Journal of Experimental Psychology:
Learning, Memory, and Cognition
1996, Vol. 22, No. 1, 143-156


0278-7393/96/$3.00

# Studies of Directed Forgetting in Older Adults

Rose T. Zacks
Michigan State University

Gabriel Radvansky
University of Notre Dame

Lynn Hasher
Duke University

Younger and older adults were compared in 4 directed forgetting experiments. These varied in the use of categorized versus unrelated word lists and in the use of item by item versus blocked remember–forget cueing procedures. Consistent with L. Hasher and R. T. Zacks's (1988) hypothesis of impaired inhibitory mechanisms in older adults, a variety of findings indicated that this age group is less able than younger adults to suppress the processing and retrieval of items designated as to be forgotten (TBF). Specifically, in comparison with younger adults, older adults produced more TBF word intrusions on an immediate recall test (Experiments 1A and 1B), took longer to reject TBF items (relative to a neutral baseline) on an immediate recognition test (Experiment 3), and recalled (Experiments 1A, 1B, and 2) and recognized (Experiments 1B and 2) relatively more TBF items on delayed retention tests in which all studied items were designated as targets.

In this article, we present four experiments comparing the performance of younger and older adults on directed forgetting tasks. In this type of task (e.g., see Bjork, 1989), participants are presented items to study, some of which they are told to remember and others of which they are told to forget. Because the cueing as to which items are to be remembered (TBR items) and which are to be forgotten (TBF items) occurs after the items have been presented for study, participants must pay some attention to each item as it is presented. Thus, the directed forgetting paradigm investigates the ability to forget some inputs that one has recently attended to while at the same time remembering others presented in the same context and near the same time. To the degree that one is successful at this task, as younger adults generally are, the following trends are seen: The presence of TBF items on a list does not reduce recall or recognition of TBR items; there are few intrusions of TBF items when participants are asked to report only TBR items; and performance on TBF items is relatively poor when, on a later retention test, participants are asked to report TBF as well as TBR items.

Rose T. Zacks, Department of Psychology, Michigan State University; Gabriel Radvansky, Department of Psychology, University of Notre Dame; Lynn Hasher, Department of Psychology: Social and Health Sciences, Duke University.

Portions of the research reported in this article were presented at the 1991 meeting of the Midwestern Psychological Association in Chicago, Illinois. This research was supported in part by National Institute on Aging Grant AGO 4306 and by National Institute of Mental Health Quantitative Training Grant MH14257. We thank Beth Chittenden, Heather Oonk, Dan Spieler, and Christy Stewart for their assistance in collecting the data. We also thank Cynthia May and Michael Kane for their very helpful comments on drafts of this article.

Correspondence concerning this article should be addressed to Rose T. Zacks, Department of Psychology, Michigan State University, East Lansing, Michigan 48824-1117. Electronic mail may be sent via Internet to zacksr@pilot.msu.edu.

Our interest in age comparisons on directed forgetting tasks has both empirical and theoretical sources. The empirical motivation for this research derives from the notion that the directed forgetting paradigm can tell us something about mental processing in situations in which it is beneficial to be able to forget information that turned out to be wrong or that is no longer relevant. Because such situations are likely to be common in the life experience of older as well as younger adults, it is perhaps surprising that the literature contains few studies of directed forgetting in older adults. The ones that we are aware of include a study that used the unusual procedure of cueing participants before study about the TBR–TBF status of individual items (Pavur, Comeaux, & Zeringue, 1984), a second that was reported at a conference and is briefly summarized in a book chapter (Camp & McKitrick, 1989), and a third that was the basis of another conference report (Giambra & Howard, 1994). Given this situation (few studies and incomplete reports of all but one), it seems premature to reach even tentative conclusions about directed forgetting effects in older adults without additional research. Suffice it to say that to the extent that comparisons are possible, the results of the previous studies do not contradict those reported in this article.

The theoretical motivation for our studies relates to the framework proposed by Hasher and Zacks (1988). A central concept in this framework is that of attentional inhibition. In particular, building upon models of attention that emphasize selection for action (e.g., Allport, 1989; Keele & Neill, 1979; Neumann, 1987), we proposed that there are two basic mechanisms of selective attention: activation and inhibition. Our conception is that inhibition operates in the service of goals by suppressing the activation of goal-irrelevant information so that such information is less likely to have access to working memory and so that irrelevant information that does enter working memory, as well as previously relevant information that is no longer useful, is quickly removed. Furthermore, analogous to the mechanism of inhibition of return in visual search (e.g., Klein & Taylor, 1994), attentional inhibition may also have the function of preventing the return of attention to a previously rejected item, whether that item is an external stimulus event or a thought (May, Kane, & Hasher, 1995; Stoltzfus, Hasher, Zacks, Ulivi, & Goldstein, 1993).

A focus of our work on attentional inhibition has been in the area of adult age differences in inhibitory efficiency. On the basis of a close reading of the cognitive gerontology literature, we (Hasher & Zacks, 1988) suggested that older adults may have deficient inhibitory mechanisms. Subsequently, this proposal has been supported by a variety of findings. One important finding relates to the *negative priming* effect, which is the finding from selective attention tasks that, relative to a new stimulus, responding is slowed to a stimulus that served as a selected against distractor on a preceding trial (e.g., Tipper, 1985). Several studies have found that, in contrast to younger adults, older individuals do not consistently show this key marker of attentional inhibition (Hasher, Stoltzfus, Zacks, & Rypma, 1991; Kane, Hasher, Stoltzfus, Zacks, & Connelly, 1994; McDowd & Oseas-Kreger, 1991; Stoltzfus et al., 1993; Tipper, 1991). Furthermore, older adults have been found to show a broad spectrum of effects consistent with the notion of deficient inhibitory mechanisms. These include increased susceptibility to interference from concurrent environmental distraction (Connelly, Hasher, & Zacks, 1991) and from concurrently activated goal-irrelevant thoughts (Gerard, Zacks, Hasher, & Radvansky, 1991). To be sure, the relevant data may not be completely supportive (e.g., see Sullivan & Faust, 1993, for an exception to the general pattern of reduced negative priming in older adults), and some of the findings (including possibly even those reported in this article) may have alternative interpretations (see Light, 1991, for a general critique of our viewpoint). Nonetheless, we believe that, overall, there is good support for the viewpoint that there is an aging-related decline in the efficiency of attentional inhibition mechanisms. Of central importance to the present studies, previous research has also provided evidence of delayed or weakened suppression of ideas that turn out to be wrong in the context of processing of garden-path sentences (Hartman & Hasher, 1991) and longer texts (Hamm & Hasher, 1992). (A summary of these findings can be found in Zacks & Hasher, 1994.)

The present research extends the application of this theory to aging-related differences in directed forgetting. This extension is motivated in large part by theoretical views (Bjork, 1989; Geiselman & Bagheri, 1985; MacLeod, 1989) arguing that directed forgetting involves multiple mechanisms, including two of an inhibitory nature. Specifically, these are the stopping of the rehearsal of an item following the presentation of a forget cue and the inhibition of TBF item retrieval. To the degree that this approach is valid, and the bulk of the evidence is supportive (Bjork, 1989), our view predicts that older adults will be less successful than younger adults in complying with forget cues. Evidence confirming this prediction would be supportive both of the notion that inhibitory mechanisms decline in old age and of the hypothesized role of inhibition in directed forgetting effects.

The four experiments reported in this article involve a variety of procedures, materials, and dependent measures. In particular, item by item forget–remember cueing was used in two experiments (1A and 1B), whereas blocked cueing was used in two others (Experiments 2 and 3). Likewise, the experimental materials consisted of categorized word lists for two experiments (1A and 1B) and random word lists for the other two (Experiments 2 and 3), and the dependent measures included immediate and final free-recall tests (Experiments 1A, 1B, and 3), final recognition tests (Experiments 1B and 2), and immediate probe recognition tests (Experiment 3). Although the effects were sometimes subtle, in all experiments there was evidence that older adults were less able than younger adults to comply with directed forgetting cues.

## Experiments 1A and 1B

The first two experiments used highly similar materials and procedures and are reported together. The learning materials were categorized word lists in which some of the words from a particular category were associated with a remember cue and others with a forget cue. All experimental lists were 24 words long and remember–forget cueing was done on a word-by-word basis after each word had been studied for several seconds. On the immediate recall tests that followed each study list, participants tried to recall all the TBR items from the current list and none of the TBF items. Our hope was that this situation would create a particularly sensitive task environment for demonstrating some of the predicted age differences in directed forgetting. Specifically, given that the preexperimental associative connections among members of a category are generally quite strong, it should be more difficult, relative to unrelated words, to suppress the TBF items from a category when other items from the same category have to be remembered. Both age groups would be expected to intrude some TBF items on the immediate recall tests, but the intrusion rate should be higher for older adults.

In addition to the immediate recall tests, Experiments 1A and 1B included final recall tests on which participants were asked to recall all the words from all of the study lists, regardless of whether they had previously been designated as TBR or TBF items. In Experiment 1B, there was also a final recognition test on which participants were told to include TBF as well as TBR items in the old category. Although both younger and older adults were expected to show greater retention of TBR than TBF items on these delayed tests, the reduced ability of older adults to inhibit processing of TBF items at encoding should result in smaller TBR–TBF differences for this age group. In other words, relative to each age group's performance on the TBR items, older adults were expected to show better long-term retention of TBF items than younger adults.

