Journal of Gerontology: PSYCHOLOGICAL SCIENCES
2005, Vol. 60B, No. 2, P92–P97

Intrusions in Episodic Recall: Age Differences in Editing of Overt Responses

Michael J. Kahana,[1] Emily D. Dolan,[1] Colin L. Sauder,[1] and Arthur Wingfield[2]

[1]Department of Psychology, University of Pennsylvania, Philadelphia.
[2]Volen Center for Complex Systems, Brandeis University, Waltham, Massachusetts.

Two experiments compared episodic word-list recall of young and older adults. In Experiment 1, using standard free-recall procedures, older adults recalled significantly fewer correct items and made significantly more intrusions (recall of items that had not appeared on the target list) than younger adults. In Experiment 2, we introduced a new method, called externalized free recall, in which participants were asked to recall any items that came to mind during the recall period but to indicate with an immediate key press those items they could identify as intrusions. Both age groups generated a large number of intrusions, but older adults were significantly less likely than young adults to identify these as nonlist items. Results suggest that an editing deficit may be a contributor to age differences in episodic recall and that externalized free recall may be a useful tool for testing computationally explicit models of episodic recall.

TO UNDERSTAND the mechanisms underlying episodic memory, we must understand both correct responses and memory errors. In word-list recall tasks, errors reflect both omissions from a to-be-recalled memory set plus intrusions in the form of erroneous recall of items that had not appeared in the target list. To experimentally induce intrusions, Deese (1959) presented participants with word lists comprising items that were semantically related to a nonpresented critical item. Following study of these lists, participants typically recall the nonpresented critical item with a probability as high (or higher) than that of the target list items. Significant new attention has been directed toward understanding the factors underlying semantic intrusions in this ''false-memory'' paradigm (Roediger & McDermott, 1995). Among the findings are that older adults are more likely than their younger counterparts to show false memories for words semantically associated with target items (Norman & Schacter, 1997; Tun, Wingfield, Rosen, & Blanchard, 1998).

Our focus here is on intrusions committed during standard free-recall tasks, in which lists are not specifically designed to evoke false memories. In such lists, participants do occasionally make intrusions, though they are very few in number. Older adults, who recall fewer correct items on average than their younger counterparts, have also been reported to make a greater number of intrusions than younger adults (Hartley & Walsh, 1980; Kahana, Howard, Zaromb, & Wingfield, 2002; Stine & Wingfield, 1987). Although many studies have been done to help researchers understand older adults' recall deficit (e.g., Kahana et al., 2002; Kausler, 1994; Naveh-Benjamin, 2000; Wingfield, Lindfield, & Kahana, 1998), relatively little is known about the age-related increase in intrusions observed in standard recall tasks. This is because there are often too few intrusions available for analysis (Kahana & Wingfield, 2000; Keppel, 1968; Wingfield et al., 1998).

Process models of episodic recall (e.g., the search of associate memory, or SAM, model of Raaijmakers & Shiffrin, 1981; see also Raaijmakers, 2003) posit that overt responses reflect an initial process of sampling candidates for recall—a process that is largely associative—followed by a subsequent editing process that limits responses to those items that were in the target list. SAM thus embraces the classic generate–recognize view that participants do in fact think of many items that were not on the target list, but that they reject those responses prior to producing them for overt recall (Anderson & Bower, 1972; Keppel 1968; Raaijmakers & Shiffrin, 1980; Wixted & Rohrer, 1994). Thus, older adults' increased tendency to produce intrusions may reflect either a tendency to generate more intrusions or an inability to recognize that a generated intrusion did not come from the target list. Theoretically, such recognition failures may reflect a deficit in source monitoring insofar as recognition judgments rely, in part, on an ability to remember the context in which an item occurred (Henkel, Jonhson, & De Leonardis, 1998; Johnson, Hashtroudi, & Lindsay, 1993; Kelley & Sahakyan, 2003). On the basis of studies of list discrimination, it would seem that older adults do indeed have difficulty remembering the source of a presented item (e.g., McCormack, 1984), but it is not known whether this difficulty also applies to items that are being overtly produced. One may also view older adults' tendency to commit intrusions as a consequence of a deficit in inhibitory processes (Hasher & Zacks, 1988; Hasher, Stoltzfus, Zacks, & Rypma, 1991). In the case of recall from episodic memory, inhibition may be viewed as a set of mechanisms or control processes that help to focus retrieval on target information by inhibiting the activation of extraneous thoughts or the retrieval of associated information that was experienced in a different context. Effective use of these control processes may depend on the retrieval of information indicating the source of a generated item.

