

# Age-related differences in immediate serial recall: Dissociating chunk formation and capacity

MOSHE NAVEH-BENJAMIN, NELSON COWAN, ANGELA KILB, AND ZHIJIAN CHEN  
*University of Missouri, Columbia, Missouri*

We assessed the contribution of two hypothesized mechanisms to impaired memory performance of older adults in an immediate serial recall task: decreased temporary information storage in a capacity-limited mechanism, such as the focus of attention, and a deficit in binding together different components into cohesive chunks. Using a method in which paired associations between words were taught at varying levels to allow an identification of multiword chunks (Cowan, Chen, & Rouder, 2004), we found that older adults recalled considerably fewer chunks and, on average, smaller chunks than did young adults. Their performance was fairly well simulated by dividing attention in younger adults, unlike what has been found for long-term associative learning. Paired-associate knowledge may be used in an implicit manner in serial recall, given that younger adults under divided attention and older adults use it well despite the relatively small chunk capacities displayed by these groups.

The decline in memory with old age (e.g., Craik & Jennings, 1992) seems to affect only some memory functions. Episodic recall is especially affected, a finding for which competing explanations have been suggested (see Light, 1991). In the present study, we will examine the effects of aging on working memory, defined broadly as the limited amount of information kept accessible concurrently for use in cognitive tasks (Cowan, 1999; cf. Baddeley & Logie, 1999). Older adults usually show a deficit in the short-term or immediate memory tasks used to examine working memory, although the size of the deficit depends on the specific task (Kausler & Puckett, 1979; Salthouse & Babcock, 1991; Wingfield, Stine, Lahar, & Aberdeen, 1988).

We will focus on the effect of aging on immediate serial recall. In such tasks, recall is to follow the order in which items were presented. Although Miller (1956) pointed out that young adults can recall lists of about seven items, subsequent work has qualified that statement in various ways. Baddeley, Thomson, and Buchanan (1975) noted that the accuracy of recall depended on the number of items that an individual could recite in about 2 sec, and subsequent work indicated that the recitation rate slowed with aging (Kynette, Kemper, Norman, & Cheung, 1990), in keeping with a cognitive-slowing theory of aging (Salthouse, 1996). Other factors that are critical in serial recall, for which the effects of aging remain uncertain, include (1) a basic capacity limit and (2) a process of chunking. It is age-related changes in these two factors that will be investigated here.

## A Basic Capacity Limit

Broadbent (1975) argued that there is a basic working memory capacity of three separate items. Miller's (1956)

estimate of about seven items was said to depend on the contribution of mnemonic processes that complement the basic capacity. Illustrations of this basic capacity in young adults included the number of items that could be repeated with 0% error and the number of items recalled without hesitation in a single burst from a large category in long-term memory. For example, individuals asked to recall all of the states within the United States do so in bursts of about three items (cf. Graesser & Mandler, 1978). Cowan (2001) obtained a similar estimate of average young adult capacity of between three and five separate items, in a literature review of diverse procedures related to working memory. The effect of aging on this basic capacity in serial recall is unknown, although there is evidence of a decline in capacity with aging, as estimated in a visual array comparison task (Cowan, Naveh-Benjamin, Kilb, & Saults, 2006), and evidence that visual array comparisons and verbal serial recall may depend, in part, on a common, core amodal capacity (Morey & Cowan, 2005).

## Chunking

A key factor that influences immediate recall is chunking. Miller (1956) showed that performance in such tasks depends on how stimuli are associated with one another to produce larger chunks. For example, if the letter series B-B-S-F-B-I-C-I-A is coded as three acronyms of well-known government agencies—BBS, FBI, and CIA—the load on working memory is reduced from nine separate letters to three multiletter chunks. Subsequent work confirmed that long-term associative knowledge is critically important in serial recall tasks (e.g., Ericsson, Chase, & Faloon, 1980; Hulme, Stuart, Brown, & Morin,

---

M. Naveh-Benjamin, navehbenjaminm@missouri.edu

---

2003; Saint-Aubin & Poirier, 1999). Cowan (2001) suggested that the estimate of Miller (1956) could result from the rapid, online grouping of about seven random items to form three or four multiitem chunks, which would fit within the basic capacity limit (Broadbent, 1975; Cowan, 2001). The typical presentation of telephone numbers in groups of three and four digits may serve the purpose of facilitating that online grouping.

If older adults form weaker item-to-item associations than do young adults, that should leave the older adults less able to use these associations to form multi-item chunks in serial recall. Indeed, a binding deficit hypothesis has been suggested. Chalfonte and Johnson (1996) and Mitchell, Johnson, Raye, Mather, and D'Esposito (2000) showed that older adults are particularly deficient in memory that requires the binding of information to contextual elements. Extending this proposal, Naveh-Benjamin (2000) suggested an associative deficit hypothesis (ADH), which focuses on the distinction between memory for single units and memory for associations among units. The basic idea is that complex episodes comprise multiple types of information that are interrelated. Remembering such an episode requires that at least some of the components be retained, along with their relationships to each other. The ADH claims that a major factor leading to poorer episodic memory in older adults is their difficulty in creating and retrieving links between single units of information. Research with independent assessments of item and associative information has supported the ADH and has shown that an associative deficit exists in older adults for both inter- and intraitem relationships (e.g., Bastin & Van der Linden, 2005; Castel & Craik, 2003; Glisky, Rubin, & Davidson, 2001; Light, Patterson, Chung, & Healy, 2004; Naveh-Benjamin, Guez, Kilb, & Reedy, 2004; Naveh-Benjamin, Guez, & Shulman, 2004; Naveh-Benjamin, Hussain, Guez, & Bar-On, 2003). An associative deficit could result in a decrease in the likelihood that the individual would retrieve and use previously learned associations between items in serial recall. It also could result in a decrease in the likelihood that the individual would make use of opportunities for the formation of new associations between items in serial recall. We will examine both of these types of chunking.

The interitem associations that allow chunking are unlikely to be the sole basis of aging effects on immediate serial recall. The ordinary serial recall procedure does not include preexisting interitem associations, and there is evidence suggesting that the online formation of such associations is not a key factor in serial recall in young adults, for lists presented in a regular, ungrouped fashion (Henson, Norris, Page, & Baddeley, 1996). Unlike the usual serial recall procedures, though, in ordinary life one might expect presentations in which the online formation of interword associations is promoted. For example, traffic directions might include an ordered list of street names, each associated with a landmark (such as a nearby store) and a left versus right direction. In the present study, we presented lists of words in pairs, for two reasons. The first was to maximize the use of pretaught associations between items in a pair so that these could be examined,

and the second was to assess the use of new associations formed at the time of list presentation. It was expected, on the basis of the literature, that older adults would be poorer than young adults in both of these bases of interitem associations.

Our procedure was designed to address the need to estimate the contributions of capacity and chunk formation despite serious obstacles in measuring them. One obstacle is that there is little agreement in the field on the question of whether chunks can be reliably observed and, if so, how they can be observed. (For some previous relevant work, see Anderson & Matessa, 1997; Johnson, 1978; Slak, 1970; Tulving & Patkau, 1962.) The factors also may interact. For example, if larger chunks are formed, the potential memory load that the items impose is decreased.

### The Present Empirical Approach

We extended to older adults a procedure (Cowan, Chen, & Rouder, 2004) that can be used to distinguish between capacity, chunk formation, and serial ordering. With young adults, Cowan et al. (2004) presented lists comprising pairs of items that had been taught at various levels of association. Each word had been presented four times in a training session, except for a control set of words that were not included in this phase of the experiment. Each word could be presented four times as a singleton (in the zero-pairing condition), three times as a singleton and once in a pair (in the one-pairing condition), twice as a singleton and twice in consistent pairs (in the two-pairing condition), or four times in consistent pairs (in the four-pairing condition). In an eight-item list, four pairs were presented, all of which were from the same pairing condition or from the control condition. It was found that serial recall increased as a function of the frequency of previous pairing between words. A mathematical model of chunking indicated that the average capacity of recall in chunks was the same in every pairing condition (constant at about 3.5 chunks; a bit lower in the control condition, in which not even the words had been presented in training). There were large differences between pairing conditions in recall, but these came from the fact that higher amounts of pairing resulted in the recall of more two-word chunks and fewer one-word chunks. The results of a cued recall procedure showed that the knowledge of item-to-item associations was long-lasting, including both knowledge from the training session and some additional knowledge acquired during the list presentations.

Any differences between younger and older adults in the use of information about the pairing of items in a list could be due to several factors: differences in the encoding of the associations in training, differences in retrieving these associations, or differences in forming new associations during the list presentation. To help distinguish between these possibilities, we included two groups of young adults, one of which carried out a dual task during list presentation. The latter group had the same pair training as the other participants but should be less likely than the other young adults to notice long-term associations during presentation of the lists in the serial recall task or to form new associations during that presentation.