Other than the inclusion of a final recognition test in Experiment 1B but not in Experiment 1A, the only significant difference between the two experiments was in some of the materials used. The lists used in Experiment 1A included both exhaustive categories, which have only four members (e.g., seasons of the year), and nonexhaustive categories, which have a large number of members (e.g., fruit). The Experiment 1B lists included only nonexhaustive categories.

## Method

### Participants

The participants for all of these experiments came from the same sources. In particular, most of the younger adults were recruited from the undergraduate participant pool at Michigan State University and were given course credit for their involvement. The remaining younger adult participants were also Michigan State University undergraduates who were recruited through an advertisement in the student newspaper and were paid for their participation. The older adults were recruited from the greater Lansing, Michigan community, provided their own transportation to the university, and were paid for their participation. Table 1 presents information on the ages, education levels, and Wechsler Adult Intelligence Scale—Revised (WAIS–R; Weschler, 1981) Vocabulary scores of the 24 younger and 24 older participants in each experiment whose data are reported below.

The younger adults in Experiment 1A came from the participant pool, whereas those in Experiment 1B were paid for participating. In addition to those listed in Table 1, four other participants were tested in Experiment 1A but had to be replaced: Two participants, 1 older and 1 younger, were lost because of experimenter error, 1 younger adult was replaced for scoring too low (below 30) on the WAIS–R, and 1 older adult was replaced for failing to finish the experiment. Likewise, 5 additional participants were replaced in Experiment 1B, including 2 older adults because of experimenter error, 1 older adult for failing to finish the experiment, 1 younger adult for scoring too low on the WAIS–R, and 1 younger adult for failing to follow the instructions.

### Materials

The study lists in Experiment 1A were derived from a pool of 36 categories, 18 of which were exhaustive categories containing only four members each (e.g., north, south, east, and west), and the other 18 of which were nonexhaustive categories containing many more items. For the latter, we used 18 of the categories in the Battig and Montague (1969) norms and selected 4 words from among the 10 most frequently produced instances. These words were arranged into six lists, such that each list was 24 words long and contained all four members of each of 3 exhaustive and 3 nonexhaustive categories. In addition, each list included 1 category in which there were 0 TBR words and 4 TBF (0R-4F) words, one category with 1 remember and 3 forget (1R-3F) words, 2 categories with 2 remember and 2 forget (2R-2F) words, 1 category of 3 remember and 1 forget (3R-1F) words, and 1 category with 4 remember and 0 forget (4R-0F) words. Subject to these constraints, a computer program generated 12 sets of word lists, each of which was used for one pair of participants in each age group. Within each pair of participants, all of the TBR words for one member of the pair were TBF words for the other and vice versa.

The study lists in Experiment 1B had a similar construction to those of Experiment 1A, except that only nonexhaustive categories were used. For each of 36 categories in the Battig and Montague (1969) norms, 8 words were selected from among the 10 most frequently mentioned category members for a total of 288 words. Across the six experimental lists, each participant studied 4 of the words from each category with the other 4 serving as distractors on the final recognition test. The computer program that generated the experimental lists created six sets of word lists, each of which was used for a subgroup of 4 participants in each age group. Within these subgroups, the 144 words that were included in the study lists for 2 participants served as recognition distractors for the other two and vice versa. Also, for each of these participant pairs, the words that were TBR words for one member of the pair were TBF words for the other and vice versa. For the recognition test, all 288 words were presented in a random order in a five-page booklet. All participants received the same recognition test.

### Procedure

A green and white monitor controlled by an Apple IIe computer was used to present the study lists in 40-column mode. Participants pressed the space bar on the computer to begin the presentation of each list. Then the following sequence of events occurred for each of the words in the list: A beep sounded and a fixation point (a + sign) was presented in the center of the screen for 500 ms. The fixation point was then replaced with a word from the list that remained on the screen for 5 s. After the 5 s elapsed, the word was erased and either an F or an R was presented for 1 s. These letters indicated whether the preceding word was a TBF or a TBR word, respectively. At the end of each list (24 words), the computer paused while the participant wrote down as many TBR words as they could recall. A separate sheet of paper was used for each list. When the participant was ready, the recall sheet was

Table 1
*Characteristics of Participants in Each Experiment*

| Experiment number and participant group | Age (years) | | Education level (years) | | | WAIS–R Vocab. scores | | |
|---|---|---|---|---|---|---|---|---|
| | *M* | Range | *M* | Range | *t*(df) | *M* | Range | *t*(df) |
| Experiment 1A | | | | | | | | |
| Younger adults | 19.4 | 18–22 | 13.9 | 12–16 | | 45.1 | 30–60 | |
| Older adults | 68.2 | 59–79 | 13.4 | 8–18 | | 48.7 | 34–66 | |
| Experiment 1B | | | | | | | | 1.25 (46)† |
| Younger adults | 21.5 | 19–26 | 14.7 | 12–16 | | 47.7 | 33–63 | |
| Older adults | 70.6 | 61–81 | 14.8 | 10–20 | | 51.9 | 30–69 | |
| Experiment 2 | | | | | 1.97 (46)† | | | 2.78 (46)* |
| Younger adults | 19.3 | 18–24 | 13.0 | 12–16 | | 44.4 | 32–58 | |
| Older adults | 70.5 | 62–77 | 14.1 | 12–18* | | 50.9 | 31–61 | |
| Experiment 3 | | | | | | | | 2.76 (46)* |
| Younger adults | 21.7 | 18–29 | 15.2 | 12–17 | | 49.2 | 38–62 | |
| Older adults | 68.8 | 62–80 | 15.7 | 12–20 | | 57.6 | 33–69 | |

*Note.* $n$ = 24 for all groups. Ages and education levels are in years. WAIS–R = Wechsler Adult Intelligence Scale—Revised; Vocab. = Vocabulary.
†Age difference is nearly significant at $p < .10$. *Age difference is signficant at $p < .05$.

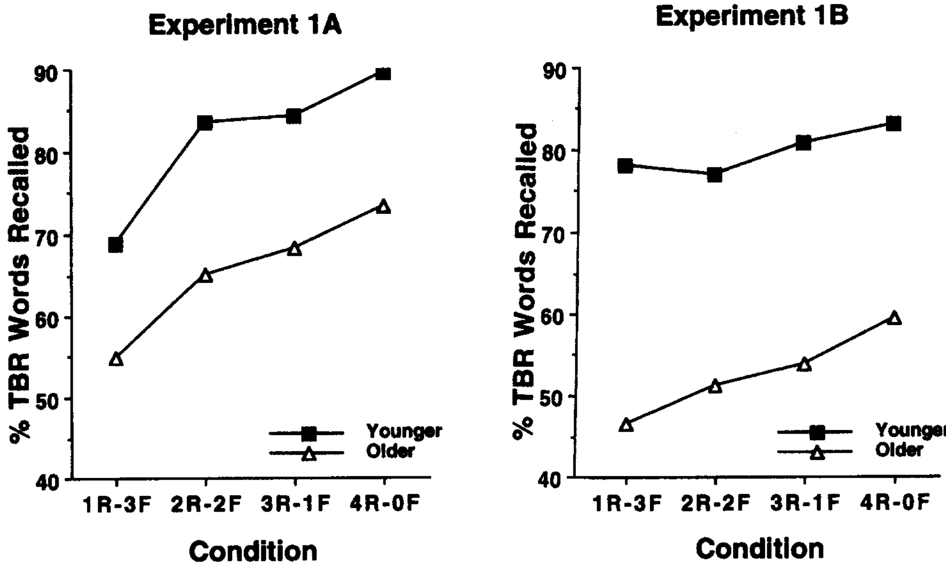


*Figure 1.* Mean percent recall of to-be-remembered (TBR) words on the immediate recall tests of Experiments 1A and 1B. 1R-3F = one remember and three forget words; 2R-2F = two remember and two forget words; 3R-1F = three remember and one forget words; 4R-0F = four remember and zero forget words.

turned over and the space bar was pressed to start the presentation of the next list.

After all six lists had been presented, the participants in both Experiments 1A and 1B were given a list of 80 three-digit addition problems (e.g., 365 + 927 = ?), which they worked on during the 5-min retention interval preceding the final recall test. (No participant completed all of the problems.) On this final test, participants were given 5 min to recall as many TBR and TBF words from the previous lists as they could. In Experiment 1B, the final recall test was followed by the final recognition test. Participants were told to circle the words they remembered having seen before regardless of whether they were TBR or TBF words. There was no time limit on the recognition test. In these two and the following two experiments reported in this article, participants were administered the WAIS-R vocabulary test as the final event in the experimental session, and the entire session lasted approximately 1 hour.

## Results

Unless otherwise indicated, a .05 alpha level was used for all statistical tests reported in this article. In initial analyses of the data from the immediate and final recall tests of Experiment 1A, we included category type (exhaustive vs. nonexhaustive) as a variable. We had originally thought that the members of exhaustive categories would tend to be more strongly interconnected than those of nonexhaustive categories, but our analyses found few significant main effects or interactions involving this factor and none that altered the conclusions derived from performance averaged over category type.[1] In hindsight, we believe that by using high frequency exemplars of the nonexhaustive categories, we minimized differences in intracategory associations across the two category types. To simplify the presentation of the results, we ignore category type in what follows (but see footnote 1).