Despite considerable theoretical and empirical research on free recall, and a general acknowledgment that some type of internal editing process must be operating, almost nothing is known about how often and under what circumstances participants censor their recalls. We devised an externalized free recall (EFR) task to potentially reveal this hypothesized editing process. Our method builds on that of Bousfield and Rosner (1970; see also Roediger & Payne, 1985), who asked participants to say aloud any words that came to mind during the recall phase of a free-recall task.

In our variant of this technique, participants were further instructed to press a key immediately following any response that they knew to be incorrect. Thus, the EFR task attempts to open to experimental scrutiny the dual processes of generating and editing potential recall responses.

The existing literature, in demonstrating striking age-related deficits in episodic memory, suggests that older adults are impaired in their ability to remember an item's temporal context or to inhibit items that have inappropriate contextual information. Such a deficit in inhibition, source monitoring, or the formation–retrieval of episodic associations would predict that older adults, after generating an intrusion, will have difficulty rejecting it as a nontarget item. This would also explain the age-related increase in intrusions committed during episodic recall tasks. Our primary interest, therefore, is in how often participants were able to correctly reject intrusions in the EFR procedure.

The EFR procedure also provides information on the total number of extralist words generated by young and older participants, whether from prior lists (prior-list intrusions, henceforth PLIs) or words that had not appeared in the experiment (extralist intrusions, henceforth XLIs). To the extent that older adults generally produce fewer items in verbal fluency tests than young adults (e.g., Kempler, Teng, Dick, Taussig, & Davis, 1998; Tombaugh, Kozak, & Rees, 1999), older adults might produce fewer intrusions than young adults. One might equally imagine scenarios in which older adults might produce more words under EFR instructions than young adults.

We report on two experiments. The first experiment is intended to demonstrate the traditional finding that, in free recall, older adults recall fewer items and make more intrusions (both PLIs and XLIs) than their younger education-matched counterparts. The second experiment uses EFR to induce greater numbers of intrusions in both age groups, but especially among younger adults who rarely make intrusions in ordinary free recall. By examining both the number of intrusions committed and the probability of accepting intrusions, we wished to determine whether older adults might be less able than their younger counterparts to correctly reject these intrusions.

Experiment 1

Our goal in Experiment 1 was to demonstrate age difference in recall accuracy, and a greater number of intrusions in a standard delayed free-recall task.

Methods

Participants

The young participants were 12 university students, 6 women and 6 men, with ages ranging from 19 to 23 years ($M = 20.5$ years, $SD = 1.3$). At time of testing, the group had a mean of 14.7 years of formal education ($SD = 0.8$) and a mean Wechsler Adult Intelligence Scale–Revised (WAIS-R) vocabulary score of 51.6 ($SD = 4.0$). The older participants were 12 healthy, community-dwelling volunteers, 10 women and 2 men, with ages ranging from 62 to 79 years ($M = 73.3$, $SD = 5.2$). At time of testing, the older participants had a mean of 14.5 years of formal education ($SD = 3.4$) and a mean WAIS-R vocabulary score of 47.0 ($SD = 8.2$). Both groups were thus well educated and with good verbal ability and did not differ significantly in either years of formal education, $t(22) = 0.16$, *ns*, or on WAIS-R vocabulary, $t(22) = 1.61$, *ns*.