To the extent that aging deficits in serial recall itself occur as a function of reduced attentional resources in the old (e.g., Craik, 1983, 1986; Rabinowitz, Craik, & Ackerman, 1982), the divided-attention group of young adults should show a pattern of serial recall similar to that for older adults with full attention on the recall task. However, effects on serial recall that stem from age differences in paired training should not be seen in the divided-attention group. The divided-attention group could be deficient in either retrieving associations or forming new ones during list presentation when attention was divided. Any deficit they showed in the use of previously learned associations could reasonably be attributed only to a failure to notice these associations during list presentation. A comparison of the three groups thus may provide valuable evidence about the sources of aging effects.

We were unable to find a mathematical model of performance (cf. Cowan et al., 2004) that could be interpreted unambiguously across our three experimental groups. Instead, we sought model-free estimates of chunk formation. To do so, we were inspired by previous work on free recall in which performance was examined in terms of the number of categories recalled and the number of items recalled within each category (Mandler & Pearlstone, 1966; Tulving & Pearlstone, 1966). Both of these measures appeared to show capacity limits. In our application of these methods to the recall of ordered pairs of words, we distinguished between *pair access* and *pair completion*. Pair access was measured as the proportion of presented word pairs in which at least one item was recalled, whereas pair completion was measured as the conditional proportion of trials in which both items in a word pair were recalled, given that at least one item in the pair was recalled. Pair access theoretically can be influenced by capacity (i.e., the number of chunks recalled), whereas pair completion theoretically can be influenced both by capacity and by the frequency with which associations between items in a pair were formed, resulting in two-word chunks (i.e., the mean size of a chunk). What we are calling a *chunk* here can entail either a forward or a backward association between the items; pairs formed from lists tend to include both directions of association (Kahana & Caplan, 2002).

To examine the effects of serial ordering, we examined the proportion of correct recall, proportion of pairs accessed, and proportion of pairs completed in two ways: using a free-scoring method in which an item is counted as correct no matter where in the list it is recalled, and using a strict-scoring method in which an item is counted as correct only if it is recalled in the correct serial position. The differences between response patterns with those two scoring methods would indicate the contribution of serial-ordering mechanisms.

The procedure also has the potential to yield important information about the interaction of factors. In particular, interword associations reduce the number of independent chunks that must be held in working memory. Theoretically, if a pair of items in a list is encoded as a single chunk, this chunking should free up one slot in working memory that can be used to encode other items in the list. We will be able to check this basic prediction on the basis

of the notion of a capacity limit and to determine whether it is altered by aging and by attention at the time of list presentation.

Aging adults often produce serial position functions different from those produced by young adults (Maylor, Vousden, & Brown, 1999; Murphy, Craik, Li, & Schneider, 2000). Given that evidence from serial position functions has been deemed critically important for an understanding of serial recall (for a review, see Farrell & Lewandowsky, 2004), it will be examined here as well.

Last, after the serial recall procedure, cued recall was examined. Evidence of age differences in cued recall for pairings that had been encountered by the participants only in the previous serial recall procedure can be taken as an indication of age differences in the ability to form new long-term associations from those once-presented pairings. We also will be able to determine how much of a boost in long-term learning occurs in each age group as a result of prior exposure to the pairings within the training phase of the study.

Given that temporal and contextual memory declines with aging (e.g., McCormack, 1982; Naveh-Benjamin, 1990), we considered it possible that a strict scoring of the serial order of responses would pick up age differences that might be overlooked using free scoring, in which credit is given for words recalled in any serial order. However, it will become clear that the results were rather consistent across both scoring methods.

## METHOD

### Participants

Of 90 participants, 60 were younger and 30 were older adults. Younger participants were undergraduates at the University of Missouri receiving course credit. Older participants were recruited from the Columbia, Missouri community. Half of the younger participants received a divided-attention task during the study phase of the list recall task. The mean ages of the three groups, in years, were as follows: younger/full-attention (11 males, 19 females), 18.36 ( $SD = 0.61$ ; range, 18–20); older/full-attention (8 males, 22 females), 71.26 ( $SD = 6.19$ ; range, 60–81); and younger/divided-attention (12 males, 18 females), 18.9 ( $SD = 0.95$ ; range 18–21). The number of years of formal education was, for the younger/full-attention group, 13.48 ( $SD = 1.58$ ); for the older/full-attention group, 13.01 ( $SD = 0.79$ ); and for the younger/divided-attention group, 12.65 ( $SD = 0.99$ ). All the participants reported normal vision and hearing, and indicated being in good health.

### Apparatus, Stimuli, and Procedure

The apparatus, stimuli, and procedure were the same as those in Cowan et al. (2004, Experiment 1), except for the response mode, which in this study was writing (list recall) or speaking (cued recall) instead of typing, given the poor typing skills of some of the elderly participants. Words were visually presented on the computer to 1 participant at a time in a sound-attenuated room. The 198 words were selected from the MRC Psycholinguistic Database (Wilson, 1988). All words were between three and four letters and between three and five phonemes in length. The Kučera–Francis word frequencies were higher than 12, and concreteness ratings were above 500. The words were assigned to three sets, each comprising 66 words. Different word sets were used for three trial blocks. For each participant, 40 words were randomly selected from each set.

Each participant was tested individually and performed all three trial blocks. Each block started with a training phase that was fol-

**Table 1**  
**Experimental Conditions**

Condition	Training Phase	List Recall	Cued Recall
No study	Not included in this phase	One list; e.g., C–D, G–H, A–B, E–F	Same pairs: G–??, etc. (answer, H)
Zero pairing	Eight word singletons, A, B, C, D, E, F, G, and H, four times each (random order)	One list; e.g., C–D, G–H, A–B, E–F	Same pairs: G–??, etc. (answer, H)
One pairing	Word Pairs A–B, C–D, E–F, G–H once, and each word three times as a singleton (random order)	One list maintaining pairs; e.g., C–D, G–H, A–B, E–F	Same pairs: G–??, etc. (answer, H)
Two pairing	Word Pairs A–B, C–D, E–F, and G–H twice, and each word twice as a singleton (random order)	One list maintaining pairs; e.g., C–D, G–H, A–B, E–F	Same pairs: G–??, etc. (answer, H)
Four pairing	Word Pairs A–B, C–D, E–F, and G–H four times each (random order)	One list maintaining pairs; e.g., C–D, G–H, A–B, E–F	Same pairs: G–??, etc. (answer, H)

Note—The letters A–H represent printed words. Within a participant, different words appeared in each training condition. Presentations of words or pairs from one condition were usually separated by other words and pairs. Specific pairing, but not the order of the pairs, was maintained throughout. Adapted from Cowan, Chen, and Rouders (2004).

lowed by five list recall trials and a delayed cued-recall task. The phases of the experiment are illustrated in Table 1 and will be explained further below.

**Training phase.** The manipulation used was the presentation method during the training phase. In this phase of the trial block, words were presented either singly or in pairs for 2 sec per presentation, to be pronounced aloud by the participant. The total number of single or paired presentations for each word was four, with all presentations randomized. Each word was assigned to the zero-, one-, two-, or four-pairing condition. This determined the number of times the words were presented in consistent pairs. Words in the four-pairing condition appeared only paired, whereas words in the one- and two-pairing conditions were paired only one and two times out of the four exposures, respectively. Words in the zero-pairing condition appeared only singly. The lag between consecutive appearances of the single words or word pairs varied from 1 to 40. In addition, there were words assigned to the no-study condition, which did not appear in the training phase but were used in the serial recall phase. When presented in a pair, words were paired in the same way, including the same order, throughout the experiment.

**List recall phase.** During the list recall phase of each trial block, which followed the training phase, five lists were presented, each for immediate serial recall. Each list comprised eight words (four pairs) drawn from the same training condition. The order of the lists and of the pairs within lists was randomized, but the same pairings as those used in the training phase were kept. Each trial started once the participant pressed a designated key. After a 1-sec delay, the four pairs were presented one at a time for 2 sec each. The recall test started after the last pair had disappeared from the screen. The participant wrote down as many of the eight presented words as he or she could remember in their order of appearance on a designated response page, which included a matrix with four two-line rows. The participants were told to recall the words in order and to leave a space if they did not remember a given word. The experimenter later keyed the participants' responses into the computer.

The young participants in the divided-attention group performed a secondary task during the presentation of the word pairs in this phase. This secondary task was a continuous three-choice reaction time task that involved a sequential presentation of auditory tones by the computer, one at a time, and a manual response on a computer keyboard to each tone. One of three tones (with fundamental frequencies of 100, 800, or 1000 Hz) was presented for 300 msec at random, and the participants' task was to press a predesignated cor-

responding key on the keyboard. A response caused the immediate presentation of one of the other two tones at random. The participants in the divided-attention group were told to pay equal attention to memorizing the words and performing the secondary digit task.

**Cued-recall phase.** During the cued recall phase of each trial block, the leftmost word in a pair appeared on the screen, and the participant was to recall orally the word that had appeared with it during the serial recall phase (and, for all but the no-study and zero-pairing conditions, during the training phase also).

