Effects of repetition and response deadline on item recognition in young and older adults

LEAH L. LIGHT
Pitzer College, Claremont, California

CHRISTIE CHUNG and REGINA PENDERGRASS
Claremont Graduate University, Claremont, California

and

JEFFREY C. VAN OCKER
Pitzer College, Claremont, California

In the present study, we examined the joint effects of aging, repetition, and response deadline in a plurality discrimination task. Young and older adults studied lists of unrelated singular and plural nouns, with half presented once (weak items) and half presented five times (strong items). Test lists contained old (same) nouns, plurality-reversed nouns (changed lures), and unstudied nouns (new lures), and the participants were asked to respond *old* only to same items. In Experiment 1, the participants were tested with both short and long deadlines. In Experiment 2, the tests were unpaced. In both experiments, repetition increased hit rates for young and older adults. Young adults tested with a long deadline or no deadline showed invariant (Experiment 1) or reduced (Experiment 2) false alarms to changed lures when the nouns were studied more often. Young adults tested with a short deadline and older adults tested with both long and short deadlines had increased false alarm rates for strong changed lures; without time pressure to respond, older adults did not have a significant increase in false alarms for changed lures. Implications of these results for theories of cognitive aging are explored.

According to dual-process models of memory, there are two processes, recollection and familiarity, that underlie recall and recognition (e.g., Atkinson & Juola, 1974; Diller, Nobel, & Shiffrin, 2001; Mandler, 1980; see Yonelinas, 2002, for a review). Although models differ somewhat in their construal of these, recollection is generally described as a conscious, attention-demanding process, with a slow rise time. Recollection entails conscious retrieval of specific episodic information, including perceptual details, the source of the information, temporal or spatial information, and emotions that accompany the events. Familiarity, on the other hand, is usually conceptualized as an unconscious, relatively automatic process that is recruited more rapidly.

On the dual-process account, word (item) recognition is based largely on familiarity when lures and targets are dissimilar. However, when lures are very similar, as is the case when singular and plural nouns (e.g., *kernel*) are studied and test lures are plurality-reversed nouns (e.g., *kernels*),

differentiation between targets and lures on the basis of familiarity is difficult unless recollection is invoked (e.g., Arndt & Reder, 2002; Curran, 2000; Hintzman, Curran, & Oppy, 1992; Malmberg, Holden, & Shiffrin, 2004; Norman, 2002; Rotello, Macmillan, & Van Tassel, 2000; Sheffert & Shiffrin, 2003). That is, participants can use a recall-to-reject strategy, recollecting studied items when confronted with familiar plurality-reversed lures; such recollection is diagnostic that plurality-reversed lures are new items and permits their rejection. Postulating a recall-to-reject strategy or some other recall-like process is needed to explain why *old/new* discrimination with highly similar lures often shows little or no improvement over a wide range of presentation frequencies at the same time that judged frequency is sensitive to number of repetitions. This phenomenon has been dubbed *registration without learning* by Hintzman et al. (1992). Single-process models have considerable difficulty in accounting for such effects, because they lack mechanisms for predicting that false alarms to dissimilar lures decrease with repetition of targets at study, whereas false alarms to similar lures do not (see Malmberg, Holden, & Shiffrin, 2004, for a discussion). In the research reported here, we examined the joint effects of aging, repetition, and response deadline on recognition in the plurality discrimination task.

There is a fair amount of consensus that familiarity-based mechanisms are relatively preserved in old age, whereas recollection is more adversely affected. We will

begin by briefly reviewing evidence from research on repetition effects that supports this view (for more complete summaries, see Light, Prull, La Voie, & Healy, 2000; Yonelinas, 2002; but see also Healy, Light, & Chung, 2005, for suggestions of age-related differences in familiarity). Jacoby (1999) investigated the influence of both repetition and response pressure in the context of aging. Participants saw a list in which some words appeared one, two, or three times and then heard a second list of words. At test, they were asked to respond *old* only to words that they had heard previously. When young adults were given long response deadlines, false alarms to previously studied words decreased with repetition. Ironically, false alarms to previously studied words increased with repetition for older adults, even with quite long deadlines. Thus, for the young adults, exclusion of seen words improved as a function of repetition, but the opposite was true for the older adults, in accord with an age-related deficit in the use of recollection to oppose the enhanced familiarity produced by repetition. Interestingly, the young adults also showed an increase in false alarms with repetition when forced to respond quickly; this finding accords well with the view that recollection has a longer rise time than does familiarity.