Because vocabulary grows with adult aging, it is not uncommon in research studies for the older participants to have higher vocabulary scores than the young participants (Nicholas, Barth, Obler, Au, & Albert, 1997). In our case, the two age groups did not differ significantly in vocabulary. We conducted a regression analysis and confirmed that, for these participants, there was no relation between vocabulary score and the intrusion rates to be described. All participants reported themselves to be in good health and to have no difficulty reading the words as they would be presented on the computer screen.

Stimuli and Procedures

Participants studied a total of 10 lists of common, two-syllable nouns, chosen at random without replacement from a subset of the Toronto Word Pool (Friendly, Franklin, Hoffman, & Rubin, 1982), with each list followed by a delayed free-recall test. Half of the lists were composed of 10 nouns (short lists) and half were composed of 20 nouns (long lists), with the order of short and long lists counterbalanced between participants.

At the start of each trial, the computer displayed each list item in capital letters for 1.4 s, followed by a 100-ms blank interstimulus interval (ISI). During list presentation, participants were required to say each word aloud. Immediately following the presentation of list items, participants performed a 16-s arithmetic distracter task consisting of addition in the form "$A + B + C = \ldots$," where A, B, and C were positive, single-digit integers. Participants were instructed to work quickly while maintaining accuracy. Errors were rare. A row of asterisks, accompanied by a tone (lasting 300 ms), then signaled that participants should begin recalling the list items aloud, in any order. Participants were allowed 90 s to recall list items; on the basis of previous research with young and older adults (e.g., Kahana et al., 2002), we have found that 90 s provides ample time for recall, with virtually all correct responses being produced during the first 60 s of this recall period. During the 90-s recall period, participants were to recall as many items as possible from the list, in any order. We considered the first four lists to be practice and we excluded them from the subsequent analyses. This served both to eliminate warm-up effects, which can differentially affect young and older participants, and to mirror the analyses to be used in Experiment 2.

Results and Discussion

The main results are summarized in Table 1, which gives both the mean number and the proportion of words correctly recalled from 10-word and 20-word lists. Also shown are the mean numbers of intrusions, which are further subdivided on the basis of whether or not the intruded word was one that appeared on an earlier list. Data for the young adults are shown on the left and data for the older adults are shown on the right.

As one can see, participants recalled a smaller number of words in short lists (although a larger proportion of the total list) than they did in long lists, and the young adults recalled a greater number of correct list items than older adults. We confirmed this by using a 2(list length: short, long) $\times$ 2(age: young, older) mixed design analysis of variance (ANOVA) conducted on the number of words correctly recalled, with list length as a within-participants variable and age as a between-participants variable. The ANOVA revealed a significant main effect of list length, $F(1, 22) = 19.44$, $MSE = 1.55$, $p < .001$, and of age, $F(1, 22) = 7.62$, $MSE = 4.82$, $p < .05$, as well as a significant List length × Age interaction, $F(1, 22) = 5.94$, $MSE = 1.58$, $p < .05$. Thus, participants of both ages made more correct recalls in the long list condition (albeit representing a smaller percentage of items recalled), which is consistent with the classic list-length effect pervasive in free recall (Murdock, 1962; Roberts, 1972; Ward, Woodward, Stevens, & Stinson, 2003). As we expected, older adults recalled fewer items in both list-length conditions (e.g., Craik, 1968; Laurence, 1967; Schonfield, 1965), an effect that was differentially greater for the long lists. Such an interaction would be expected because recalling items from longer lists is more difficult and thus likely to be differentially harder for older adults, who are impaired in the mechanisms supporting episodic memory retrieval (see Kausler, 1994 for a discussion).

In addition to looking at correct responses, we were also interested in characterizing the behavior of intrusions, both PLIs and XLIs. We submitted the number of PLIs produced by the young and older participants for the long and short lists shown in Table 1 to a 2 (list length) × 2 (age) mixed design ANOVA with list length as a within-participants variable and age as a between-participants variable. The ANOVA confirmed significant main effects of list length, $F(1, 22) = 4.84$, $MSE = 0.06$, $p < .05$, and of age, $F(1, 22) = 5.16$, $MSE = 0.42$, $p < .05$. The interaction between age and list length did not quite reach our significance threshold, $F(1, 22) = 3.64$, $MSE = 0.06$, $p = .07$.