### *Immediate Recall of TBR Items, Experiments 1A and 1B*

Mean percentages of TBR items recalled on the immediate tests are shown in Figure 1. Each set of data was submitted to a 2 (age) × 4 (number of TBR words in a category) mixed analysis of variance (ANOVA). As can be expected from the figure, the outcomes of these two ANOVAs were highly similar. In both experiments, older adults recalled fewer TBR words than younger adults: Experiment 1A, mean of 65.4% versus 81.6%, $F(1, 46) = 30.0$, $MSE = 416.3$; Experiment 1B, mean of 52.7% versus 79.4%, $F(1, 46) = 53.4$, $MSE = 639.3$. In addition, as the number of TBR words in a category increased, there was an increase in the proportion of TBR words recalled, $F(3, 138) = 20.0$, $MSE = 168.4$, for Experiment 1A, and $F(3, 138) = 4.0$, $MSE = 179.7$, for Experiment 1B. In neither experiment was the interaction of age and number of TBR words significant ($F$s $< 1$).

### *Intrusions in Immediate Recall, Experiments 1A and 1B*

Among the intrusions in immediate recall, the greatest number (and the type of primary interest) were intrusions of current-list TBF items. The TBF intrusion data for each

[1] In the immediate recall data, none of the comparisons involving category type were significant. However, on the delayed recall test, recall of TBF items from exhaustive categories was slightly higher ($M = 23.6\%$) than that from nonexhaustive categories ($M = 18.1\%$), $F(1, 46) = 9.9$, $MSE = 302.5$. The same comparison was not significant in the case of TBR items ($F < 1$). The only significant category type effect involving age group was the largely uninterpretable Age × Category Type × Number of TBR–TBF Words in category interaction in the final recall of TBF items, $F(3, 138) = 3.4$, $MSE = 375.8$.

experiment (shown in Table 2) were analyzed in a 2 (age) × 4 (number of TBR–TBF words in a category) mixed ANOVA. As can be seen in Table 2, older adults intruded a greater proportion of the TBF words on the immediate recall tests than did younger adults: Experiment 1A, $F(1, 46) = 12.6$, $MSE = 111.7$; Experiment 1B, $F(1, 46) = 21.4$, $MSE = 108.3$. In addition, both ANOVAs revealed a significant main effect of the number of TBR–TBF words and a significant Age × Number of TBF Words interaction: For Experiment 1A, $Fs(3, 138) = 4.4$ and 5.1 for the main effect and interaction, respectively, $MSE = 55.0$; for Experiment 1B, $Fs(3, 138) = 3.3$ and 3.0 for the main effect and interaction, respectively, $MSE = 100.3$. There was a similar pattern for the older adults in Experiments 1A and 1B; their rate of TBF intrusions in immediate recall increased as the number of TBF words in a category decreased (or, alternatively, as the number of TBR words increased), $F(3, 69) = 5.4$, $MSE = 96.3$, and $F(3, 69) = 3.0$, $MSE = 150.1$, for Experiments 1A and 1B, respectively. In contrast, the younger adults in the two experiments showed different patterns on this measure: no effect ($p > .10$) in Experiment 1A and a curvilinear relationship in Experiment 1B, $F(3, 69) = 3.8$, $MSE = 50.4$.

In both Experiment 1A and Experiment 1B, the rates for other types of intrusions, including unpresented words from one of the current list's categories, unpresented words from previous categories, TBR words from previous lists, TBF words from previous lists, and new words, were low and there were no significant age differences.

### *Final Recall, Experiment 1A*

Although the final recall data from Experiments 1A and 1B show similar trends, these data sets are somewhat more complicated than those already reported and we describe them separately. The final recall data from Experiment 1A are displayed in Table 3. Average recall collapsed over number of TBR–TBF items in a category was analyzed in a 2 (age) × 2 (TBR vs. TBF) mixed ANOVA. The older adults recalled a smaller percentage of all words (25.9%) than the younger adults (38.6%), $F(1, 46) = 36.5$, $MSE = 106.8$, and more TBR words were recalled (43.7%) than TBF words (20.8%), $F(1, 46) = 217.7$, $MSE = 57.4$. There was also a significant interaction of these two variables, $F(1, 46) = 5.8$, $MSE = 57.4$. Although both age groups recalled a higher percentage of TBR than TBF words, this difference was smaller for older adults (TBR = 35.4% and TBF = 16.3%) than for younger adults (TBR = 51.9% and TBF = 25.4%).

Separate analysis of the TBR and TBF items revealed the additional finding that the recall of both types of items increased as the number of TBR items in a category got larger (and as the number of TBF items got smaller): for TBR items, $F(3, 138) = 4.49$, $MSE = 230.7$, for TBF items, $F(3, 138) = 10.7$, $MSE = 159.6$.

Table 2
*Mean Percentage of To-Be-Forgotten Intrusions in Immediate Recall in Experiments 1A and 1B*

| Experiment and age group | Condition | | | | |
|---|---|---|---|---|---|
| | 0R-4F | 1R-3F | 2R-2F | 3R-1F | *M* |
| Experiment 1A | | | | | |
| Younger | 1.7 | 3.7 | 3.8 | 2.1 | 2.8 |
| Older | 4.3 | 5.8 | 8.3 | 14.6 | 8.3 |
| Experiment 1B | | | | | |
| Younger | 1.4 | 2.3 | 6.8 | 0.3 | 2.7 |
| Older | 4.3 | 8.6 | 11.1 | 14.6 | 9.7 |

*Note.* 0R-4F = zero remember and four forget words; 1R-3F = one remember and three forget words; 2R-2F = two remember and two forget words; 3R-1F = three remember and one forget words.

Table 3
*Mean Percent Recall of To-Be-Remembered (TBR) and To-Be-Forgotten (TBF) Items on the Final Recall Test of Experiment 1A*

| Age group | Condition | | | | | |
|---|---|---|---|---|---|---|
| | 0R-4F | 1R-3F | 2R-2F | 3R-1F | 4R-0F | *M* |
| TBR items | | | | | | |
| Younger | | 42.4 | 54.7 | 54.0 | 56.6 | 51.9 |
| Older | | 31.3 | 38.6 | 35.6 | 36.3 | 35.4 |
| *M* | | 36.8 | 46.6 | 44.8 | 46.4 | 43.7 |
| TBF items | | | | | | |
| Younger | 17.0 | 24.5 | 30.7 | 29.2 | | 25.4 |
| Older | 9.2 | 14.1 | 21.9 | 20.1 | | 16.3 |
| *M* | 13.1 | 19.3 | 26.3 | 24.6 | | 20.8 |

*Note.* 0R-4F = zero remember and four forget words; 1R-3F = one remember and three forget words; 2R-2F = two remember and two forget words; 3R-1F = three remember and one forget words; 4R-0F = four remember and zero forget words.

### *Final Recall, Experiment 1B*

The data from the final recall test of Experiment 1B were quite similar to those of Experiment 1A, as can be seen by comparing Table 4 with Table 3. A 2 (age) × 2 (TBR vs. TBF items) mixed ANOVA was performed on recall in Experiment 1B averaged over the number of TBR–TBF items in a category. The older adults recalled fewer words overall (18.4%) than the younger adults (27.6%), $F(1, 46) = 15.7$, $MSE = 127.7$. There were more TBR words recalled (34.7%) than TBF words (11.2%), $F(1, 46) = 214.8$, $MSE = 62.1$, and as in Experiment 1A, the interaction of these two variables was significant, $F(1, 46) = 34.1$, $MSE = 62.1$. The age difference in favor of the young group in the recall of TBR words was substantial and significant, $F(1, 46) = 30.4$, $MSE = 543.0$, whereas the small age difference in the other direction for the TBF words was not ($F < 1$). Finally, in separate analyses of the TBR and TBF items, we found that recall of each of these types of items increased as the number of TBR words in a category increased (and the number of TBF items decreased), $F(3, 138) = 4.0$, $MSE = 184.7$, and $F(3, 138) = 2.7$, $MSE = 79.8$, for TBR and TBF items, respectively.

### *Final Recognition Test, Experiment 1B*

Only Experiment 1B included a final recognition test. The hit-rate and false-alarm results from this test are shown in

Table 4
*Mean Percent Recall of To-Be-Remembered (TBR) and To-Be-Forgotten (TBF) Items on the Final Recall Test of Experiment 1B*

| Age group | 0R-4F | 1R-3F | 2R-2F | 3R-1F | 4R-0F | *M* |
|---|---|---|---|---|---|---|
| | Condition | | | | | |
| | TBR items | | | | | |
| Younger | | 38.2 | 45.0 | 42.4 | 50.7 | 44.1 |
| Older | | 22.9 | 23.6 | 26.2 | 29.3 | 25.5 |
| *M* | | 30.6 | 34.3 | 34.3 | 40.0 | 34.8 |
| | TBF items | | | | | |
| Younger | 8.3 | 10.7 | 12.8 | 12.5 | | 11.1 |
| Older | 8.7 | 11.1 | 10.9 | 14.6 | | 11.3 |
| *M* | 8.5 | 10.9 | 11.9 | 13.5 | | 11.2 |

*Note.* 0R-4F = zero remember and four forget words; 1R-3F = one remember and three forget words; 2R-2F = two remember and two forget words; 3R-1F = three remember and one forget words; 4R-0F = four remember and zero forget words.

