbolic activity*. Reading, MA: Addison-Wesley.
- Pollio, H. R., & Gerow, J. R. (1968). The role of rules in recall. *American Journal of Psychology*, 81, 303-313.
- Pollio, H. R., Richards, S., & Lucas, R. (1969). Temporal properties of category recall. *Journal of Verbal Learning and Verbal Behavior*, 8, 529-536.
- Rankin, J. L., & Firnhaber, S. (1986). Adult age differences in memory: Effects of distinct and common encodings. *Experimental Aging Research*, 12, 141-146.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 803-814.
- Rohrer, D., Wixted, J. T., Salmon, D. P., & Butters, N. (1995). Retrieval from semantic memory and its implications for Alzheimer's disease. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1127-1139.
- Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.
- Smith, A. D. (1975). Aging and interference with memory. *Journal of Gerontology*, 30, 319-325.
- Smith, A. D. (1977). Adult age differences in cued recall. *Developmental Psychology*, 13, 326-331.
- 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.
- Tulving, E. (1968). Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 2-36). Englewood Cliffs, NJ: Prentice Hall.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381-403). New York: Academic Press.
- Tulving, E. (1983). *Elements of episodic memory*. Oxford, United Kingdom: Clarendon.
- Tulving, E. (1986). What kind of a hypothesis is the distinction between

- episodic and semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 307-311.
- Tulving, E., & Arbuckle, T. Y. (1963). Sources of intratrial interference in immediate recall of paired associates. *Journal of Verbal Learning and Verbal Behavior*, 1, 321-334.
- Tulving, E., & Madigan, S. A. (1970). Memory and verbal learning. *Annual Review of Psychology*, 21, 437-484.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5, 381-391.
- Tulving, E., & Psotka, J. (1971). Retroactive inhibition in free recall: Inaccessibility of information available in the memory store. *Journal of Experimental Psychology*, 87, 1-8.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 359-380.
- Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging*, 13, 230-241.
- Tun, P. A., Wingfield, A., Stine, E. A. L., & Mecsas, C. (1992). Rapid speech processing and divided attention: Processing rate versus processing resources as an explanation of age effects. *Psychology and Aging*, 7, 546-550.
- Underwood, B. J. (1969). Attributes of memory. *Psychological Review*, 76, 559-573.
- Uyeda, K. M., & Mandler, G. (1980). Prototypicality norms for 28 semantic categories. *Behavior Research Methods & Instrumentation*, 12, 587-595.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. New York: Psychological Corporation.
- West, R. L., & Boatwright, L. K. (1983). Age differences in cued recall and recognition under varying encoding and retrieval conditions. *Experimental Aging Research*, 9, 185-189.
- Wingfield, A., & Byrnes, D. L. (1981). *The psychology of human memory*. New York: Academic Press.
- Wixted, J. T., & Rohrer, D. (1994). Analyzing the dynamics of free recall: An integrative review of the literature. *Psychonomic Bulletin and Review*, 1, 89-106.

Appendix

Stimulus Lists Used in Experiment 2

| List 1 | List 2 | List 3 |
|--|--|---|
| Money: nickel, dime, penny | Clergy: priest, minister, rabbi | Tools: saw, hammer, drill |
| Nature: mountain, volcano, canyon | Instruments: piano, guitar, trumpet | Furniture: dresser, table, desk |
| Body parts: foot, nose, ear | Weapons: gun, rifle, knife | Cloth: cotton, wool, silk |
| Readings: book, newspaper, magazine | Sports: football, swimming, soccer | Metals: copper, brass, gold |
| Clothing: pants, blouse, jacket | Trees: oak, pine, fir | Transportation: car, truck, train |
| Colors: red, yellow, green | Beverages: milk, water, juice | Countries: France, England, Italy |
| Fruits: pear, apple, banana | Animals: dog, cat, horse | Building parts: roof, floor, ceiling |

Received July 9, 1997

Revision received November 5, 1997

Accepted November 6, 1997 ■