Older adult's increased tendency to commit PLIs is consistent with several prior reports (e.g., Hartley & Walsh, 1980; Kahana et al., 2002). The finding that PLIs are more frequent in longer lists is not surprising, given the greater number of total recalls in long lists and the tendency for intrusions to be associatively evoked by list items (Roediger & McDermott, 1995). We observed a similar qualitative pattern for XLIs insofar as there was a marginally significant effect of age on number of intrusions $F(1, 22) = 3.56$, $MSE = 0.16$, $p = .07$. For XLIs, the effect of list length, and the Age × List length interaction did not approach significance. Aggregating XLIs and PLIs together, we found significant effects of both age, $F(1, 22) = 6.06$, $MSE = 0.83$, $p < .05$, and of list length, $F(1, 22) = 12.04$, $MSE = 0.01$, $p < .01$, on the total number of intrusions committed. The interaction between age and list length was not statistically significant, $F(1, 22) = 1.93$, $MSE = 0.01$, *ns*.

Repetitions of already recalled items represent another category of memory errors. Although repetition errors are not the focus of this article, we report them here for completeness. Young and older adults rarely repeated items; the mean number of repetitions per list was 0.10 ($SD = 0.13$) and 0.21 ($SD = 0.19$) for young and older adults, respectively; this difference did not reach significance, $t(22) = 1.7$, *ns*. Younger adults never repeated intrusions ($M = 0$), and older adults did so very rarely ($M = 0.03$, $SD = 0.07$).

Table 1. Experiment 1: Mean Number and Proportions of Correct Recalls and Intrusions

	Younger				Older			
	Short		Long		Short		Long	
Parameter	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Correct recalls (per list)								
No. of words	4.65	1.38	7.10	2.22	3.77	0.74	4.48	2.31
Proportion of list	0.46	0.14	0.36	0.11	0.38	0.07	0.22	0.12
Intrusions (per list)								
Prior list	0.17	0.22	0.19	0.30	0.46	0.59	0.75	0.69
Extra list	0.08	0.12	0.25	0.35	0.31	0.47	0.46	0.46
Total	0.25	0.21	0.44	0.39	0.77	0.93	1.21	0.88

Note: Proportions are grand averages across participants.

Experiment 2

In Experiment 2 we again tested free recall for a group of young and older adults, but this time we used the EFR procedure. In the EFR procedure, we asked participants to say aloud all of the words that came to mind while trying to recall the just-presented list. Our intent was to disinhibit participants from editing their responses (albeit in an artificial setting), and thereby increase the number of intrusions committed by both the younger and older participants. Our main interest was in our participants' ability to correctly reject these intrusions.

Methods

Participants

The young participants were 40 university students and staff, 23 women and 17 men, with ages ranging from 18 to 28 years ($M = 20.5$ years, $SD = 2.4$). At time of testing, the group had a mean of 14.7 years of formal education ($SD = 1.7$) and a mean WAIS-R vocabulary score of 52.2 ($SD = 6.4$). The older participants were 40 healthy, community-dwelling older adults, 26 women and 14 men, with ages ranging from 61 to 87 years ($M = 73.3$, $SD = 5.9$). At time of testing, the older participants had a mean of 16.4 years of formal education ($SD = 2.5$) and a mean WAIS-R vocabulary score of 53.8 ($SD = 5.3$). Both groups were thus well educated and with good vocabulary ability. The older adults had a mean of 1.7 more years of formal education at time of testing than the young adults, $t(74) = 3.43$, $p < .01$, but the two groups did not differ significantly in WAIS-R vocabulary scores, $t(72) = 1.19$, *ns* (Vocabulary scores were unavailable for several participants.) As in Experiment 1, the young and older participants had equivalent vocabulary scores. We again conducted a regression analysis and confirmed that, for these participants, there was no relation between vocabulary score and the intrusion rates to be described. All participants reported themselves to be in good health and to have no difficulty reading the words as they would be presented on the computer screen.