Similar ironic effects of repetition for older adults and for young adults tested with short response deadlines have been found for the Deese/Roediger–McDermott task (Benjamin, 2001; Budson, Daffner, Desikan, & Schacter, 2000; Kensinger & Schacter, 1999; Watson, McDermott, & Balota, 2004), as well as for the false-fame paradigm (Bartlett, Strater, & Fulton, 1991), for recognition of compound words (Jones & Jacoby, 2001), and for associative recognition (Light, Patterson, Chung, & Healy, 2004). Light et al. (2004) utilized a paradigm previously employed by Kelley and Wixted (2001), in which young adults studied lists of word pairs. Half of the pairs were presented once (*weak* pairs), and half were presented four times (*strong* pairs). After study, the participants received a recognition test that included strong and weak intact pairs (both words studied together), strong and weak rearranged pairs (both words studied, but with different partners), and new pairs. Kelley and Wixted found that with repetition of pairs at study, the young adults' hits to intact pairs increased, whereas their false alarms to rearranged pairs remained constant. In our study, hit rates increased with repetition for both young and older adults. We also found constancy of false alarm rates for weak and strong pairs (or even a reduction in false alarms for strong pairs), when the young adults did not need to respond quickly. However, the young adults tested with short response deadlines and the older adults tested with or without response deadlines had *higher* rates of false alarms for strong than for weak rearranged pairs.

Outcomes such as these are easily explained by the assumption that familiarity is less effectively opposed by recollection in older adults and in young adults pressured to respond quickly. In the present study, we asked whether these findings could be generalized to item recognition in the plurality discrimination task. In Experiment 1, sin-gular and plural nouns were presented once (weak items) or five times (strong items), and the participants were tested under both short- and long-deadline conditions. We predicted that without pressure to respond quickly, young adults would show the pattern of results observed in the several paradigms summarized above: higher hit rates for strong than for weak same-plurality nouns, but either no effect of repetition on false alarms to plurality-reversed lures or a decline in false alarms for strong plurality-reversed lures. We also predicted that young adults would show an increase in false alarms to plurality-reversed lures when tested with a short response deadline (Hintzman & Curran, 1994; Hintzman et al., 1992). Older adults, in contrast, should show ironic effects of repetition, manifested as higher rates of false alarms to strong than to weak plurality-reversed lures in both short- and long-deadline conditions. To preview our findings, this was what we found. In Experiment 2, responding was self-paced. The question of interest was whether giving older adults unlimited time to respond would encourage recruitment of recollection to oppose familiarity, yielding a pattern of results like that for young adults in the long-deadline condition in Experiment 1.

EXPERIMENT 1

Method

Participants. The young adults (27 women, 3 men) were students recruited from the Claremont colleges. The older adults (20 women, 8 men) resided in the Claremont area. All the participants were paid $40 for participation. The mean ages of the two groups were 19.43 years (*SD* = 1.68, range = 18–26) and 68.29 years (*SD* = 5.99, range = 60–80), respectively. Additional background information for these participants is given in Table 1. As is typically the case, reading and computation spans were higher in the young than in the older adults, although only the difference in reading span was significant in Experiment 1. Older adults often have higher vocabulary scores than do young adults, and the advantage here was also to the older adults, although the difference was not statistically significant. The data of 1 additional young adult were excluded because of a very high false alarm rate on new lures in the long-deadline condition. The data from 23 additional older adults were also excluded. Thirteen did not meet a criterion of 50% for reaction times within the response window in each deadline condition of the speeded recognition task, 6 did not complete the experiment, 3 fell below criterion on one or more auxiliary tasks (described below), and the data of 1 were lost due to equipment failure. Analyses including the data from those participants who were excluded because they did not respond frequently enough within the response window or performed poorly on the auxiliary tasks led to conclusions consistent with those reported here.