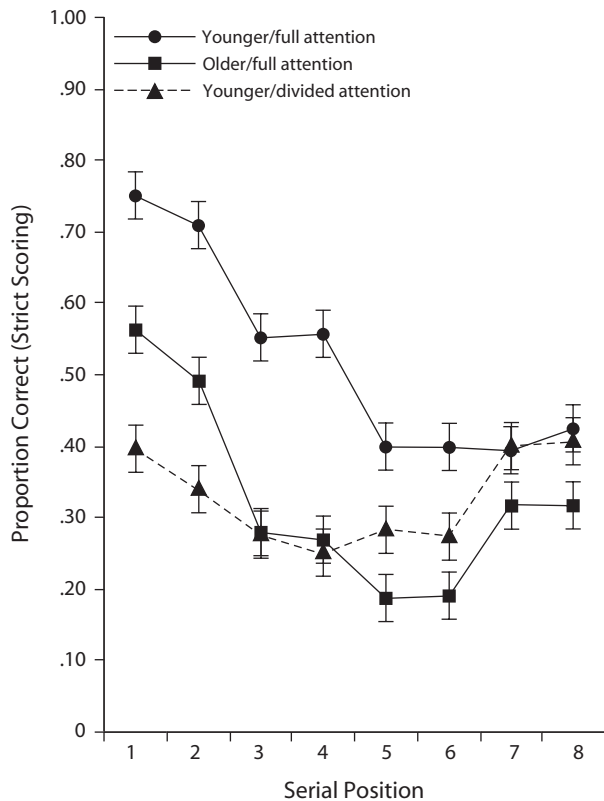
## RESULTS

The pattern of results for the three trial blocks was similar in all three groups, so, in the analyses to be reported, the data have been collapsed across blocks for simplicity. Reporting of results for the two test procedures, list and cued recall, will appear in turn.

### List Recall

The results will be presented here for the overall proportion correct and then for the theoretically derived measures, pair access and pair completion. In each case, both the free-scoring and the strict serial order scoring results will be discussed for the sake of comparison. These are followed by theoretical analyses of serial ordering and of the mean number and size of chunks recalled. For the strict serial scoring method, a correct response was counted if a given word was reported as being in exactly the same position in which it had been presented during the study phase of the list. For the free-scoring method, a correct response was counted if a word from the study list was reported in the recall phase, regardless of its recall position.

**Proportion correct.** An ANOVA of the proportion correct in serial recall, using a strict serial order scoring, included all three participant groups and within-participants factors of condition (control and zero-, one-, two-, and four-pairing training) and serial position within the list (one to eight). Performance was highest in the

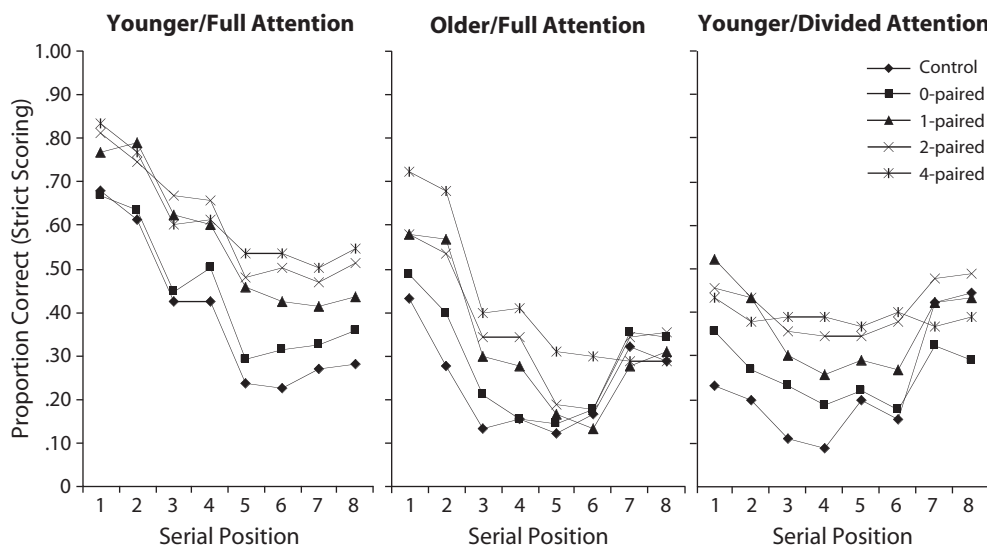


**Figure 1.** Proportion correct for each group in each serial position according to strict serial scoring, collapsed across conditions. Error bars are standard errors.

younger/full-attention group ( $M = .52$ ,  $SEM = .02$ ), and was lower in the older/full-attention and younger/divided-attention groups (in both groups,  $M = .33$ ,  $SEM = .02$ ),

so that the main effect of group was significant [ $F(2,87) = 20.85$ ,  $MS_e = 0.72$ ,  $p < .01$ ]. There was a steady increase in performance across conditions [ $F(4,348) = 41.03$ ,  $MS_e = 0.12$ ,  $p < .01$ ]. Performance levels in the control and zero- through four-pairing conditions were .29, .33, .42, .46, and .48, respectively ( $SEM = .02$  in each case).

Figures 1 and 2 show that there were some interesting differences between the serial position functions in different groups. There was a main effect of serial position [ $F(7,609) = 36.52$ ,  $MS_e = 0.12$ ,  $p < .01$ ] and a condition  $\times$  serial position interaction [ $F(28,2436) = 2.12$ ,  $MS_e = 0.05$ ,  $p < .01$ ]. More important, these effects were modified by group  $\times$  serial position [ $F(14,609) = 7.98$ ,  $MS_e = 0.12$ ; Figure 1] and group  $\times$  condition  $\times$  serial position [ $F(56,2436) = 1.46$ ,  $MS_e = 0.05$ ,  $p < .05$ ; Figure 2] interactions. The results show that the older adults were, in some ways similar to those in the younger/divided-attention group and in some ways different from them. Their overall levels of performance were quite similar and much lower than those for the younger/full-attention group. Also, Figure 2 shows that whereas the younger/full-attention participants benefited from word pair training throughout all the serial positions, the participants in the older/full-attention and younger/divided-attention groups benefited much in Serial Positions 1–6 but showed little or no benefit in the final two serial positions. However, both figures also show that the effect of dividing attention was to decrease the primacy effect (early serial position performance) and increase the recency effect (late serial position performance, relative to earlier positions) in comparison with the other two groups. The basis of these differences is not entirely clear and will be discussed later, although it should be emphasized that the divided-attention group could use full attention except during the list presentation phase of the trial, whereas any loss of processing ability due to aging presumably was not similarly restricted to the



**Figure 2.** Proportion correct in each group in each serial position according to strict serial scoring, separately for each condition (graph parameter). Left, younger/full-attention group; center, older/full-attention group; and right, younger/divided-attention group.

list presentation phase. Finally, note that the pattern for the control condition (which included words that had not appeared in the training phase) indicates that both the older adults and the younger/divided-attention group showed a lowered primacy and enhanced recency effect, relative to the younger/full-attention group.

The mean proportions correct, using free scoring, were .64 for the younger/full-attention condition, .46 for the older/full-attention condition, and .47 for the younger/divided-attention condition ( $SEM = .02$  in each case). The statistical pattern of results was the same as that for strict serial order scoring, except that the three-way group  $\times$  condition  $\times$  serial position interaction did not approach significance ( $p = .11$ ).

**Pair access.** Access to a pair of words within a list was judged to have occurred if either or both of the words in the presented pair were recalled. An ANOVA included all groups and included the training condition within participants. Results of the analysis using strict serial order scoring are illustrated in Figure 3. The analysis produced a main effect of condition [ $F(4,348) = 23.45$ ,  $MS_e = 0.02$ ,  $p < .01$ ]. The proportion of pairs accessed increased steadily across the control (.37), zero-pairing (.40), one-pairing (.46), two-pairing (.50), and four-pairing (.52) conditions. Post hoc Newman-Keuls tests indicated that both the control and the zero-pairing conditions differed from the remaining three conditions and that the one-pairing and four-pairing conditions also differed from one another. These findings are theoretically important inasmuch as the paired training pertained to the association between items within a pair, not associations between pairs. The

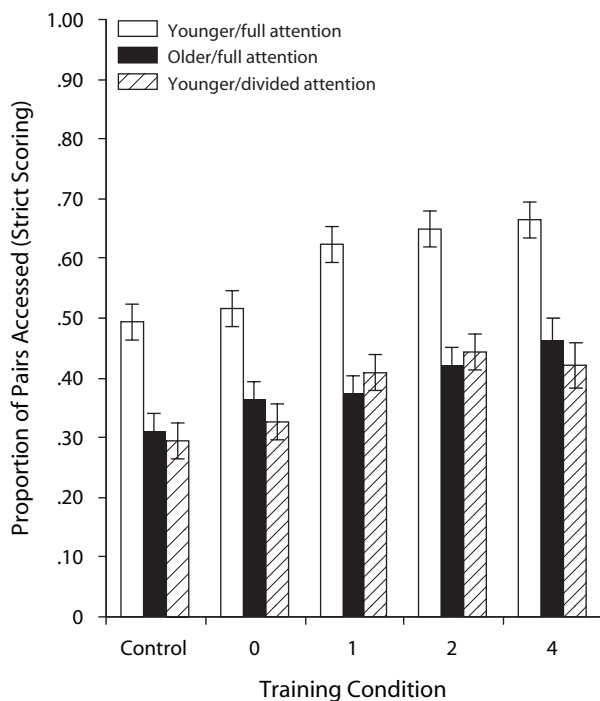


Figure 3. Proportion of pairs in a list accessed in each group, according to strict serial scoring, separately for each condition. Error bars are standard errors.