Procedures

The method for Experiment 2 was the same as for Experiment 1 with one key addition: On trials 4–10 participants were instructed, in addition to recalling words from the just-presented list, to recall aloud all of the words that came to mind as they were attempting to give their list recall. Participants were told, however, that if a word they had just given had not been on the just-presented list, then they were to press the spacebar on the computer keyboard immediately after recalling the word but prior to recalling the next word.

Table 2. Mean Number of Responses, by Type, and the Probability of Accepting Them

	Younger				Older			
	Short		Long		Short		Long	
Parameter	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Correct responses (per list)								
No. accepted	4.45	1.57	6.80	2.10	3.13	1.03	4.26	1.36
No. rejected	0.37	0.40	0.59	0.57	0.38	0.45	0.44	0.44
Probability of accepting	0.92	0.09	0.91	0.08	0.90	0.12	0.91	0.09
Prior-list intrusions (per list)								
No. accepted	0.51	0.62	0.58	0.67	0.76	0.74	0.62	0.62
No. rejected	3.06	3.84	2.78	3.57	1.80	2.30	1.32	2.01
Probability of accepting	0.18	0.21	0.27	0.28	0.39	0.33	0.43	0.36
Extra-list intrusions (per list)								
Mean no. accepted	0.18	0.28	0.37	0.36	0.73	0.87	0.67	0.78
Mean no. rejected	1.24	1.23	1.19	1.05	1.29	1.22	1.13	1.46
Probability of accepting	0.13	.021	0.25	0.23	0.37	0.32	0.38	0.32
All intrusions (per list)								
Mean no. accepted	0.69	0.72	0.94	0.70	1.49	1.28	1.29	1.06
Mean no. rejected	4.29	3.87	3.97	3.58	3.09	2.52	2.45	2.46
Probability of accepting	0.15	0.14	0.24	0.18	0.37	0.25	0.42	0.29

Note: Probabilities of accepting were estimated separately for each participant and then averaged across participants.

Participants were given 90 s to recall the list items. During this time they were to recall as many items as possible from the list, in any order they wished. As in Experiment 1, participants were tested on a total of 10 lists (half of which were short, i.e., 10 items, and other half long, i.e., 20 items). We did not include performance on the first three lists, on which participants performed standard delayed free recall, and on the first list of EFR in the analyses reported. Thus, as in Experiment 1, we conducted the analyses of correct recall and intrusions on the data from lists 5–10.

Results and Discussion

The results of Experiment 2 are summarized in Table 2, which shows the mean number of accepted and rejected correct responses per list, and the probability that participants accepted correct responses. Table 2 also shows the mean number of accepted and rejected PLI and XLIs, along with the probability of accepting each type of intrusion.

When considering correct responses, one can either consider all correct responses or restrict the analyses to only those correct responses that were accepted as correct. Table 2 shows that, in both cases, participants recalled a smaller number (but greater proportion) of the words in short lists than in long lists, and younger adults recalled a greater number of correct list items than older adults. In what follows, we report results in terms of accepted correct responses. (We also obtained the pattern of significant effects to be described for all correct responses taken together.)

A two-way mixed design ANOVA that we conducted on the number of correct items that were not rejected revealed significant main effects of list length, $F(1, 78) = 116.09$, $MSE = 1.04$, $p < .001$, and of age, $F(1, 78) = 38.97$, $MSE = 3.84$, $p < .001$, as well as a significant List length × Age interaction $F(1, 78) = 14.23$, $MSE = 1.04$, $p < .001$. This replicates with EFR instructions the basic effects of list length and age on correct recall demonstrated in Experiment 1 using standard free-recall procedures.