Materials. We selected 240 low-frequency nouns and their plural forms from the MRC Psycholinguistic Database (Coltheart, 1981) to serve as experimental items. These words met three criteria. First, the plural version of each noun could be created by adding *s*. Second, the dominant meaning of each item was the same in its singular and plural forms. Third, the singular and plural forms were similar in frequency, with singular and plural nouns having mean frequencies of 2.12 and 1.87 per million, respectively. No word was longer than 12 letters. The words were assigned randomly to lists for each pair of participants (1 young and 1 old). In each study list, each of 10 singular and 10 plural nouns was presented once (weak items), and each of 10 singular and 10 plural nouns was presented five times (strong items); in addition, there were 2 buffer words at the begin-

Table 1
Means for Young and Older Participants on Selected Measures
(With Standard Deviations)

Measure	Experiment 1				Experiment 2			
	Young		Old		Young		Old	
	M	SD	M	SD	M	SD	M	SD
Years of education	13.23	1.25	15.75**	2.63	13.93	1.51	14.10	2.70
Subjective health rating[a]	8.37	1.07	8.04	1.17	8.33	1.50	7.69	1.53
Word fluency[b]	14.74	3.63	13.92	3.21	15.02	3.17	13.29	3.93
Vocabulary[c]	17.63	3.62	18.96	3.75	16.48	3.14	17.21	3.83
Reading span[d]	3.20	1.19	2.46*	1.11	3.48	1.38	2.45**	1.18
Computation span[d]	3.17	1.09	2.64	0.95	3.31	1.39	2.62*	1.15

[a]1 = poor; 10 = excellent.   [b]Mean number of words generated beginning with the letters C, F, and L in 60 sec.   [c]Maximum = 25.   [d]Maximum = 7.   *$p < .05$.   **$p < .01$.

ning of each study list and 2 at the end, so each study list contained a total of 124 words. Within a study list, the words in the strong condition were spaced so that each word occurred once in each fifth of the list and repetition of the same word was never consecutive. Furthermore, no more than 2 singular words or plural words were presented consecutively.

The experimental design called for three types of test items: same plurality (old), plurality reversed (changed), and new. The old words were presented on the study and test lists in the same form (e.g., singular on the study list and singular on the test list). The changed words had their plurality reversed between study and test (e.g., singular on the study list and plural on the test list). The new words were presented only on the test list. Each 60-word test list contained 20 same items (10 weak and 10 strong), 20 changed items (10 weak and 10 strong), and 20 new words (10 singular and 10 plural). No more than 2 words of each word type (singular or plural, old or new) appeared consecutively on a test list.

Design and Procedure. In each of two 2-h sessions separated by a maximum of 7 days, the participants began by practicing use of the 6-point confidence rating scale that would be used later for the recognition test (1 = *sure new* to 6 = *sure old*). Two practice study–test blocks and two experimental study–test blocks followed. For the practice blocks, the long deadline always preceded the short deadline. For the experimental blocks, the order of deadline conditions was counterbalanced across participants for each age group and was different in each session. For all practice and experimental study lists, presentation rate was 2 sec per word, with the test following immediately after list study. For each test item, the participants indicated their confidence that the word was old.

In the first session, the participants read two pages of detailed instructions about the main task. The instructions clearly described the three types of test words that they would experience. In particular, people were told that the test item types were *old* words, *changed* words, and *new* words, that they should respond *old* to old words and *new* to both changed and new words, that some words on a study list would be repeated several times within that list, and that words appeared only on a single study–test block. Finally, they were instructed to wait for the response cue and then to respond as quickly as possible at test. They next completed two practice blocks (long and short deadlines) and two experimental blocks. In the second session, the participants again completed the training, reviewed the detailed instructions for the experimental task, and completed two practice and two experimental blocks.