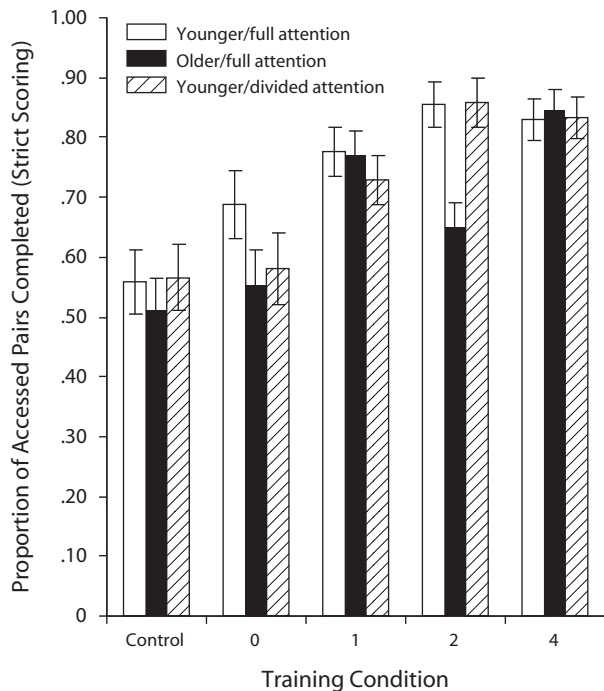
only way that access to pairs could be affected by paired training is if the formation of two-item chunks relieved the participant from recalling those two items separately and, therefore, freed up capacity, allowing access to additional pairs that otherwise would have been forgotten.

There was a main effect of group [ $F(2,87) = 21.07$ ,  $MS_e = 0.10$ ,  $p < .01$ ]. The proportion of pair access was .59 in the younger/full-attention group, as compared with .39 in the older/full-attention group and .38 in the younger/divided-attention group ( $SEM = .03$  in each case). Apparently, it is possible to simulate the effect of aging on the access to pairs by dividing attention during list presentation. However, there was no hint of a group  $\times$  condition interaction ( $p = .48$ ). The access to pairs depended on the participant characteristics to the same extent for all the training conditions.

An ANOVA with pair access, using free scoring, yielded the same statistical effects as the ANOVA strict scoring. There was again a main effect of condition [ $F(4,348) = 42.83$ ,  $MS_e = 0.01$ ,  $p < .01$ ]. Across pairing conditions (control and zero- to four-pairing), the proportion of pairs accessed was .50, .57, .65, .67, and .70, respectively. Newman-Keuls tests showed that all the conditions differed from one another, except for the two- and four-pairing conditions. Again, this strengthens the notion that original pair learning freed up capacity. There was again a main effect of group [ $F(2,87) = 23.53$ ,  $MS_e = 0.07$ ,  $p < .01$ ]. Access to the pairs was highest in the younger/full-attention group ( $M = .74$ ) and much lower in the other two groups (older/full-attention,  $M = .55$ ; younger/divided-attention,  $M = .56$ ;  $SEM = .02$  in each case). Finally, there was again no indication of a group  $\times$  condition interaction ( $p = .59$ ). In order to understand these results fully, they have to be examined in the context of pair completion.

**Pair completion.** Pair completion was measured as the conditional proportion of pairs for which both items were recalled, given that at least one of the two items in a pair was recalled (i.e., given successful pair access). Figure 4 shows the results, using strict serial order scoring. The participants who did not access at least one pair in every training condition did not have a complete data set and had to be omitted from this analysis, which included  $n = 29$  for the young/full-attention group and  $n = 27$  for each of the other two groups (out of 30). There was a strong effect of condition [ $F(4,320) = 23.17$ ,  $MS_e = 0.06$ ,  $p < .01$ ]. For the control and zero- through four-pairing conditions, the proportions were .54, .61, .76, .79, and .84, respectively. Newman-Keuls tests showed that the control and zero-pairing conditions differed from all of the other conditions but that no other conditions differed. Thus, it took only a single exposure to a pairing in the training condition for a large benefit of pairing to occur, as measured in this manner. There was only a marginal effect of group [ $F(2,80) = 2.65$ ,  $MS_e = 0.08$ ,  $p = .08$ ].

Given that most of the missing data points were in the nonstudied word control condition, another analysis was conducted using only the zero- through four-pairing conditions. There were still missing cells for 1 participant in the young/full-attention condition and 2 in the young/divided-



**Figure 4.** Proportion of accessed pairs in a list, in each group, that were completed, according to strict serial scoring, separately for each condition. Error bars are standard errors.

attention condition. Each of these was filled with expected means obtained by multiplying the participant average by the condition average and dividing by the grand average. The effect of group was now significant [ $F(2,87) = 8.74$ ,  $MS_e = 0.09$ ,  $p < .01$ ]. Newman-Keuls tests indicated that pair completion in the younger/full-attention group ( $M = .76$ ) exceeded that in both the older group and the younger/divided-attention group ( $M = .62$  in both cases). The group effect did not interact with pairing condition [ $F(6,261) = 1.64$ ,  $MS_e = 0.05$ , n.s.].

Once more, an analysis based on free scoring yielded similar effects. In this analysis, it was possible to include  $n = 30$  in the younger/full-attention condition,  $n = 29$  in the older/full-attention condition, and  $n = 30$  in the younger/divided-attention condition. This analysis again yielded a main effect of training condition [ $F(4,344) = 37.33$ ,  $MS_e = 0.04$ ,  $p < .01$ ]. For the control and zero-through four-pairing conditions, the proportions were .52, .59, .72, .75, and .83, respectively. Newman-Keuls tests showed that all the proportions differed from one another, except for those for the one- and two-pairing conditions. Once more, there was a marginal effect of group [ $F(2,86) = 2.58$ ,  $MS_e = 0.08$ ,  $p = .08$ ]. An analysis using only the zero- through four-pairing conditions included a complete data set and showed a significant effect of group [ $F(2,87) = 5.62$ ,  $MS_e = 0.07$ ,  $p < .01$ ]. Newman-Keuls tests showed a significant advantage for the younger/full-attention group ( $M = .73$ ) over both the older group ( $M = .62$ ) and the younger/divided-attention group ( $M = .65$ ), which did not differ significantly from

one another. The group effect did not interact with pairing condition [ $F(6,261) = 1.52$ ,  $MS_e = 0.04$ , n.s.].

**Serial ordering.** Throughout our analyses, the observed performance levels were diminished when the participants were made responsible for the exact serial order of responses. Yet similar statistical effects related to aging emerged from both strict serial order and free scoring of the results. This similarity seems to suggest that although older adults have trouble forming interitem associations (Naveh-Benjamin, 2000), serial ordering in immediate recall falls in a different domain of skills and is not selectively impaired with aging. This finding warrants further basic research on how serial ordering in immediate recall differs from associative learning.

**Capacity and chunk size analyses.** A primary research question is whether the capacity of working memory changes with age and whether any change with age can be simulated by dividing attention at the time of the list presentation. Any estimate of capacity depends on some tentative assumptions. Suppose that paired associations acquired before the list recall task are useful for decreasing the working memory load imposed by the list, whereas paired associations formed at the time of list recall do not reduce the memory load. This supposition seems sensible because any pair learning that occurs online during the list presentation must require working memory capacity. One can then estimate the benefit from pair learning in the zero-, one-, two-, and four-pairing conditions as the pair completion score in each of those conditions minus the pair completion score in the zero-pairing condition. If  $A_i$  is the proportion of pair access in the  $i$ -pairing condition and  $C_i$  is the conditional proportion of pair completion in that condition given pair access, the estimate of working memory capacity in chunks in that condition ( $K_i$ ) is

$$K_i = 4 * [A_i * (1 - C_i) + A_i * (C_i - C_0) + (2 * A_i * C_0)]. \quad (1)$$

The rationale for Equation 1 is that, of the four pairs in a list, each of the pairs for which only a singleton was recalled [ $4 * A_i * (1 - C_i)$ ] counts as a single chunk; each of the pairs for which both items were recalled as a result of paired-associate training [ $4 * A_i * (C_i - C_0)$ ] also counts as a single chunk; and each of the pairs for which both items were recalled, but not as a result of paired-associate training ( $4 * A_i * C_0$ ) counts as two chunks and is multiplied by 2 in Equation 1.