Our specific question pertains to the editing process in recall: The probability that, when a participant produces an intrusion, it will be erroneously accepted as a correct response. Assuming that participants' behavior in EFR at least partially mirrors the generation and editing of responses hypothesized to underlie successful episodic recall, we would expect to find that older adults are impaired at editing their responses. Such impairment predicts that older adults would be more likely than their younger counterparts to accept as correct their intrusion errors.

We separately analyzed young and older adults probabilities of accepting PLIs and XLIs (see Table 2). For the probability of accepting PLIs, there was a significant main effect of age, $F(1, 70) = 11.59$, $MSE = 0.12$, $p < .01$, but the effects of list length, $F(1, 70) = 1.02$, $MSE = 0.05$, *ns*, and the List length × Age interaction, $F(1, 70) = 1.29$, $MSE = 0.05$, *ns*, did not approach significance. Thus, older adults were more likely than younger adults to accept a PLI as correct, regardless of list length.

For the probability of accepting XLIs, there was a significant main effect of age, with older adults exhibiting a greater tendency to accept XLIs as correct, $F(1, 64) = 7.32$, $MSE = 0.12$, $p < .01$. Here, however, the main effect of list length was also significant, $F(1, 64) = 4.15$, $MSE = 0.04$, $p < .05$, indicating that participants were more likely to accept intrusions as correct in long lists than they were in shorter lists. There was no significant Age × List length interaction, $F(1, 64) = 2.63$, $MSE = 0.04$, $p = .11$.

Treating PLIs and XLIs together, we obtained significant main effects of both age, $F(1, 76) = 18.86$, $MSE = 0.08$, $p < .001$, and list length, $F(1, 76) = 8.44$, $MSE = 0.02$, $p < .01$, but no significant Age × List length interaction, $F(1, 76) < 1$, *ns*. Thus, older adults are more likely than younger adults to accept an intrusion as correct, and both groups were more likely to accept intrusions in recalling longer lists.

Although our interest was primarily in the probability that young and older adults would accept intrusions as correct, in Table 2 we also give the mean raw numbers of accepted and rejected intrusions per list. For the total number of accepted intrusions (PLIs and XLIs combined), there was a significant main effect of age, $F(1, 78) = 8.10$, $MSE = 1.60$, $p < .01$, and there was no main effect of list length, $F(1, 78) = 0.08$, $MSE = 0.30$, *ns*, but there was a significant Age × List length interaction, $F(1, 78) = 6.79$, $MSE = 0.30$, $p < .05$. For the total number of rejected intrusions (PLIs and XLIs together), there was a significant main effect of age, $F(1, 78) = 3.87$, $MSE = 19.09$, $p = .05$, and of list length, $F(1, 78) = 9.37$, $MSE = 1.00$, $p < .05$, but the Age × List length interaction did not approach significance, $F(1, 78) = 1.01$, $MSE = 1.00$, *ns*.

In addition to considering correct responses, PLIs and XLIs, we also report the number of times participants repeated items during recall. As in Experiment 1, young and older adults did not differ significantly in the number of times they repeated correct list items: younger, $M = 0.56$, $SD = 1.23$; older, $M = 0.44$, $SD = 0.45$; $t(78) = 0.6$, *ns*. However, older adults did show a trend toward repeating intrusions more frequently than the young: younger, $M = 0.11$, $SD = 0.23$; older, $M = 0.25$, $SD = 0.42$; $t(78) = 2.0$, $p = .05$.

General Discussion

Older adults' difficulty in learning new information is seen most easily in recall tasks that rely on the ability to forge new associations among novel items, and that supply little in the way of contextual support during recall (Naveh-Benjamin, 2000; Kahana et al., 2002; Wingfield et al., 1998). As demonstrated in Experiment 1, older adults recall fewer correct items, and they also produce more errors in the form of intrusions of items that had not been in the list than do their younger counterparts. This experiment adds to an emerging body of evidence showing that, although intrusions are very few in number, reliable age differences are easily observed (Kahana et al., 2002).