Table 5. The hit-rate data were submitted to ANOVAs similar to those for the final recall data. Older and younger adults recognized similar numbers of studied items; the overall proportion of hits was 61.1% for the older group and 64.4% for the younger group ($F < 1$). More TBR words were recognized (78.4%) than TBF words (47.1%), $F(1, 46) = 225.2$, $MSE = 104.1$. However, similar to the findings on the final recall test, the difference between the TBR and TBF hit rates was greater for the younger adults than for the older adults, $F(1, 46) = 14.4$, $MSE = 104.1$, for the Age × TBR versus TBF interaction. Separate analyses of the hit rates on TBR and TBF words revealed that the age difference was significant for TBR items, $F(1, 46) = 7.9$, $MSE = 769.1$, but not for TBF items ($F < 1$). Also, for TBF items, the hit rate was higher when there were two TBR and two TBF words in the category than for other

Table 5
*Mean Percentage of Hits on the Final Recognition Test of Experiment 1B*

| Age group | 0R-4F | 1R-3F | 2R-2F | 3R-1F | 4R-0F | *M* |
|---|---|---|---|---|---|---|
| | Condition | | | | | |
| | TBR items | | | | | |
| Younger | | 79.9 | 85.1 | 85.1 | 86.1 | 84.0 |
| Older | | 74.3 | 68.4 | 72.9 | 75.3 | 72.7 |
| *M* | | 77.1 | 76.7 | 79.0 | 80.7 | 78.4 |
| | TBF items | | | | | |
| Younger | 45.3 | 45.6 | 52.4 | 36.1 | | 44.9 |
| Older | 48.6 | 49.3 | 52.4 | 47.2 | | 49.4 |
| *M* | 47.0 | 47.4 | 52.4 | 41.7 | | 47.1 |

*Note.* The false alarm percentages were 5.2% for the younger group and 8.7% for the older group. 0R-4F = zero remember and four forget words; 1R-3F = one remember and three forget words; 2R-2F = two remember and two forget words; 3R-1F = three remember and one forget words; 4R-0F = four remember and zero forget words. TBR = to be remembered; TBF = to be forgotten.

Table 6
*Mean Sensitivity (A′) and Bias (B″) Scores for Performance on the Final Recognition Tests of Experiments 1B and 2*

| Age group | Sensitivity (A′) All items | Sensitivity (A′) TBR items | Sensitivity (A′) TBF items | Bias (B″) All items | Bias (B″) TBR items | Bias (B″) TBF items |
|---|---|---|---|---|---|---|
| | Experiment 1B | | | | | |
| Younger | .89 | .94 | .82 | .64 | .41 | .64 |
| Older | .86 | .89 | .81 | .52 | .42 | .54 |
| | Experiment 2 | | | | | |
| Younger | .70 | .85 | .74 | .42 | .47 | .42 |
| Older | .66 | .79 | .75 | .37 | .47 | .45 |

*Note.* TBR = to be remembered; TBF = to be forgotten.

TBR–TBF combinations, $F(3, 138) = 5.4$, $MSE = 170.7$. Given that this finding was limited to the TBF items and that it has no apparent explanation, we do not pursue it further.

Although the two age groups did not differ in overall hit rate, the older adults produced more false alarms (8.7%) than the younger adults (5.2%), $t(46) = 2.3$, $SE = 1.5$. Consequently, we also performed signal-detection analyses on these data. Sensitivity and bias measures for overall recognition and for recognition of TBR and TBF items separately can be found in the top half of Table 6. We used the nonparametric signal-detection measures, $A'$ and $B''$, as our respective measures of sensitivity and bias (Snodgrass & Corwin, 1988). For $A'$, a value of .5 is associated with chance discrimination, and higher values indicate greater sensitivity. Collapsed across item type, the mean $A'$ of the older adults was smaller than that of the younger adults, but this difference was only marginally significant, $t(46) = 2.0$, $SE = .015$, $p < .06$.[2] When the data were examined separately for TBR and TBF items, there was a significant age difference in $A'$ for TBR items, $t(46) = 3.5$, $SE = .014$, but not for the TBF items ($t < 1$). The measure of bias, $B''$, ranges from $-1.0$ to 1.0, with 0 indicating a neutral bias, values greater than 0 indicating a liberal bias, and values less than 0 indicating a conservative bias. In the present case, both age groups displayed moderately liberal biases. Collapsed over item type, the age difference was marginally significant, $t(46) = 1.9$, $SE = .063$, $p < .07$, but it was not significant for either the TBR ($t < 1$) or TBF trials, $t(46) = 1.7$, $SE = .060$, $p > .10$, individually.

### *Discussion*

The typical finding of an age-related decline in performance on episodic memory tasks was obtained in both Experiments 1A and 1B. Older adults recalled fewer TBR items than younger adults on the immediate recall tests, and they recalled fewer total items on the delayed recall tests. Also, in Experiment 1B, their overall hit rate on the delayed recognition test was lower than that of young adults. However, in addition to these unsurprising findings, we see a number of indications

[2] The statistical analyses of $A'$ and $B''$ were limited to between-subject comparisons, because with only one false-alarm estimate for each participant, the within-subject comparisons are redundant with the analyses of the hit and false-alarm data already reported.

that the older age group was less able than the younger group to differentially process TBR and TBF items. In this respect, one important result of Experiments 1A and 1B was that older adults were more prone to report TBF words on the immediate recall test than were younger adults. Furthermore, unlike the younger adults, the older adults showed a greater tendency to intrude TBF items in immediate recall as the number of TBR items in the category increased from zero out of four to three out of four. One interpretation of this outcome is based on the presumption that older adults find it especially difficult to suppress processing of TBF items when there are closely associated TBR items from the same category in the list. A consequence of this might be a tendency to treat an entire category as a TBR category when two or more items from that category are TBR items. Younger adults, in contrast, are better able to focus on individual items despite strong preexperimental associations to other items in the list (cf. Gerard et al., 1991, for a similar argument regarding the fan effect.)

Consistent with the findings suggesting that older adults had greater difficulty than younger adults in inhibiting the retrieval of TBF items on the immediate tests were other suggestions from the delayed tests indicating that older adults processed TBR and TBF items more equally than younger adults. In particular, on the final recall tests of both experiments, the difference in the percentage of TBR and TBF words recalled was smaller for the older group. In fact, in Experiment 1B the typical age difference in recall was not found at all for TBF items. The final recognition data of Experiment 1B showed similar patterns in both the hit-rate and signal-detection analyses. Taken together, the indications from these experiments are that older adults have more difficulty suppressing TBF items than do younger adults on immediate tests and are relatively more likely to produce them on delayed memory tests.

## Experiment 2

In contrast to Experiments 1A and 1B, the next two experiments used lists of unrelated words and a procedure in which remember–forget cueing was carried out on blocks of words. In Experiment 2, a block of zero, two, or four TBF words was followed by a block of three to seven TBR words. With this type of procedure and materials, younger adults generally show a performance pattern in which there are few, if any, TBF intrusions in immediate recall, and in which the immediate recall of TBR words is unaffected by the number of TBF words in the list (e.g., Bjork, 1989). The pattern demonstrates the ability of younger adults to prevent TBF items from interfering with recall of TBR items. In contrast, we expected that for older adults the presence of TBF items in a list would have a negative impact on TBR recall and that the severity of the impact would increase with increasing numbers of TBF items in the list.

### Method

#### *Participants*

Table 1 presents information on the 24 participants in each age group who provided the following data. Four additional participants were replaced, 1 older adult for failing to follow the instructions, 1 younger adult for scoring too low on the WAIS–R (Wechsler, 1981) vocabulary test, and 2 additional younger adults as a result of experimenter error.

#### *Materials*

To form the experimental lists, 159 high-frequency (51–83 occurrences per million) nouns, five to seven letters in length, were selected from the Francis and Kučera (1982) word frequency norms. Six sets of study lists were formed from these words. Each set comprised 15 lists, 1 list for each of the combinations of 0, 2, or 4 TBF words with 3 to 7 TBR words. The study lists thus used 105 words, with the remaining 54 words serving as distractors on a recognition test at the end of the experiment. Across the six sets of lists, individual words were rotated among the TBR, TBF, and nonpresented distractor conditions so that each word occurred in each condition at least once. Otherwise, the words were randomly assigned to particular lists. There were two different orders of presentation for the conditions representing the different combinations of numbers of TBR and TBF items. Each order was used with three sets of lists. Finally, all participants saw two additional lists of very high frequency (467–832 occurrences per million) words, each consisting of 4 TBR and 4 TBF items. One of these preceded the experimental lists and served as a practice list and primacy buffer for the final retention tests; the other followed the experimental lists and served as a recency buffer for the final retention tests. Performance on these two lists was not scored. Approximately equal numbers of participants in each age group received each of the sets of experimental lists.

A stack of 3 in. × 5 in. unlined index cards held together by a large binder ring was used to present the 17 study lists that a particular participant saw. Except for the blank cards that signaled the end of each study list, there were two labels in the center of each index card. The smaller of these two labels was white and had a word printed on it in large print. A colored border was created around this white label by placing it on top of a larger label that was either green or red. Border color was used as the remember–forget cue. Specifically, a change from red to green borders or vice versa was the signal that the preceding items in that list could be forgotten, whereas no change in border color was the cue that all the words in that list were TBR items. Each color was used as the starting color or the only color for approximately half of the lists that a particular participant saw. No participants reported any difficulty differentiating between the two border colors.

A final recognition test was generated by randomly listing all 159 words in a two page booklet. All participants were given the same recognition test.

#### *Procedure*

Participants were tested in groups of one, two, or three, depending on availability. In addition to the stack of cards, participants were provided with 17 sheets of paper for immediate recalls, 1 for each of the 2 buffer and 15 experimental lists. Participants were instructed that whenever the color of the border changed within a list, they were to forget all of the words in the previous color and to remember only the words in the new color.