An alternative analysis relies on the assumption that pairs recalled in the zero-pairing condition are, in fact, two-word chunks that have formed online, conserving capacity in the process. The corresponding alternative formula in this case replaces Equation 1's ( $2 * A_i * C_0$ ) with ( $A_i * C_0$ ), and the resulting equation simplifies to

$$K_i = 4 * A_i. \quad (2)$$

On the basis of either formula, it is possible to obtain an estimate of the average size of the chunk in Condition  $i$ ,  $S_i$ , expressed in items per chunk:

$$S_i = N_i / K_i, \quad (3)$$

**Table 2**  
**Mean Estimates of Capacity and Chunk Size**

Scoring Method	Chunking Assumption	Age/Attention Group		
		Younger/Full	Older/Full	Younger/Divided
Capacity Estimate ( $K_i$ )				
Strict	No new chunks <sup>1</sup>	4.10	2.47	2.40
	New chunks <sup>2</sup>	2.45	1.62	1.60
Free	No new chunks <sup>1</sup>	4.92	3.39	3.38
	New chunks <sup>2</sup>	3.08	2.33	2.35
Chunk Size ( $S_i$ )				
Strict	No new chunks <sup>1,3</sup>	1.12	1.16	1.26
	New chunks <sup>2,3</sup>	1.82	1.70	1.75
Free	No new chunks <sup>1,3</sup>	1.13	1.18	1.25
	New chunks <sup>2,3</sup>	1.77	1.66	1.71

Note—These estimates are averaged across the zero-, one-, two-, and four-pairing conditions. Strict-scoring estimates are based on data with several missing cells filled as in the analysis of pair completion scores. <sup>1</sup> According to Equation 1 in text. <sup>2</sup> According to Equation 2 in text. <sup>3</sup> According to Equation 3 in text.

where  $N_i$  is the mean number of items correct. The estimate depends on which capacity formula is used and also on whether free or serial scoring of recall is used.

In all likelihood, the truth lies somewhere between the assumptions embodied in Equations 1 and 2. New chunks may be formed sometimes, whereas two items from a pair may be recalled as separate chunks sometimes. However, the mathematical model of Cowan et al. (2004) suggests that new chunks usually do form. We will see also that the data from the present study looks more orderly with that assumption (i.e., Equation 2).

The estimates according to various assumptions, collapsed across the zero- through four-pairing conditions, are shown in Table 2. According to all analyses, there was a difference between groups in capacity: with the assumption of no new chunks (Equation 1),  $F(2,87) = 18.04$  with strict scoring and 19.79 with free scoring ( $MS_e = 6.15$  and 4.77, respectively;  $p < .01$  in both cases), and with the assumption of new chunks (Equation 2),  $F(2,87) = 19.21$  with strict scoring and 23.09 with free scoring ( $MS_e = 1.49$  and 0.95, respectively;  $p < .01$  in both cases). In all cases, Newman-Keuls analyses indicated that the younger/full-attention group had a higher capacity than did the other two groups, which did not differ from one another. In no case did the interaction of group with pairing condition approach significance.

The finding of group differences in chunk size was dependent on the assumptions, however. With the assumption of no new chunks (Equation 1), using free scoring, the group effect was marginally significant [ $F(2,87) = 2.86$ ,  $MS_e = 0.15$ ,  $p = .06$ ]. With strict scoring, the effect did reach significance [ $F(2,87) = 3.13$ ,  $MS_e = 0.19$ ,  $p < .05$ ]. However, in both cases, Newman-Keuls tests indicate that the younger/divided-attention condition produced larger chunk sizes than did the younger/full-attention condition (see Table 2), with the older group intermediate. There is no theoretical reason to believe that dividing attention should increase chunks size, and therefore, these results cast further doubt on the validity of Equation 1 (cf. Cowan et al., 2004), which determines what gets counted as a single chunk.

The finding was more reasonable under the assumption that new chunks can form online (Equation 2; cf. Cowan et al., 2004). With free scoring of the results, this produced a main effect of group [ $F(2,87) = 3.16$ ,  $MS_e = 0.10$ ,  $p < .05$ ]. Newman-Keuls tests indicated that the chunk size of the older adults was significantly below that of those in the younger/full-attention group. There also was an effect of the pairing condition, with increasing chunk size across conditions [for zero- through four-pairing trials,  $M = 1.58$ , 1.72, 1.75, and 1.82 items/chunk, respectively;  $F(3,261) = 21.69$ ,  $MS_e = 0.04$ ,  $p < .01$ ]. Newman-Keuls analyses indicated that all pairs of conditions differed significantly, except the one- and two-pairing conditions.

With the strict scoring of results, again there was a main effect of group [ $F(2,87) = 3.43$ ,  $MS_e = 0.12$ ,  $p < .05$ ]. As with free scoring, a Newman-Keuls analysis indicated that the chunk size of the older adults was significantly below that of the younger/full-attention group. Again there was an effect of pairing condition, with increasing chunk size across conditions [for zero- through four-pairing trials,  $M = 1.61$ , 1.79, 1.79, and 1.83 items/chunk, respectively;  $F(3,261) = 15.90$ ,  $MS_e = 0.06$ ,  $p < .01$ ]. Newman-Keuls analyses indicated that the zero-pairing condition produced a significantly smaller chunk size than did any of the other conditions, which did not differ.

What distinguished the last analysis from the others was a group  $\times$  pairing condition interaction [ $F(6,261) = 2.27$ ,  $MS_e = 0.06$ ,  $p < .05$ ]. To investigate this interaction, separate ANOVAs were carried out for each pairing condition, and it was found that the group effect was significant only in the two-pairing condition [ $F(2,87) = 6.53$ ,  $MS_e = 0.06$ ,  $p < .01$ ]. Newman-Keuls tests showed that the chunk size was smaller in the older adults ( $M = 1.66$  items/chunk) than in either of the other two groups, which did not differ (younger/full-attention,  $M = 1.87$ ; younger/divided-attention,  $M = 1.85$ ). This result is entirely parallel to that for pair completion shown in Figure 4 and probably results from the same aspects of the data.

In sum, the conclusion that chunk capacity is lower in older adults than in younger adults under full attention is

firm, regardless of the assumptions about whether online chunks are formed, and it also seems clear that the effect of aging on chunk capacity is well simulated by dividing attention at the time of list presentation in young adults. The same pattern of results was obtained for both strict serial scoring of the data and free scoring. The conclusions about average chunk size (the use of information from paired associations) are more situation dependent. The assumption that no chunks were formed online in the zero-pairing condition led to the nonsensical result that dividing attention raised chunk size, as compared with full attention. With the assumption that new chunks could form online, the finding was that aging reduces chunk size. The mean chunk size in younger adults with divided attention was intermediate and was not significantly different from that in the other two groups in most instances.

### Cued Recall

The cued recall results are shown in Figure 5. One can see a dramatic increase in cued recall across training conditions. An analysis with group and condition as factors produced a main effect of condition [ $F(4,348) = 245.87$ ,  $MS_e = 0.02$ ,  $p < .01$ ]. Across the control and zero- through four-pairing conditions, the mean proportions correct were .11, .11, .37, .50, and .61, respectively ( $SEM = .01$ , .01, .02, .03, and .02). Newman-Keuls tests indicated that all the values were different from one another, except the control and zero-pairing values.

As Figure 5 also shows, there were important group differences. These resulted in a main effect of group

[ $F(2,87) = 5.36$ ,  $MS_e = 0.10$ ,  $p < .01$ ], as well as a group  $\times$  condition interaction [ $F(8,348) = 2.89$ ,  $MS_e = 0.02$ ,  $p < .01$ ]. The figure suggests a pattern in which, in the absence of paired training (control and zero-pairing conditions), the younger/full-attention group outperformed the other two groups. That is to be expected if the younger/full-attention group was able to pick up the pairing from the list presentation better than could the other two groups. With more paired presentations in training, however, the younger/divided-attention group (which had the same full-attention training session as the other groups) came to look more like the younger/full-attention group. This also is expected on the supposition that, with more pairings during training, the participant becomes less reliant on pairings that occurred during list recall to assist performance in the subsequent cued recall task. Consequently, older adults fell behind the other groups by the two-pairing condition.

These suggestions, based on Figure 5, were statistically supported by simple-effect one-way ANOVAs on individual conditions. The effect of group was significant in all the conditions [control,  $F(2,87) = 9.37$ ,  $MS_e = 0.01$ ,  $p < .01$ ; zero-pairing,  $F(2,87) = 5.78$ ,  $MS_e = 0.02$ ,  $p < .01$ ; one-pairing,  $F(2,87) = 3.86$ ,  $MS_e = 0.04$ ,  $p < .05$ ; two-pairing,  $F(2,87) = 3.88$ ,  $MS_e = 0.06$ ,  $p < .05$ ; and four-pairing,  $F(2,87) = 3.27$ ,  $MS_e = 0.05$ ,  $p < .05$ ]. However, Newman-Keuls tests indicated that the basis of this effect changed across conditions. In both the control and the zero-pairing conditions, the younger/full-attention group differed from both of the other groups. In the one-pairing condition, however, only the younger and the older/full-attention groups differed from one another, with the younger/divided-attention group intermediate. In the two-pairing condition, it was the older adults who differed from the other two groups, which did not differ from one another. Similarly, in the four-pairing condition, only the younger/divided-attention and the older/full-attention groups differed from one another.