In the short-deadline condition, the participants saw a focal cue (+ + + + +) for 500 msec, followed by presentation of a test word. A response cue (*****) appeared 400 msec after presentation of each test word. At that time, the participants had 600 msec to respond. Thus, a total of 1,000 msec elapsed from item presentation until the response deadline lapsed. A beep followed, which signaled a 1,000-msec intertrial interval (blank screen). In the long-deadline

condition, the response cue did not appear until after the test word had been presented for 2,400 msec. Again, the participants had 600 msec to respond (3,000-msec total response time). If the participants responded before the response cue, they were reminded to wait for that cue before pressing a key. If the participants did not respond in time, a beep signaled the end of the response window.

A small battery of secondary measures was also administered to characterize the samples of young and older participants. Speed and executive function were assessed with the Trail Making Test (Parts A and B; Reitan, 1992). Verbal fluency was measured with the Word Fluency Test (Benton & Hamsher, 1976). A locally developed letter identification task in which the participants circled as many letter *A*s as they could in 3 min was used as a measure of perceptual–motor speed. Finally, working memory was assessed with computational span and reading span tasks (Salthouse, 1992). These tasks were interspersed with the word recognition study–test blocks in order to reduce interference between lists. The last task for both groups in the second session was a 25-item version of the Nelson and Denny (1961) vocabulary test.

Results

We began by examining the percentage of responses that fell within the long- and short-deadline response windows for the young and older adults. The young adults made .92 of all responses within the response window for both short and long deadlines, whereas the older adults had .82 and .86 of their responses within the response window for the two deadlines. A 2 (age) × 2 (deadline) ANOVA indicated that main effects of both age and deadline were significant [$F(1,56) = 43.01$, $p < .001$, $\eta^2 = .43$, and $F(1,56) = 5.06$, $p < .05$, $\eta^2 = .08$, respectively]. The interaction between these variables was marginally significant [$F(1,56) = 3.27$, $p < .08$; $\eta^2 = .06$]. Given these effects, we carried out parallel analyses on all the responses, as well as on just those responses that fell within the designated time windows. Conclusions drawn from these analyses were identical, so to conserve space we will report only the analyses of on-time responses. An α of .05 was used for all the analyses.

Responses of 4, 5, or 6 for studied nouns were treated as hits, whereas responses of 4, 5, or 6 to plurality-reversed and new lures were treated as false alarms (see Figure 1). Separate 2 (age: young vs. older) × 2 (deadline: short vs. long) × 2 (strength: one vs. five repetitions) ANOVAs were carried out for hits and false alarms on plurality-reversed nouns. Overall, there were more hits in the long-
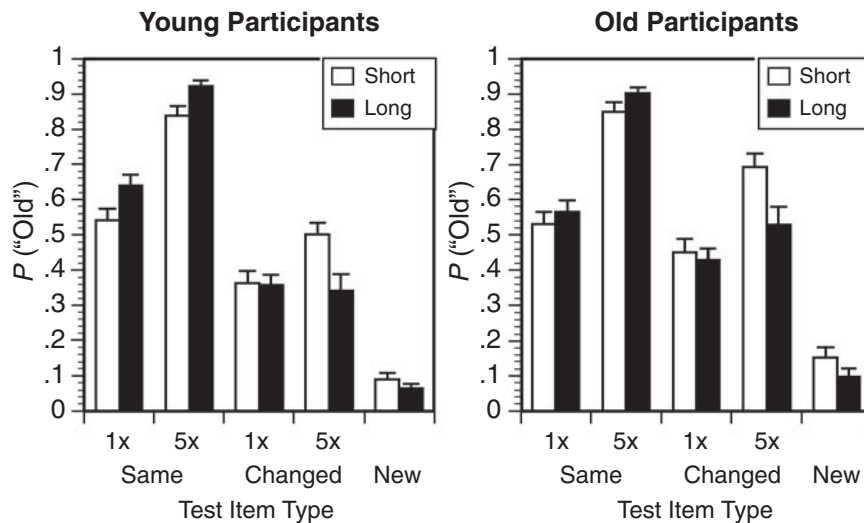
Young Participants   **Old Participants**