A tape recording of beeps emitted at a one every 5 s was used to control study time. No participants reported any difficulty hearing the tape. The cards were advanced each time the tape beeped. The younger adults flipped their own cards while the experimenter flipped the cards for the older adults because they seemed to experience more difficulty with this aspect of the procedure. When a blank card was encountered, 30 s were given to write down as many TBR words as could be recalled from the most recent list. Participants were informed that they could recall the words in any order they chose. After the words had been recalled, the recall sheet was turned over.

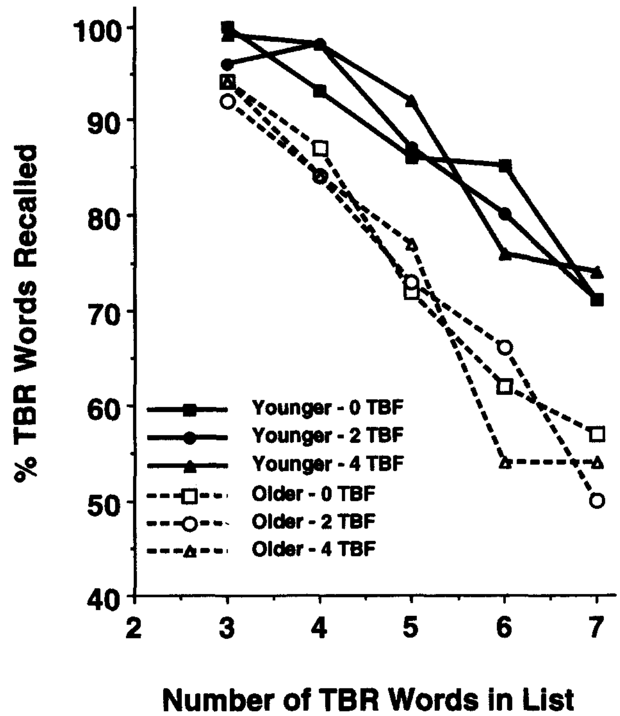


*Figure 2.* Mean percent recall of to-be-remembered (TBR) words on the immediate recall tests of Experiment 2. TBF = to be forgotten.

Following the immediate recall test on the last list, participants were presented with a sheet of paper and were asked to recall as many words as they could, regardless of whether the words were TBR or TBF words. Participants were given 5 min to perform this task. After the final recall test, participants were given the recognition test. On the recognition test, they were asked to circle as many words that they could identify as having been presented earlier. There was no time limit on the recognition test.

## *Results*

### *Immediate Recall of TBR Items*

Because of the varying numbers of TBR items across lists, recall of TBR items was measured in terms of the percentage of TBR items recalled. These data are presented in Figure 2, where it can been seen that the older adults generally recalled a lower percentage of the target items and that, contrary to our expectations, neither age group showed much of an effect of the number of TBF items in the list.

The immediate recall data were submitted to a 2 (age) × 5 (number of TBR words) × 3 (number of TBF words) mixed ANOVA. The older adults recalled a lower proportion of words (73.3%) than the younger adults (87.5%), $F(1, 46) = 20.9$, $MSE = 1{,}728.5$, and greater numbers of TBR words resulted in lower proportions of words recalled, $F(4, 184) = 97.4$, $MSE = 278.9$. There was an Age × Number of TBR Words interaction, with older adults showing a greater deficit with increasing number of TBR words, $F(4, 184) = 5.1$, $MSE = 278.9$, but interpretation of this interaction is complicated by the apparent ceiling effect for young adults in conditions with small numbers of TBR items. The main effect of the number of TBF words was not significant ($F < 1$), but this factor did interact with the number of TBR items in the list, $F(8, 368) = 2.2$, $MSE = 192.7$. Inspection of Figure 2 suggests that this reflects some irregularities in the performance functions rather than any systematic trends.

### *Intrusions in Immediate Recall*

Because there were few erroneous recalls overall (an average of 2.7 per participant over 15 lists), and because each participant received only one list at each Number of TBR Words × Number of TBF Words combination, it was not feasible to analyze the errors as a function of these list conditions. Collapsed over list conditions, erroneous recalls were categorized as new (i.e., never presented) words, TBF words from the current list, TBF words from previous lists, and TBR words from previous lists. Older and younger adults did not differ in total erroneous recalls or in any of the particular types of intrusion errors, except for those in the previous-list TBR category. Older adults produced more of these (an average of 1.13 per participant) than did younger adults (an average of .29 per participant), $t(46) = 2.4$, $SE = 2.4$.

### *Final Recall and Recognition*

Performance on the final recall and recognition tests was initially examined as a function of the numbers of TBR and TBF items in the list. As was the case for the analysis of the errors in immediate recall, and probably for the same reasons, these analyses did not reveal any systematic trends not seen in the analyses of performance averaged over list-length conditions. Consequently, only the averaged data are considered in what follows. These data are displayed in Table 7.

The final recall data were submitted to a 2 (age) × 2 (item type: TBR vs. TBF words) mixed ANOVA. As is apparent in the top half of Table 7, older adults recalled fewer words than younger adults, $F(1, 46) = 8.1$, $MSE = 34.0$, and both ages recalled more TBR than TBF words, $F(1, 46) = 27.9$, $MSE = 25.2$. Additionally, there was a significant Age × Item Type interaction, $F(1, 46) = 8.4$, $MSE = 25.2$. The difference in the

Table 7
*Mean Percent Recall and Recognition Hits on the Delayed Tests of Experiment 2*

| Test and items | Younger adults | Older adults | *M* |
|---|---|---|---|
| Final recall test | | | |
| TBR items | 12.4 | 6.1 | 9.3 |
| TBF items | 4.1 | 3.6 | 3.9 |
| *M* | 8.3 | 4.9 | |
| Final recognition test | | | |
| TBR items (hits) | 60.0 | 46.8 | 53.4 |
| TBF items (hits) | 39.2 | 36.5 | 37.8 |
| *M* | 49.6 | 41.6 | |

*Note.* The false-alarm percentages on the final recognition test were 9.7% for the younger group and 11.7% for the older group. TBR = to be remembered; TBF = to be forgotten.

rate of the recall of the TBF and TBR words was smaller for the older adults than for the younger adults. In fact, in a separate analysis of the recall of TBF items, younger and older adults were found not to differ in the recall of these items ($F < 1$). In contrast, the age difference in recall of TBR items was highly significant, $F(1, 46) = 16.3$, $SE = 29.6$. Intrusions of nonpresented items were few in number and did not differ across age groups.

The recognition hit-rate and false-alarm results are also presented in Table 7, where it can be seen that the younger adults had higher hit rates and lower false-alarm rates than the older adults, although the age differences were less striking than on the recall test. The recognition hit-rate data were submitted to an ANOVA similar to the final recall data and the results were parallel. The overall age difference in hit rate was not significant, $F(1, 46) = 2.7$, $MSE = 572.6$, $p = .11$, but the higher recognition rate of TBR words as compared with TBF words was significant, $F(1, 46) = 55.0$, $MSE = 105.7$. In addition, the Age × Item Type interaction was significant, $F(1, 46) = 6.3$, $MSE = 105.7$, reflecting a smaller age difference in the hit rates for TBF versus TBR items. Indeed, separate analyses of the hit rates for TBR and TBF items indicated a significant age difference for the former item type, $F(1, 46) = 7.4$, $MSE = 425.5$, but not for the latter ($F < 1$). The age difference in false-alarm rate was not significant ($t < 1$).

Using the same measures as were used for Experiment 1B, we also carried out signal-detection analyses on the final recognition data of Experiment 2. The relevant means are shown in the bottom half of Table 6. Similar to the results for the hit-rate analysis, the age difference in $A'$ scores was not significant for TBR and TBF items combined, $t(46) = 1.6$, $SE = .027$. However, when $A'$ scores were calculated separately for the TBR and TBF items, there was a significant age difference favoring the young adults for the TBR items, $t(46) = 3.0$, $SE = .018$, but not for the TBF words ($t < 1$). As for $B''$, it can be seen in Table 6 that, similar to Experiment 1B, the younger and older adults exhibited similar, moderately liberal biases, and indeed none of the age differences were significant.

### *Discussion*

The data for younger adults in Experiment 2 replicated previous findings from experiments using similar methods. In particular, there were few TBF intrusions in immediate recall, and the recall of TBR words on the immediate recall tests was unaffected by the number of TBF words in the study list. Furthermore, memory for TBF words showed incomplete recovery on the final recall and recognition tests.

As for age differences, the older adults generally did not perform as well as the younger adults. However, they showed little evidence on the immediate recall tests of the expected negative impact of the presence of TBF items on the study lists. Their immediate recall was unaffected by the number of TBF items in the list, and unlike in Experiments 1A and 1B, they showed no tendency to intrude TBF items to a greater extent than did younger adults. However, this latter result may reflect a floor effect in that TBF intrusion rates were very low for both age groups. An unpredicted but interesting result was the older adults' greater production of previous-list TBR intrusions in immediate recall; this suggests some failure on the part of older adults to follow the implicit instruction to forget earlier lists.