Last, the cued recall results can be used to assess more closely whether capacity changes with age. The conclusion that chunk capacity is smaller in older adults might be brought into question if their recall of fewer pairs was somehow an indirect consequence of having less information about associations between words in a pair (as the ADH predicted). Although we do not have a direct measure of learning during the training phase, we compared the performance of the 23 older adults who had the highest cued recall scores with that of the 23 younger adults who had the lowest such scores. Cued recall was exactly matched between these groups ( $M = .34$  in both cases). The results in this subsample indicate that all of the effects of *pair access* (a measure of chunk capacity) remain. In contrast, the statistical effects of *pair completion* are wiped out by this matching. This makes sense, since cued recall and pair completion should reflect similar knowledge base related to the amount of training. This analysis strongly suggests that the older adults' lower chunk capacity was not due to poorer training.

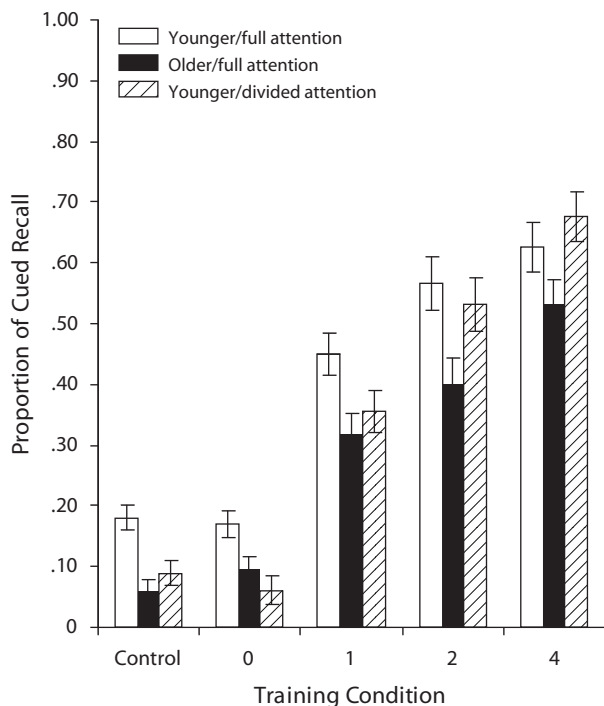


Figure 5. Proportion of correct cued recall in each group, separately for each condition. Error bars are standard errors.

## DISCUSSION

The present study provides some important clues toward understanding the effects of aging on capacity limits in serial recall, given multiple sources of influence, including associations between items based on their pairing in a training session and in the lists to be recalled. Key points in the results are as follows.

1. The level of performance was higher in the younger/full-attention group than in the older/full-attention group or in a group of younger adults whose attention was divided during list presentation. This divided-attention manipulation provided a reasonably good analogue to the aging process.

2. In the young/full-attention group, the benefit of word pairings familiar from training occurred throughout all serial positions in list recall. In contrast, for the other two groups, the benefit was observed primarily for the first six of eight serial positions (Figure 2). Figure 2 also shows that, despite the absence of a difference between pairing conditions in the last few serial positions for older adults, these participants still showed better performance in the last few positions than in previous positions. The explanation for this could be that there are multiple sources of information contributing to performance (cf. Baddeley, 2001; Baddeley & Logie, 1999). In particular, participants might maintain both abstract and phonological codes for words in a list, which are used in combination (Chen & Cowan, 2005). Learned associations should improve the abstract code, but only if the associations can be retrieved from long-term memory and maintained in a capacity-limited store (i.e., kept in mind), in which case they can be used in short-term serial recall. If the abstract code was impaired for the last few items without similar impairment in the phonological code, then for those items, there might be little or no role for associative information from learned pairs.

Theoretically, the aforementioned deficit in older adults (absence of training effects at the end of the list) could occur because of deficits in the use of associative information either during the presentation of the list or during the recall period. The use of associative information during presentation of the list would help to maintain a learned pair in working memory in its entirety, whereas the use of that information during the recall period would help to restore a word from long-term memory, given retention of the other pair member in working memory. One can find potential reasons for an aging deficit in the use of associative information during presentation of the list (from a lowered capacity limit, along with the increasing memory load as the list progresses) or during recall (from a vulnerability to output interference, which increases as recall of the list progresses; see Cowan, Saults, Elliott, & Moreno, 2002).

Results from dividing attention in young adults can help to address the basis of the aging deficit. The younger/divided-attention participants, like the older adults, failed to use associative information for the last pair only. Given that the divided-attention manipulation occurred only during list presentation, the deficit in both the aging and the younger/divided-attention groups in the use of associative

information at the end of the list can be attributed to some sort of encoding failure. An attention-demanding type of rehearsal might be needed to reinstate the grouping, from the associative knowledge, during list presentation. When that reinstatement failed to occur, the participants may not have used the grouping during recall. Although we know of no relevant data to support this interpretation of the aging deficit in list recall, subjective reports of young adults, but not second-grade children, suggest that they remember lists by grouping items together at the time of their presentation (Cowan et al., 2006). Aging participants under a large memory load also may fail to use grouping information.

A remaining question is why an individual who had learned a word pair during training still might not use the pairing information within list recall, as we are suggesting for the last pair of items in the elderly and younger/divided-attention participants. If the participants suspected that the pairing within the lists did not always match the pairing within training, they might have avoided using the pairing information, unless they were able to confirm it by attending to it during list presentation. That might not have occurred for the last pair in many lists. Alternatively, the demands of list recall might not have left enough attention to reinstate the last pair from long-term memory. In either case, performance was fairly good for that last pair in all the training conditions (probably because of available phonological information), so it might not make sense to rely on uncertain pairing information at that point in the recall of the list.

3. One difference between the younger/divided-attention and the older/full-attention groups is in the overall serial position curve (Figure 1). The two full-attention groups displayed serial position curves with large primacy and smaller recency effects, whereas dividing attention appeared to lower performance in the earlier part of the list more than it lowered performance in the latter part of the list. An explanation for this pattern of results is that attention was not divided during output. When a list ended, the participants in the younger/divided-attention group might have been able to turn attention to a lower level phonological or visual memory of the last few items and carry out the processing that they could not do during the list. Such retroactive processing may not be possible for items presented earlier in the list. This retroactive processing apparently does not include associative information and may involve simple phonological rehearsal of those items near the end of the list (Baddeley & Logie, 1999).

4. The access to pairs within the list, judged as the recall of at least one member of a pair, increased markedly across training conditions in all the groups. The natural interpretation of this finding appears to be that strengthened paired associations allowed the participants to use less capacity to retrieve two items in a pair and freed up some of this capacity to allow access to more pairs.

5. The access to pairs within a list was far superior in younger/full-attention participants than in both of the other groups, which did not differ from one another in this regard. There is still a great deal that is not known about this effect. We cannot examine response latencies, given

our use of handwritten responses, but they could carry important information. Inspection of serial positions curves for pair access and pair completion (which we have omitted from the results, for the sake of simplicity) showed that the large primacy effect in the older adults could be attributed mostly to pair access, and not to pair completion. Therefore, it is possible that the older adults delayed the recall of later pairs while trying to complete earlier pairs. In the young adults under full attention, in contrast, both pair access and pair completion were highest in the early list positions, suggesting that responding may have been faster. Investigation of response latencies should be a high priority for further research.

6. Notably, the group differences in pair access did not change across training conditions (see Figure 3). If paired associate training related directly to pair access in serial recall, one would have expected that those individuals who were more completely trained (the young adults) would show larger benefits across training conditions in serial recall than would the less well trained individuals (the older adults), because they would have more learned pairs to apply. However, that does not appear to be the mechanism of pair access. Instead, it appears that the ability to attend to the list (younger/full-attention group) afforded a fixed advantage in pair access over the other two groups. Thus, in a comparison of the younger/divided-attention group, which presumably learned associations in training better than did the older/full-attention participants (given that attention was undivided during training), both of these groups showed the same amount of benefit of pair training on pair access in the lists (see Figure 3). Perhaps the contribution of paired associate training to pair access occurs from retrieval of paired associations on an implicit level of memory at which the older adults are not impaired, relative to young adults. This account will be explained further below.

7. The probability of completing a pair, conditional upon access to it, was selectively impaired in older adults and in younger adults under divided attention. This is what one would expect if pair completion depends on prior paired associate learning, as could occur if this learning allows pairs of items to be combined into two-word chunks in serial recall.

8. For both pair access and pair completion, the same statistical effects emerged, no matter whether a strict serial order scoring or a free scoring was used. One explanation of this finding is that, across conditions, the items recalled in the correct serial position were a relatively fixed percentage of the items recalled anywhere in the list (control, 76%; zero-pairing, 73%; one-pairing, 75%; two-pairing, 78%; and four-pairing, 75%). This suggests that pair training and chunk capacity had little to do with the correct serial order of recall, which, instead, might depend primarily on phonological information (cf. Chen & Cowan, 2005).