There are several possible causes of older adults' increased tendency to commit intrusions during recall in standard recall tests such as demonstrated in Experiment 1. Theories of memory (e.g., Raaijmakers & Shiffrin, 1980, 1981) traditionally assume that, during recall, participants internally generate items as they search memory but censor their responses when they believe that the internally generated item had not occurred on the list, perhaps because of a mismatch in source information (e.g., Johnson et al., 1993). Within this framework, older adults' increased tendency to commit intrusions could result from either a greater tendency to internally generate nonlist items during recall or a diminished ability to censor the internally generated nonlist items, perhaps because of a failure of retrieving the source of the retrieved items (e.g., Henkel et al., 1998). Finally, a third possibility is that older adults' increased tendency to commit intrusions arises simply because they are more willing to "think out loud," producing responses that they know were not on the list because, in standard recall, there is no penalty for doing so. According to this last possibility, older adults' increased tendency to overtly commit intrusions does not necessarily reflect an increased tendency to internally generate intrusions or a lack of knowledge concerning the source of these internally generated intrusions. Clearly these factors are not mutually exclusive: Older adults' tendency to commit intrusions in the standard recall task could result from some combination of a conscious think-out-loud strategy, a reduced ability to distinguish between intrusions versus items on the target list, or an inability to inhibit these intrusions from response.

In standard free recall, all of the aforementioned factors are obscured from experimental scrutiny—one can only measure the frequency of participants' intrusions. We therefore designed a new method, called externalized free recall, to unveil the process of editing. EFR provides participants with the opportunity to speak aloud any items that they may think of and to overtly reject any responses they deem to be errors.

Experiment 2 demonstrated that older adults are impaired in their ability to edit intrusions, both those that came from prior lists and those that had not appeared on any list. That is, older adults are more likely to accept intrusions as list items than their younger counterparts. This finding is consistent with an age-related deficit in the ability to inhibit from response nonlist items that come to mind during recall, perhaps because of a failure to remember the source of those items.

The EFR method can only approximate the covert editing process hypothesized to take place during recall. Whereas some participants may not be comfortable responding overtly, others may excessively free associate during recall in order to be compliant with the experimenter. Another potential concern is that instructing participants to press a key to reject responses may interfere with the normal processes operating during recall. Further research using the EFR procedure is needed to determine how well it approximates the normal recall process.

Albeit with these caveats and cautions, data generated using the EFR method may be particularly suitable for testing computationally explicit models of episodic recall. Many such models (e.g., Becker & Lim, 2002; Kahana, 1996; Raaijmakers, 2003; Raaijmakers & Shiffrin, 1981) posit that, during recall, participants generate candidate responses for recall and then edit these responses to reduce or eliminate intrusion errors. For example, one could extend the classic SAM model to use the item-to-context strength in the model to "recognize" sampled items (e.g., Sirotin, Kimball & Kahana, 2004). On the basis of this postsampling recognition process, items may be rejected or accepted. Data from EFR could then be used to test such enhanced versions of SAM.

Using the EFR method, our data suggest that older adults are less effective in editing their intrusions than their younger counterparts, with older adults accepting 36% of their overt intrusions and younger adults accepting only 21%. Although an inability to edit may underlie older adults' tendency to make intrusions, this is only one of multiple mechanisms that are likely to account for the large age-related deficits seen in episodic retrieval.

Acknowledgments

We acknowledge support from the National Institutes of Health under Grants MH55687 and AG15852 and support from the W. M. Keck Foundation.

Address correspondence to M. J. Kahana, Department of Psychology, University of Pennsylvania, 3401 Walnut Street, Philadelphia, PA 19104 or to A. Wingfield, Volen National Center for Complex Systems, MS 013, Brandeis University, Waltham, MA 02454-9110. E-mail: kahana@psych.upenn.edu or wingfield@brandeis.edu

References

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review, 79*, 97–123.

Becker, S., & Lim, J. (2003). A computational model of prefrontal control in free recall: Strategic memory use in the California Verbal Learning Task. *Journal of Cognitive Neuroscience, 15*, 821–832.