The data from the delayed retention tests were more supportive of our expectations than the immediate recall results. Specifically, the consistent finding of smaller retention differences between TBR and TBF items for older as compared with younger adults and the nonsignificant age differences on the TBF items indicate that the older group suppressed the TBF items at encoding to a lesser extent than the young group. It can also be noted that the pattern of performance on the delayed test was similar to those described by Camp and McKitrick (1989) and Giambra and Howard (1994). The next experiment presents another attempt, with what was hoped to be a more sensitive measure, to obtain evidence of age differences in suppression of TBF items on an immediate test. Specifically, Experiment 3 used immediate probed recognition tests and measured reaction times (RTs) to make yes or no decisions.

## Experiment 3

Experiment 3 used a variant of the Sternberg (1966) short-term memory search paradigm, in which the experimental lists had a blocked forget–remember construction similar to that used in Experiment 2. Specifically, each list consisted of zero to four TBF words followed by one to four TBR words, but instead of asking for recall of the TBR words after list presentation, a single word was presented as a recognition probe. Across lists, the recognition probe was one of the TBR words from the current list, one of the TBF words from the current list, or a new word. Participants were required to respond yes to the first type of recognition probe and no to the other two types.

In developing our predictions for Experiment 3, we continued to assume that given the uncertainty about if and when a forget cue might occur, participants would accord each word full attention and some amount of rehearsal effort as it was presented. Consequently, all words, including TBF words, would achieve strong initial activations in working memory. The presentation of a forget cue should result in an attempt to inhibit the activations of the words presented before the forget cue, but in general the suppression would be incomplete and the activation levels of the pre-cue words would remain above the nonpresented baseline. This incomplete suppression should be reflected in slowed *no* responses to TBF recognition probes relative to *no* responses to new (i.e., nonpresented) recognition probes (e.g., Atkinson & Juola, 1973).

Previous experiments with designs similar to that of Experiment 3 have confirmed this line of argument in the case of young adults. One of these earlier studies (Bjork, Abramowitz, & Krantz, 1970; as reported in Bjork, 1989) used digit lists, and two others (Neumann, Cherau, Hood, & Steinnagel, 1993; Neumann & DeSchepper, 1992) used word lists with similar constructions to those in this study. In all cases (including, as will be seen, our data), *no* responses to TBF probes have been found to be slower than *no* responses to new probes. The new result predicted for Experiment 3 is that under the assumption of weaker inhibitory mechanisms for older adults, the activation level of TBF items will be correspondingly higher in this age group and the difference in RTs to TBF and new recognition probes will be larger for older than for younger adults.

### *Method: Participants, Materials, and Procedure*

Data on the Experiment 3 participants are presented in Table 1. No participants had to be replaced in this experiment. The 432 words used in Experiment 3 were drawn from the Francis and Kučera (1982) word frequency norms. The words were five-to-seven letter nouns, ranging in frequency from 35 to 135 occurrences per million. Each word was used three times, once in each of three blocks of 96 trials. Within each block, each trial was composed of zero to four TBF words and one to four TBR words, with the probe word being a TBR, TBF, or new word. The 16 TBF trials in each block included one instance of each possible combination of one to four TBF items and one to four TBR items. The 48 TBR trials in a block included two instances of each TBF–TBR combination except for the zero TBF combinations, which each occurred four times. Finally, for the 32 new trials, each TBF–TBR combination was presented once except for the zero TBF combinations which were presented four times.

On each trial, a series of words was presented one at a time in the center of the video monitor of an IBM-compatible computer. Each word was presented for 1,300 ms, with an 80-ms intertrial interval. On a given trial there were zero to four TBF words and one to four TBR words. The actual selection of words for each trial was determined through a random selection procedure with a different randomization used for each participant. Words were assigned randomly without replacement to the different trial conditions so that the whole set of 432 words was used in each block. As in Experiment 2, a blocked-list construction was used for the presentation of the TBF and TBR words. At the beginning of a series, the words were presented in one color, either red or green. Participants were told that if the color of the words switched to the other color, they should forget the words in the first color and only concern themselves with remembering the words in the second color. After all of the words in the sequence had been presented, there was a 3,000-ms blank interval, and then a probe word appeared in white. This word was one of the TBR words, one of the TBF words, or a new word that had not appeared in the sequence.

Participants indicated whether the probe word was one they had to remember or not by pressing one of two buttons on a computer mouse that was held in the right hand. The left mouse button indicated "Yes, this was a word I had to remember," whereas the right mouse button indicated "No, this was not a word I had to remember." The probe word remained on the screen until the participant responded. Participants were encouraged to respond as quickly and as accurately as possible. Response times and error rates were recorded by the computer. In order to ensure that participants understood the procedure, they were given a three-trial practice period. The conditions presented during the practice period were 2F-3R, TBF probe; 3F-1R, new probe; and 0F-2R, TBR probe. In the practice period, a set of high-frequency, five-to-seven letter nouns was used. During the practice period, participants were provided with feedback concerning the correctness of their responses. Feedback was not given during the actual test. Participants were given a self-timed break every 48 trials.

### *Results*

Response times to TBF probes were slower than those to new or TBR probes, which did not differ from one another. In addition, this TBF–new difference was greater for older adults than for younger adults. This supports the notion that older adults have difficulty inhibiting irrelevant (TBF) words.

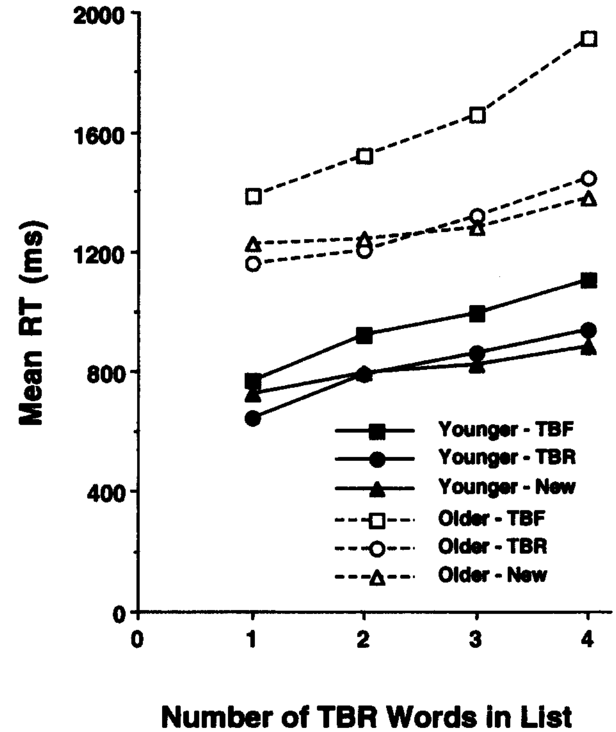


*Figure 3.* Mean reaction times (RTs) in Experiment 3 for to-be-forgotten (TBF), to-be-remembered (TBR), and new probes as a function of number of TBR words in the list (averaged over number of TBF items in the list).

The RT data from this experiment are summarized in Figures 3 and 4. In Figure 3, the results are presented as a function of the number of TBR items, averaged over the number of TBF items, whereas in Figure 4, the results are presented as a function of the number of TBF items in the list, averaged over the number of TBR items. The RT data for correct responses were submitted to a 2 (age) × 3 (probe type: TBF, TBR, or new) × 4 (number of TBR words) × 4 (number of TBF words) analysis that excluded conditions with no TBF items. This was done because these conditions of course had no trials with TBF probes. (Other analyses that did include the zero-TBF conditions were performed on the RTs to TBR and new probes. These analyses revealed no additional interesting effects.)

### *RTs*

*Overall results.* Responses to TBF probes were slower ($M$ = 1,284 ms) than those to new and TBR probes, which did not differ from one another ($M$s = 1,044 and 1,046 ms, respectively), $F(2, 92) = 46.4$, $MSE$ = 315,322. There were also significant main effects of number of TBR items, $F(3, 138) = 49.6$, $MSE$ = 177,685, and number of TBF items, $F(3, 138) = 3.9$, $MSE$ = 134,182. The first of these main effects reflects an increase in RT with increasing numbers of TBR items in the list, whereas the second reflects a decrease in RT with

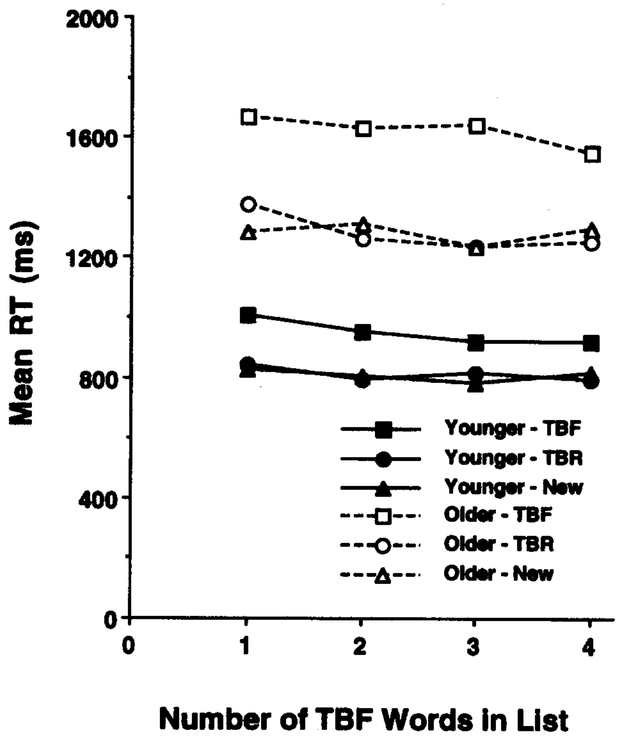


*Figure 4.* Mean reaction times (RTs) in Experiment 3 for to-be-forgotten (TBF), to-be-remembered (TBR), and new probes as a function of number of TBF words in the list (averaged over number of TBR items in the list).

increasing numbers of TBF items in the lists. The first of these effects is considerably larger than the second (differences of 292 and −64 ms, respectively, between conditions involving one item in the relevant set and those involving four items in the relevant set). The Type of Probe × Number of TBR Words interaction was also significant, $F(6, 276) = 8.4$, $MSE = 77{,}609$. The increase in RT with increasing numbers of TBR words in the list was smaller for new probes (a 157-ms increase, overall) than for TBR (290 ms) and TBF probes (430 ms). However, simple effects tests indicated that the effect of number of TBR items was significant for each probe type: TBF, $F(3, 138) = 34.2$, $MSE = 183{,}982$; TBR, $F(3, 138) = 40.5$, $MSE = 73{,}598$; new, $F(3, 138) = 11.2$, $MSE = 75{,}324$. There were no significant interactions involving TBF probes (all $F\text{s} \leq 1.4$), though it appears that the number of TBF items had almost no impact on responses to new probes.