9. Finally, a theoretical analysis of chunk capacity and chunk size further supported the finding that most of the aging difference in serial recall could be modeled by dividing attention at the time of the list presentation and that the difference emerged primarily as a decrease in the number

of chunks recalled and, secondarily, as a decrease in the chunk size. For the number of chunks recalled, the aging effect was clearly simulated by dividing attention in younger adults; for the chunk size, this was less clearly so.

The results seem consistent with those in Taub (1974), which showed age-related differences in chunking, especially when the materials allowed the use of organizational schemes. Although Allen and Coyne (1988, 1989) and Allen and Crozier (1992) found no difference in chunk size across groups of younger and older adults, those results were obtained with lists with only three to six items, which is probably not taxing enough to elicit the use of maximum-sized chunks.

The present findings on aging effects in immediate serial recall are in line with previous findings on aging effects in long-term learning. Research supporting the ADH attributes older adults' deficient long-term episodic memory to their difficulty in creating and retrieving associations (Li, Naveh-Benjamin, & Lindenberger, 2005; Naveh-Benjamin, 2000; Naveh-Benjamin, Guez et al., 2004, Naveh-Benjamin, Guez, & Shulman, 2004; Naveh-Benjamin et al., 2003). One might have expected this finding to generalize to the creation of interitem associations in serial recall, and the present results indicate this pattern. One interpretation of these results is that older adults' problems in encoding and retrieval of long-term associations, noted in the above-mentioned studies, start early in the processing stream, in short-term working memory activity, and extend to their ability to retrieve associations after longer periods of time (e.g., in the cued recall task in the present experiment).

One aspect of the data does not seem compatible with the evidence that older adults are deficient in encoding paired associations. In particular, we would have expected them to benefit to a lesser degree than did younger ones from the original paired-associate training in their performance in the serial recall and cued recall tasks. However, in neither serial recall nor cued recall did we find a pattern of results in which the benefit of paired training increased more across pairing conditions in younger adults than in older adults. All the groups appeared to have increased benefits across pairing conditions to a similar degree (see Figures 3–5). One possible explanation for this result is that paired associations created during the training phase may have been partially used as implicit knowledge in our immediate serial recall and cued recall tasks. Older adults are not particularly deficient in implicit knowledge (Hay & Jacoby, 1999; Jennings & Jacoby, 1993, 1997; Light, 2000; Yonelinas, 1997), which also is relatively unimpaired in the case of divided attention (e.g., Jacoby, 1991). Dividing attention and age may remove the use of explicit knowledge involved in the creation of associations, but not implicit use of previously learned associations.

Overall, older adults seem to have a problem in the explicit encoding and retrieval of new associations both in short- and in long-term memory. For example, the advantage of the young/full-attention participants over the older participants in serial recall for the zero-pairing condition, in which the first pairing occurred within the list, was 0.61 standard deviation units (using the  $N = 30$  data set with

free scoring, which we deemed most appropriate). This is quite similar to the effect size for this condition in the subsequent cued recall test, which was 0.53 standard deviation units. However, the reliance on previously learned word-to-word associations on an implicit, rather than an explicit, level seems to be relatively intact, and this could explain why all three groups were able to benefit about equally from previously learned associations in immediate serial recall and in the cued recall tasks.

With respect to the effects of divided attention in younger adults, research suggests that dividing attention may impair some aspects of learning (such as elaboration and consolidation processes—*Craik & Kester, 2000; Naveh-Benjamin, Craik, Gavrilescu, & Anderson, 2000*) but does not selectively impair item-to-item associations (*Naveh-Benjamin, 2002; Naveh-Benjamin, Guez, & Marom, 2003*) and, therefore, does not adequately simulate the effects of aging on memory. However, the results of the present study show that the prevention of mnemonic strategies by dividing attention in young adults may provide a reasonable analogue to the effects of aging on immediate serial recall, as indicated by the similarity between the patterns of responses in the older/full-attention and the younger/divided-attention groups. To be sure, there were subtle differences between the exact patterns of responses across serial positions in these two groups, but they can be explained by the possibility that dividing attention during the list presentation left the participants free to recoup some of the losses by rehearsing or retrieving later items in the list after the trial's divided-attention task had ended, in the response phase, on the basis of a phonological or sensory memory stream. Furthermore, it is possible that the ability to retrieve previously learned associations by this group was enhanced by the implicit use of such associations; as was mentioned above, such an ability is not much compromised by divided attention. The difference between older adults and young ones under divided attention may show up when the explicit encoding of associations is required, as has been shown in previous research (e.g., *Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2003*).

In sum, we have shown that it is possible to learn more about age differences in serial recall by applying a perspective in which the focus is on the process of chunking highlighted in the seminal work of *Miller (1956)* and by considering the process of chunking in comparison with the process of interitem associations known to be deficient in older adults (e.g., *Castel & Craik, 2003; Chalfonte & Johnson, 1996; Naveh-Benjamin, 2000*). The results suggest that these processes might be partially mediated by the same binding mechanisms in which older adults are deficient. Interestingly, the present study showed that the ability to create item-temporal-context associations might be based on a different mechanism, since age effects in serial recall were comparable using strict serial scoring and free-scoring methods. Although studies of short-term recognition of lists have shown that older adults are deficient specifically in serial order information (*Hartman & Warren, 2005*), the implications for our study are unclear. In our task of serial recall of lists comprising only four

word pairs, the serial order information may be strongly linked to item information, so that the loss of order information typically is accompanied by omission of the item.

We also have provided further evidence for a capacity limit of about three independent chunks of information in young adults (*Cowan, 2001*), including a visual array comparison procedure (*Luck & Vogel, 1997*) that has been tied to specific posterior brain loci producing activity closely related to individual differences in storage capacity (*Todd & Marois, 2004; Vogel & Machizawa, 2004*). This capacity limit seems smaller in older adults.

To explain capacity limits in both age groups, neural modelers have pointed to the possibility that the representations of features bound together to form a chunk of information could be confused with those forming another chunk, if the features of the two chunks were activated too close together in time (e.g., *Luck & Vogel, 1998; Usher, Haarmann, Cohen, & Horn, 2001*). The present finding of relatively constant capacity for multiword bindings, as opposed to intraobject feature bindings, may lead to an important extension of this type of model. The finding of separable age differences in capacity, expressed in chunks, and in chunk formation places important constraints on cognitive and developmental theories.

#### AUTHOR NOTE

This research was supported by NIH Grant R01 HD-21338. We thank Beth Weinholt for data collection. Correspondence concerning this article should be addressed to M. Naveh-Benjamin, Department of Psychological Sciences, University of Missouri, 106 McAlester Hall, Columbia, MO 65211 (e-mail: [navehbenjaminm@missouri.edu](mailto:navehbenjaminm@missouri.edu)).

#### REFERENCES

- ALLEN, P. A., & COYNE, A. C. (1988). Age differences in primary organization or processing variability? Pt. I: An examination of age and primary organization. *Experimental Aging Research*, *14*, 143-149.
- ALLEN, P. A., & COYNE, A. C. (1989). Are there age differences in chunking? *Journal of Gerontology*, *44*, P181-P183.
- ALLEN, P. A., & CROZIER, L. C. (1992). Age and ideal chunk size. *Journal of Gerontology*, *47*, P47-P51.
- ANDERSON, J. R., & MATESSA, M. (1997). A production system theory of serial memory. *Psychological Review*, *104*, 728-748.
- BADDELEY, A. D. (2001). The magic number and the episodic buffer. *Behavioral & Brain Sciences*, *24*, 117-118.
- BADDELEY, A. D., & LOGIE, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). Cambridge: Cambridge University Press.
- BADDELEY, A. D., THOMSON, N., & BUCHANAN, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior*, *14*, 575-589.
- BASTIN, C., & VAN DER LINDEN, M. (2005). The effects of aging on the recognition of different types of associations. *Experimental Aging Research*, *32*, 61-77.
- BROADBENT, D. E. (1975). The magic number seven after fifteen years. In A. Kennedy & A. Wilkes (Eds.), *Studies in long-term memory* (pp. 3-18). New York: Wiley.
- CASTEL, A. D., & CRAIK, F. I. M. (2003). The effects of aging and divided attention on memory for item and associative information. *Psychology & Aging*, *18*, 873-885.
- CHALFONTE, B. L., & JOHNSON, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, *24*, 403-416.
- CHEN, Z., & COWAN, N. (2005). Chunk limits and length limits in immediate recall: A reconciliation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *31*, 1235-1249.
- COWAN, N. (1999). An embedded-processes model of working memory.