Bousfield, W. A., & Rosner, S. R. (1970). Free vs. uninhibited recall. *Psychonomic Science, 20*, 75–76.

Craik, F. I. (1968). Two components in free recall. *Journal of Verbal Learning & Verbal Behavior, 7*, 996–1004.

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

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto World Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods and Instrumentation, 14*, 375–399.

Hartley, J. T., & Walsh, D. A. (1980). The effect of monetary incentive on amount and rate of free recall in older and younger adults. *Journal of Gerontology, 35*, 899–905.

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.
Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17*, 163–169.
Henkel, L. A., Johnson, M. K., & De Leonardis, D. M. (1998). Aging and source monitoring: Cognitive processes and neuropsychological correlates. *Journal of Experimental Psychology, 127*, 251–268.
Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3–28.
Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition, 24*, 103–109.
Kahana, M. J., Howard, M. W., Zaromb, F., & Wingfield, A. (2002). Age dissociates recency and lag recency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*, 530–540.
Kahana, M. J., & Wingfield, A. (2000). A functional relation between learning and organization in free recall. *Psychonomic Bulletin & Review, 7*, 516–521.
Kausler, D. M. (1994). *Learning and memory in normal aging.* San Diego, CA: Academic Press.
Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language, 48*, 704–721.
Kempler, D., Teng, E. L., Dick, M., Taussig, I. M., & Davis, D. S. (1998). The effects of age, education, and ethnicity on verbal fluency. *Journal of the International Neuropsychological Society, 4*, 531–538.
Keppel, G. (1968). Retroactive and proactive inhibition. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 172–213). Englewood Cliffs, NJ: Prentice-Hall.
Laurence, M. W. (1967). Memory loss with age: A test of two strategies for its retardation. *Psychonomic Science, 9*, 209–210.
McCormack, P. D. (1984). Temporal coding by young and elderly adults in a list-discrimination setting. *Bulletin of the Psychonomic Society, 22*, 401–402.
Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64*, 482–488.
Nicholas, M., Barth, C., Obler, L. K., Au, R., & Albert, M. L. (1997). Naming in normal aging and dementia of the Alzheimer's type. In H. Goodglass & A. Wingfield (Eds.), *Anomia: Neuroanatomical and cognitive correlates* (pp. 166–188). Boston: Academic Press.
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.
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.
Raaijmakers, J. G. W. (2003). Spacing and repetition effects in human memory. *Cognitive Science, 27*, 431–452.
Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 207–262). New York: Academic Press.
Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88*, 93–134.
Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 803–814.
Roediger, H. L., & Payne, D. G. (1985). Recall criterion does not affect recall level or hypermnesia: A puzzle for generate/recognize theories. *Memory & Cognition, 13*, 1–7.
Roberts, W. A. (1972). Free recall of word lists varying in length and rate of presentation: A test of total time hypothesis. *Journal of Experimental Psychology, 92*, 365–372.
Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403–428.
Schonfield, D. (1965). Memory changes with age. *Nature, 208*, 918.
Sirotin, Y. B., Kimball, D. R., & Kahana, M. J. (2004). *Going beyond a single list: Modeling the effects of prior experience on episodic free recall.* Manuscript submitted for publication.
Stine, E. L., & Wingfield, A. (1987). Process and strategy in memory for speech among younger and older adults. *Psychology and Aging, 2*, 272–279.
Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology, 14*, 167–177.
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.
Ward, G., Woodward, G., Stevens, A., & Stinson, C. (2003). Using overt rehearsals to explain word frequency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 29*, 186–210.
Wingfield, A., Lindfield, K. C., & Kahana, M. J. (1998). Adult age differences in the temporal characteristics of category free recall. *Psychology and Aging, 13*, 256–266.
Wixted, J. T., & Rohrer, D. (1994). Analyzing the dynamics of free recall: An integrative review of the empirical literature. *Psychonomic Bulletin & Review, 1*, 89–106.

Received March 30, 2004
Accepted October 13, 2004
Decision Editor: Thomas M. Hess, PhD