*Aging results.* Overall, the older adults responded more slowly ($M = 1{,}395$ ms) than did the younger adults ($M = 855$ ms), $F(1, 46) = 29.1$, $MSE = 315{,}322$. More important, although the relative ordering of RTs to the three probe types was the same for the two age groups, the difference in RTs to TBF versus new probes was greater for the older (340 ms) than for the younger adults (140 ms). This is evidenced by a significant Age × Probe Type interaction, $F(2, 92) = 8.2$, $MSE = 315{,}322$. Separate ANOVAs for each age group showed that the main effect of probe type was significant for both the younger, $F(2, 46) = 37.7$, $MSE = 65{,}554$, and the older adults, $F(2, 46) = 26.1$, $MSE = 565{,}090$.

An additional analysis showed that the critical Age × Probe Type interaction reflected a significant age difference even when the overall slower responding of older adults was taken into account. Specifically, for each Number of TBR Items × Number of TBF Items cell, a participant's mean RT to TBF probes was expressed as a proportion of his or her mean RT to new probes. These proportions were then analyzed in a 2 (age) × 4 (number of TBR words) × 4 (number of TBF words) ANOVA. The important outcome of this analysis was that the age difference was significant, $F(1, 46) = 5.0$, $MSE = 0.5$. The mean TBF to new proportions were 1.18 and 1.31 for the younger and older groups, respectively. That is, relative to their RTs to new probes, younger participants were 18% slower on TBF probes, whereas older adults were 31% slower.

### *Errors*

Because the error rate in Experiment 3 was very low (2.1%), it was not feasible to examine the errors as a function of the numbers of TBR and TBF items in the list. Instead the error data were submitted to a 2 (age) × 3 (probe type) mixed ANOVA. The older adults made more errors ($M = 3.2\%$) than the younger adults ($M = 1.0\%$), $F(1, 46) = 11.7$, $MSE = 15.0$. There was also a main effect of probe type, $F(2, 92) = 18.06$, $MSE = 2.4$, with the most errors occurring for TBF word probes ($M = 2.7\%$), followed by TBR ($M = 2.5\%$) and new probes ($M = 1.0\%$). The Age × Probe Type interaction was not significant, $F(2, 92) = 2.40$, $MSE = 2.4$.

## *Discussion*

Disregarding age differences for the moment, the results of Experiment 3 replicate the usual findings in this modification of the Sternberg task. In particular, the experiments of Bjork et al. (1970, as reported in Bjork, 1989), Neumann et al. (1993), and Neumann and DeSchepper (1992) and the current experiment have all found (a) that mean RTs to TBF probes are slower than those to TBR and new probes; (b) that RTs to the latter two types of probes are quite similar; (c) that slower responses are associated with increasing numbers of TBR items for all probe types; and (d) that if anything, there is a small drop in RT associated with increasing numbers of TBF items in the list. The first of these findings suggests that in addition to comparing the probe item against each of the TBR items, as participants are believed to do in the typical Sternberg study, they do a final re-check of some sort when a probe item (specifically, a TBF probe) has a high activation level. The second and third results are consistent with those typically obtained in the original version of the Sternberg task, and the last finding agrees with the general pattern of results in other types of directed forgetting experiments (e.g., Experiments 2 and 3 agree in showing minimal negative impact on performance of increasing the number of TBF items in the list).

The main new finding of Experiment 3 was that the older adults showed a larger difference than the younger adults in response times to TBF versus new probes. We interpret this result as indicating that following the presentation of a forget cue, the activation level of TBF words remained relatively higher in the older group. In other words, because the elderly participants were less efficient in their suppression of the TBF items, they had more difficulty rejecting them when they appeared as memory probes at the end of the list. This result is further bolstered by a similar experiment (Zacks, Radvansky, & Hasher, 1993) we have conducted using lists of digits rather than words. In that study, we also found that responses to TBF items were slower than those to new items and that this difference in RT was greater for older adults. In general, we view the age differences in this experiment and in the one using digits as being consistent with the predictions generated from Hasher and Zacks's (1988) hypothesis of an age-related decline in the ability to inhibit processing of irrelevant, or no longer relevant, information.

In contrast, the age-related differences of Experiment 3 provide little support for an alternative account of directed forgetting effects in the modified Sternberg task that has been proposed by Bjork (1989), Neumann et al. (1993), and Neumann and DeSchepper (1992). Citing a presumed parallel to the negative priming paradigm described earlier, Neumann et al. (1993), and Neumann and DeSchepper (1992) concur with Bjork (1989) in making a claim opposite to our assumption that a forget cue would result in less than complete suppression of TBF items even for young adults. These authors have argued that presentation of a forget cue results in suppression of TBF items to a below-baseline level of activation. The consequence is slowed encoding of the TBF probes relative to TBR and new probes, and consequently slowed decision times for TBF items.

We find this alternative account less than compelling on a number of grounds. First, we think there is a weak basis for the presumed parallel to the negative priming effect. In the typical selective-attention task showing negative priming effects, participants are asked to make speeded responses to target items and to ignore concurrently presented distractors. Although it is clear that distractors in a selective attention task are processed to the level of meaning activation, participants certainly do not give them the kind of attention and rehearsal that TBF items in the short-term memory task receive prior to the presentation of a forget cue. That is, the amount of activation that must be suppressed in the case of distractors in a selective-attention task is much less than the amount that must be suppressed to go below baseline for TBF items in the modified Sternberg task. Another point is that the required responses differ importantly in two cases. The typical response in selective attention tasks is to identify (e.g., by naming) the target stimulus, whereas in the modified Sternberg task participants are required to say "no" to TBF items. It is easily seen how below-baseline inhibition of the representation of the target stimulus could impede the former response, but if anything, the opposite seems more likely for the latter response: In particular, some accounts of recognition memory (e.g., Atkinson & Juola, 1973) would suggest that *no* responses should benefit from, rather than be hindered by, lower activation levels (i.e., lower familiarity) of the probe items.

Most critical in the current context is the fact that the two views make different predictions about age patterns. Because older adults do not reliably show negative priming on selective-attention tasks requiring target identification (Hasher et al., 1991; McDowd & Oseas-Kreger, 1991; Stoltzfus et al., 1993; Tipper, 1991), a finding that we interpret as support for the assumption of deficient inhibition in older adults, the Bjork (1989) and Neumann (Neumann & DeSchepper, 1992) view would seem to predict that the difference in RT between TBF and new recognition probes should be smaller for older than for younger adults. This is because if the slowed response to TBF probes is due to a negative priming effect that delays encoding of TBF items, then a group of participants who show reduced negative priming effects should also show smaller delays in the encoding of the TBF probes and a smaller TBF–new probe difference. Contrary to this expectation and consistent with our view, older adults showed a larger, rather than a smaller, TBF–new difference.

## General Discussion

In each of the four experiments reported in this article, there is evidence that older adults are less able than younger adults to suppress items designated as to be forgotten. In particular, in comparison with the younger participants, the older participants produced more TBF word intrusions on an immediate recall test (Experiments 1A and 1B), took longer to reject TBF items (relative to a neutral baseline) on an immediate recognition test (Experiment 3), and showed relatively greater recall (Experiments 1A, 1B, and 2) and recognition (Experiments 1B and 2) of TBF items on delayed retention tests on which all studied items were designated as targets.[3]

Thus, these experiments provided straightforward answers to the empirical questions asked regarding age differences in the ability to forget information that is designated as no longer relevant. Clearly, older adults were less able than younger adults to ignore such information. Two aspects of the results seem particularly noteworthy: (a) Especially when information designated as irrelevant had associative connections to relevant information, there was an increased tendency for older individuals to report the irrelevant information at inappropriate times; and (b) relative to their poorer overall retention, older adults showed an elevated level of later retrieval of previously irrelevant information once the irrelevant designation was removed. The finding that suppression was particularly difficult when the no longer relevant information was associatively related to relevant information is consistent with earlier results using a reading aloud task (Connelly et al., 1991). Various types of distraction were distributed amongst

[3] One question that might be asked about our results is whether the age differences are at all attributable to the generally higher verbal abilities of the older participants (see the WAIS–R vocabulary scores in Table 1). Correlational analyses suggest that this is not the case. In particular, vocabulary score did not correlate significantly with any of a variety of presumed measures of inhibitory lapses, whether the correlations were calculated for each group separately or for the combined groups in an experiment. The measures that we correlated with vocabulary score included the number of TBF intrusions in immediate recall in Experiments 1A and 1B; the ratio of TBF to TBR delayed recall in Experiments 1A, 1B, and 2; the ratio of TBF to TBR recognition hits in Experiments 1B and 2; and the RT difference in responding to new and TBF probes in Experiment 3. Most of the correlations hovered around zero, especially in the older group.

the target text. Relative to younger adults, older adults were slower to read in the face of all distraction. They were particularly bothered, however, by distraction that was meaningfully related to the text. Suppression of related distraction, whether it was never relevant (as in Connelly et al., 1991) or whether it was once relevant (as in the present studies; see also Hamm & Hasher, 1992; Hartman & Hasher, 1991) poses a particular difficulty for older adults.