- In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge: Cambridge University Press.
- COWAN, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral & Brain Sciences*, **24**, 87-185.
- COWAN, N., CHEN, Z., & ROUDER, J. N. (2004). Constant capacity in an immediate serial-recall task: A logical sequel to Miller (1956). *Psychological Science*, **15**, 634-640.
- COWAN, N., ELLIOTT, E. M., SAULTS, J. S., NUGENT, L. D., BOMB, P., & HISMJATULLINA, A. (2006). Rethinking speed theories of cognitive development: Increasing the rate of recall without affecting accuracy. *Psychological Science*, **17**, 67-73.
- COWAN, N., NAVEH-BENJAMIN, M., KILB, A., & SAULTS, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult? *Developmental Psychology*, **42**, 1089-1102.
- COWAN, N., SAULTS, J. S., ELLIOTT, E. M., & MORENO, M. (2002). Deconfounding serial recall. *Journal of Memory & Language*, **46**, 153-177.
- CRAIK, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London: Series B*, **302**, 341-359.
- 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 performance* (pp. 409-422). Amsterdam: North-Holland.
- CRAIK, F. I. M., & JENNINGS, J. M. (1992). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51-110). Hillsdale, NJ: Erlbaum.
- CRAIK, F. I. M., & KESTER, J. D. (2000). Divided attention and memory: Impairment of processing or consolidation? In E. Tulving (Ed.), *Memory, consciousness, and the brain: The Tallinn Conference* (pp. 38-51). New York: Psychology Press.
- ERICSSON, K. A., CHASE, W. G., & FALOON, S. (1980). Acquisition of a memory skill. *Science*, **208**, 1181-1182.
- FARRELL, S., & LEWANDOWSKY, S. (2004). Modelling transposition latencies: Constraints for theories of serial order memory. *Journal of Memory & Language*, **51**, 115-135.
- GLISKY, E. L., RUBIN, S. R., & DAVIDSON, P. S. R. (2001). Source memory in older adults: An encoding or retrieval problem? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 1131-1146.
- GRAESSER, A., II, & MANDLER, G. (1978). Limited processing capacity constrains the storage of unrelated sets of words and retrieval from natural categories. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 86-100.
- HARTMAN, M., & WARREN, L. H. (2005). Explaining age differences in temporal working memory. *Psychology & Aging*, **20**, 645-656.
- HAY, J. F., & JACOBY, L. L. (1999). Separating habit and recollection in young and older adults: Effects of elaborative processing and distinctiveness. *Psychology & Aging*, **14**, 122-134.
- HENSON, R. N. A., NORRIS, D. G., PAGE, M. P. A., & BADDELEY, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology*, **49A**, 80-115.
- HULME, C., STUART, G., BROWN, G. D. A., & MORIN, C. (2003). High- and low-frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory & Language*, **49**, 500-518.
- JACOBY, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, **30**, 513-541.
- JENNINGS, J. M., & JACOBY, L. L. (1993). Automatic versus intentional uses of memory: Aging, attention, and control. *Psychology & Aging*, **8**, 283-293.
- JENNINGS, J. M., & JACOBY, L. L. (1997). An opposition procedure for detecting age-related deficits in recollection: Telling effects of repetition. *Psychology & Aging*, **12**, 352-361.
- JOHNSON, N. F. (1978). The memorial structure of organized sequences. *Memory & Cognition*, **6**, 233-239.
- KAHANA, M. J., & CAPLAN, J. B. (2002). Associative asymmetry in probed recall of serial lists. *Memory & Cognition*, **30**, 841-849.
- KAUSLER, D. H., & PUCKETT, J. M. (1979). Effects of word frequency on adults' age differences in word memory span. *Experimental Aging Research*, **5**, 161-169.
- KYNETTE, D., KEMPER, S., NORMAN, S., & CHEUNG, H. (1990). Adults' word recall and word repetition. *Experimental Aging Research*, **16**, 117-121.
- LI, S., NAVEH-BENJAMIN, M., & LINDENBERGER, U. (2005). Aging neuromodulation impairs associative binding. *Psychological Science*, **16**, 445-450.
- LIGHT, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology*, **43**, 333-376.
- LIGHT, L. L. (2000). Memory changes in adulthood. In S. H. Qualls & N. Abeles (Eds.), *Psychology and the aging revolution: How we adapt to longer life* (pp. 73-97). Washington, DC: American Psychological Association.
- LIGHT, L. L., PATTERSON, M. M., CHUNG, C., & HEALY, M. R. (2004). Effects of repetition and response deadline on associative recognition in young and older adults. *Memory & Cognition*, **32**, 1182-1193.
- LUCK, S. J., & VOGEL, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.
- LUCK, S. J., & VOGEL, E. K. (1998). Response from Luck and Vogel [A response to "Visual and auditory working memory capacity," by N. Cowan, in the same issue.] *Trends in Cognitive Sciences*, **2**, 78-80.
- MANDLER, G., & PEARLSTONE, Z. (1966). Free and constrained concept learning and subsequent recall. *Journal of Verbal Learning & Verbal Behavior*, **5**, 126-131.
- MAYLOR, E. A., VOUSDEN, J. I., & BROWN, G. D. A. (1999). Adult age differences in short-term memory for serial order: Data and a model. *Psychology & Aging*, **14**, 572-594.
- MCCORMACK, P. D. (1982). Temporal coding by young and elderly adults: A test of the Hasher-Zacks model. *Developmental Psychology*, **17**, 509-515.
- MILLER, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, **63**, 81-97.
- MITCHELL, K. J., JOHNSON, M. K., RAYE, C. L., MATHER, M., & D'ESPOSITO, M. (2000). Aging and reflective processes of working memory: Binding and test load deficits. *Psychology & Aging*, **15**, 527-541.
- MOREY, C. C., & COWAN, N. (2005). When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 703-713.
- MURPHY, D. A., CRAIK, F. I. M., LI, K. Z. H., & SCHNEIDER, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychology & Aging*, **15**, 323-334.
- NAVEH-BENJAMIN, M. (1990). Coding of temporal order information: An automatic process? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 117-126.
- NAVEH-BENJAMIN, M. (2000). Adult-age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 1170-1187.
- NAVEH-BENJAMIN, M. (2002). The effects of divided attention on encoding processes: Underlying mechanisms. In M. Naveh-Benjamin, M. Moscovitch, & H. L. Roediger III (Eds.), *Perspectives on human memory and cognitive aging: Essays in honor of Fergus Craik* (pp. 193-207). New York: Psychology Press.
- NAVEH-BENJAMIN, M., CRAIK, F. I. M., GAVRILESCU, D., & ANDERSON, N. D. (2000). Asymmetry between encoding and retrieval processes: Evidence from divided attention and a calibration analysis. *Memory & Cognition*, **28**, 965-976.
- NAVEH-BENJAMIN, M., GUEZ, J., KILB, A., & REEDY, S. (2004). The associative memory deficit of older adults: Further support using face-name associations. *Psychology & Aging*, **19**, 541-546.
- NAVEH-BENJAMIN, M., GUEZ, J., & MAROM, M. (2003). The effects of divided attention at encoding on item and associative memory. *Memory & Cognition*, **31**, 1021-1035.
- NAVEH-BENJAMIN, M., GUEZ, J., & SHULMAN, S. (2004). Older adults' associative deficit in episodic memory: Assessing the role of decline in attentional resources. *Psychonomic Bulletin & Review*, **11**, 1067-1073.
- NAVEH-BENJAMIN, M., HUSSAIN, Z., GUEZ, J., & BAR-ON, M. (2003). Adult-age differences in episodic memory: Further support for an

- associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 826-837.
- RABINOWITZ, J. C., CRAIK, F. I. M., & ACKERMAN, B. P. (1982). A processing resource account of age differences in recall. *Canadian Journal of Psychology*, **36**, 325-344.
- SAINT-AUBIN, J., & POIRIER, M. (1999). The influence of long-term memory factors on immediate recall: An item and order analysis. *International Journal of Psychology*, **34**, 347-352.
- SALTHOUSE, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, **103**, 403-428.
- SALTHOUSE, T. A., & BABCOCK, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, **27**, 763-776.
- SLAK, S. (1970). Phonemic recoding of digital information. *Journal of Experimental Psychology*, **86**, 398-406.
- TAUB, H. A. (1974). Coding for short-term memory as a function of age. *Journal of Genetic Psychology*, **125**, 309-314.
- TODD, J. J., & MAROIS, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, **428**, 751-754.
- TULVING, E., & PATKAU, J. E. (1962). Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. *Canadian Journal of Psychology*, **16**, 83-95.
- TULVING, E., & PEARLSTONE, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning & Verbal Behavior*, **5**, 381-391.
- USHER, M., HAARMANN, H., COHEN, J. D., & HORN, D. (2001). Neural mechanism for the magical number 4: Competitive interactions and non-linear oscillations. *Behavioral & Brain Sciences*, **24**, 151-152.
- VOGEL, E. K., & MACHIZAWA, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, **428**, 749-751.
- WILSON, M. (1988). The MRC Psycholinguistic Database: Machine-readable dictionary, Version 2. *Behavior Research Methods, Instruments, & Computers*, **20**, 6-10.
- WINGFIELD, A., STINE, E. L., LAHAR, C. J., & ABERDEEN, J. S. (1988). Does the capacity of working memory change with age? *Experimental Aging Research*, **14**, 103-107.
- YONELINAS, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, **25**, 747-763.

(Manuscript received December 30, 2005;  
revision accepted for publication March 27, 2006.)