Figure 1. Mean proportions of *old* responses (+*SE*) as a function of age, item type, repetition, and response deadline in Experiment 1.

($M = .76$) than in the short- ($M = .69$) deadline condition [$F(1,56) = 12.35, p = .001; \eta^2 = .18$]. Hits increased with strength [$F(1,56) = 267.93, p < .001; \eta^2 = .83$]. Across age groups, there were more hits for strong words ($M = .88$) than for weak words ($M = .57$). Neither the main effect of age [$F(1,56) = 1.23, p = .27; \eta^2 = .02$] nor any of the interactions between factors (all $Fs \leq 1.54$) reached significance. Overall, the false alarm rate on changed lures was higher for the older adults ($M = .53$) than for the young adults ($M = .39$) [$F(1,56) = 9.88, p < .01; \eta^2 = .15$]. There were more false alarms at the short deadline ($M = .50$) than at the long deadline ($M = .42$) [$F(1,56) = 22.70, p < .001; \eta^2 = .29$]. False alarms increased with strength ($Ms = .40$ and $.52$) [$F(1,56) = 33.71, p < .001; \eta^2 = .38$]. These main effects were qualified by significant two-way interactions of strength with age [$F(1,56) = 7.68, p < .01; \eta^2 = .12$] and strength with deadline [$F(1,56) = 18.80, p < .001; \eta^2 = .25$].

Although the three-way interactions of age, deadline, and strength were not significant for either hits or false alarms on changed lures (both $Fs < 1$), our a priori hypotheses dictated that separate analyses be carried out to examine the joint effects of deadline and strength for each age group. For hits, the young adults had significant main effects of both deadline and strength [$F(1,29) = 15.36, p < .001, \eta^2 = .35$, and $F(1,29) = 137.38, p < .001, \eta^2 = .83$, respectively]. The young adults had more hits at the long deadline ($M = .78$) than at the short deadline ($M = .69$) and more hits on strong words ($M = .88$) than on weak words ($M = .59$). The young adults made more false alarms on changed lures in the short- ($M = .43$) than in the long- ($M = .35$) deadline condition [$F(1,29) = 11.48, p < .01; \eta^2 = .28$], and their false alarms increased with strength ($Ms = .36$ and $.42$) [$F(1,29) = 4.47, p < .05; \eta^2 = .13$]. These main effects were qualified by a significant interaction between deadline and strength [$F(1,29) =$

$11.88, p < .01; \eta^2 = .29$]. False alarms increased with repetition at the short deadline ($Ms = .36$ and $.50$) [$t(29) = -4.28, p < .001$], but not at the long deadline ($Ms = .36$ and $.34$) [$t(29) = 0.40, p = .69$].

The same analyses were carried out for the older adults, who also had more hits on strong items ($M = .88$) than on weak items ($M = .55$) [$F(1,27) = 130.44, p < .001; \eta^2 = .83$] and more false alarms in the short- ($M = .57$) than in the long- ($M = .48$) deadline condition [$F(1,27) = 11.19, p < .01; \eta^2 = .29$]. There were also more false alarms on strong lures ($M = .61$) than on weak lures ($M = .44$) [$F(1,27) = 38.29, p < .001; \eta^2 = .59$]. However, these main effects were qualified by a significant interaction [$F(1,27) = 7.48, p = .01; \eta^2 = .22$]. There were more false alarms on strong changed items at both deadlines, but, not surprisingly, the effect of strength was greater at the short deadline [$t(27) = -7.27, p < .001$] than at the long deadline [$t(27) = -2.36, p < .05$]. At the short deadline, the false alarm rates were $.69$ and $.45$ for strong and weak changed lures, respectively, and at the long deadline, the corresponding values were $.53$ and $.43$.