A central motivation for this research was theoretical. In particular, we sought to apply Hasher and Zacks's (1988) model of the role of inhibition in general cognitive functioning to a situation (i.e., directed forgetting) that we take to be representative of a range of tasks that require active suppression of irrelevant or once-relevant information for effective compliance with task demands (see Zacks & Hasher, 1994, for a discussion of other such tasks). Two of Hasher and Zacks's fundamental assumptions were directly relevant to the present studies: (a) One function of inhibition is to dampen no longer relevant information, and (b) older adults have inefficient inhibition mechanisms. Thus, we predicted and found that TBF items are not as effectively eliminated from older adults' response repertoires as is the case for younger adults. Likewise, in a previous text comprehension study (Hamm & Hasher, 1992) and in an indirect memory study (Hartman & Hasher, 1991), older adults had difficulty abandoning initial interpretations that turned out to be wrong. Taken together, such findings are consistent with key assumptions of the overall theoretical framework.

Having made the case that the aging trends in the present data are consistent with the reduced inhibition view, it must be acknowledged that alternative views can provide at least partial accounts of the current results. For example, on the basis of various findings of source-monitoring deficits and poorer contextual memory in older adults (see Craik & Jennings, 1992), it might be argued that older adults have greater difficulty than younger adults in keeping track of which are the TBR and which are the TBF items at encoding and at retrieval. At encoding, this could be seen as an effect of their slowed processing producing delays or lapses in registering the meaning of the remember–forget cues (especially with an item-by-item cueing procedure); alternatively, as the list goes on, older adults might be more likely to forget which items have been cued as TBF and which as TBR. In either case, older adults would tend to rehearse TBF items more than young adults and therefore show relatively higher delayed retention of TBF items. Although such arguments seem plausible, we think that the reduced inhibition view is preferable in part because it predicted the results obtained in this research. Also, some of the specific findings in these experiments seem to argue against a straightforward source-monitoring explanation of what appears to be evidence of relatively greater rehearsal of TBF items by older adults. In particular, accurate initial encoding of the TBR–TBF status of items on the part of older adults is suggested by the facts that they produced few TBF intrusions in immediate recall in Experiment 2 and that their error rate for TBF items was not elevated relative to those for other probe types in Experiment 3. Difficulties with discriminating between TBR and TBF items might arise after a delay, but these could be interpreted as a consequence of failure to inhibit rehearsal of TBF items, so the strength of TBF items is more equal to that of TBR items than it ought to be. In this way, we do not see that the source-monitoring and decreased inhibition views are necessarily incompatible; the question is whether source-monitoring problems are primary in the current situation. Our view suggests that they are not, but a definite answer will have to await further research.

One important point about our data is that although suppression of TBF words was not as effective for older as for younger adults, none of the experiments demonstrated a drastic decline in the ability to comply with directed forgetting instructions. In a few instances, as on the immediate recall tests of Experiment 2, the age differences in response to forget cues were minimal. Other data in this and the remaining experiments did show the expected age differences, but in no case did the older adults' performance fail to discriminate between TBR and TBF items. In other words, the present data indicate some reduction in the ability to forget items designated as irrelevant rather than anything like a complete loss in this ability. Presumably, this is reflective of the age trends in the underlying inhibitory mechanisms. Nonetheless, the positive findings are consistent with the conclusion that older adults have a reduced ability to control the contents of current processing. It appears that information that has been clearly designated as irrelevant remains more available in working memory and can thereby interfere with the processing of relevant information. Consequently, we can expect that older adults will be more likely to make errors in processing, to fail to make appropriate inferences because critical components cannot be readily retrieved from working memory, and to follow lines of mental processing that rely on the irrelevant information that is accessible. In addition, transitions to new topics, frames of reference, or mental models will pose particular problems for older adults, because the ability to stop the processing of no longer relevant information is impaired (see Stoltzfus, Hasher, & Zacks, in press, for an elaboration of this argument). On another front, it is possible that older adults' difficulty in inhibiting the continued processing of information designated as irrelevant or wrong could have an impact on their ability as jurors to comply with a judge's instructions to ignore testimony that that has been stricken from the record.[4]

[4] We thank R. Edward Geiselman for suggesting this implication of our findings to us.

## References

Allport, A. (1989). Visual attention. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 631–682). Cambridge, MA: MIT Press.

Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 583–612). New York: Academic Press.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph, 80,* 1–46.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger, III, & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309–330). Hillsdale, New Jersey: Erlbaum.

Camp, C. J., & McKitrick, L. A. (1989). The dialectics of forgetting and remembering across the adult lifespan. In D. A. Kramer & M. Bapp (Eds.), *Transformations in clinical and developmental psychology* (pp. 169–187). New York: Springer-Verlag.

Connelly, S. L., Hasher, L., & Zacks, R. T. (1991). Age and reading: The impact of distraction. *Psychology and Aging, 6,* 533–541.

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.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Geiselman, R. E., & Bagheri, B. (1985). Repetition effects in directed forgetting: Evidence for retrieval inhibition. *Memory & Cognition, 13,* 57–62.

Gerard, L. D., Zacks, R. T., Hasher, L., & Radvansky, G. A. (1991). Age deficits in retrieval: The fan effect. *Journal of Gerontology: Psychological Sciences, 46,* 131–136.

Giambra, L. M., & Howard, D. V. (1994, April). *Directed forgetting in older adults: Some data on inhibition.* Poster presented at the 1994 Cognitive Aging Conference, Atlanta, GA.

Hamm, V. P., & Hasher, L. (1992). Age and the availability of inferences. *Psychology and Aging, 7,* 56–64.

Hartman, M., & Hasher, L. (1991). Aging and suppression: Memory for previously relevant information. *Psychology and Aging, 6,* 587–594.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 163–169.

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* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.

Kane, M. J., Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Connelly, S. L. (1994). Inhibitory attentional mechanisms and aging. *Psychology and Aging, 9,* 103–112.

Keele, S. W., & Neill, W. T. (1979). Mechanisms of attention. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 9, pp. 3–47). New York: Academic Press.

Klein, R. M., & Taylor, T. L. (1994). Categories of inhibition with reference to attention. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 53–112). San Diego, CA: Academic Press.

Light, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology, 42,* 737–747.

MacLeod, C. M. (1989). Directed forgetting affects both direct and indirect tests of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 13–21.

May, C. P., Kane, M. J., & Hasher, L. (1995). Determinants of negative priming. *Psychological Bulletin, 118,* 35–54.

McDowd, J. M., & Oseas-Kreger, D. M. (1991). Aging, inhibitory processes, and negative priming. *Journal of Gerontology: Psychological Sciences, 46,* 340–345.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 361–394). Hillsdale, NJ: Erlbaum.

Neumann, E., Cherau, J. F., Hood, K. L., & Steinnagel, S. L. (1993). Does inhibition spread in a manner analogous to spreading activation? *Memory, 1,* 81–105.

Neumann, E., & DeSchepper, B. G. (1992). An inhibition based fan effect: Evidence for an active suppression mechanism in selective attention. *Canadian Journal of Psychology, 46,* 1–40.

Pavur, E. J., Comeaux, J. M., & Zeringue, J. A. (1984). Younger and older adults' attention to relevant and irrelevant stimuli in free recall. *Experimental Aging Research, 10,* 59–60.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34–50.

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153,* 652–654.

Stoltzfus, E. R., Hasher, L., & Zacks, R. T. (in press). Working memory and aging: Current status of the inhibitory view. In J. T. E. Richardson (Ed.), *Working memory and human cognition.* New York: Oxford University Press.

Stoltzfus, E. R., Hasher, L., Zacks, R. T., Ulivi, M. S., & Goldstein, D. (1993). Investigations of inhibition and interference in younger and older adults. *Journal of Gerontology: Psychological Sciences, 48,* 179–188.

Sullivan, M. P., & Faust, M. E. (1993). Evidence for identity inhibition during selective attention in old adults. *Psychology and Aging, 8,* 589–598.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 37(A),* 571–590.

Tipper, S. P. (1991). Less attentional selectivity as a result of declining inhibition in older adults. *Bulletin of the Psychonomic Society, 29,* 45–47.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale–Revised.* New York: Psychological Corporation.

Zacks, R. T., & Hasher, L. (1994). Directed ignoring: Inhibitory regulation of working memory. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 241–264). San Diego, CA: Academic Press.

Zacks, R. T., Radvansky, G., & Hasher, L. (1993). [Response times for younger and older adults in a modification of the Sternberg task]. Unpublished raw data.



Received June 27, 1994
Revision received December 28, 1994
Accepted January 25, 1995 ■