A two (age) $\times$ 2 (deadline) ANOVA on new lures revealed that the older adults made more false alarms ($M = .13$) than did the young adults ($M = .08$) on this class of items [$F(1,56) = 4.83, p < .05; \eta^2 = .08$]. In addition, there were more false alarms in the short- ($M = .13$) than in the long- ($M = .08$) deadline condition [$F(1,56) = 7.90, p < .01; \eta^2 = .12$]. The interaction between these two factors was not significant [$F(1,56) = 1.13, p = .29; \eta^2 = .02$]. To investigate whether the significance of age differences in hits and false alarms on changed lures was due merely to age differences in false alarm rates on new lures, the hit rate and the false alarm rate for changed items were corrected by subtracting out the false alarms to new items for the appropriate cell, and the analyses reported above were repeated for the corrected measures.

For hits, there was now an age effect; the young adults ($M = .66$) had significantly more hits than did the older adults ($M = .58$) [$F(1,56) = 5.47, p < .05; \eta^2 = .09$]. In the analyses of the young adults' hits and false alarms on plurality-reversed lures, correcting for baseline guessing by subtracting out new lure false alarm rates and repeating the ANOVAs yielded identical results. Correcting the older adults' hits and false alarms on changed items to plurality-reversed lures and repeating the ANOVAs made the main effect of deadline for hits significant, with more hits in the long- ($M = .63$) than in the short- ($M = .53$) deadline condition [$F(1,27) = 7.51, p = .01; \eta^2 = .22$]; the main effect of deadline for false alarms to plurality-reversed lures was not significant in the analysis of corrected scores [$F(1,27) = 1.04, p = .32; \eta^2 = .04$]. The critical finding of higher false alarms by the older adults on changed lures for strong items at both deadlines was, of course, unaffected by this correction, which does not alter the pattern or magnitude of the within-participants effects.

Discussion

The results of Experiment 1 are easily summarized. Both the young and the older adults exhibited higher hit rates for strong than for weak same-plurality test items, regardless of deadline. The young adults tested at the shorter deadline and the older adults tested at both deadlines had higher false alarms for plurality-reversed lures whose complements were studied five times than for those presented once, but no strength effect was seen for plurality-reversed lures when the young adults were tested at the longer deadline. These outcomes are consistent with earlier work in which the plurality discrimination task was used (e.g., Hintzman & Curran, 1994; Hintzman et al., 1992), as well as with predictions made by dual-process models of memory. Our results also are relevant for questions of registration without learning. The increase in false alarms that we observed for repeated nouns at short deadlines shows that repetitions are registered without appealing to judgments of frequency, and our study thus extends the range of manipulations that can be used to investigate registration without learning.[1]

Further implications of these findings will be discussed in conjunction with those of Experiment 2.

EXPERIMENT 2

In Experiment 1, the young adults showed changes with deadline in the effects of repetition on false alarms for plurality-reversed lures. It could be argued that the older adults' performance might have been equivalent to that of the young adults shifted by some temporal offset and that we simply did not include a deadline condition for the older adults that was sufficiently long to demonstrate this. The primary goal of Experiment 2 was to investigate the question of whether removing time constraints altogether would eliminate age differences in strength effects for false alarms to plurality-reversed lures, as might be expected if the age differences in effects of repetition ob-

served in Experiment 1 were simply the result of an injudicious selection of deadline conditions.

Method

Participants. The young adults (22 women, 7 men) were students from the Claremont colleges. The older adults (23 women, 6 men) resided in the Claremont area. Mean ages for the two groups were 20.24 years ($SD = 2.23$, range = 18–29) and 69.07 years ($SD = 5.12$, range = 60–80), respectively. The participants were paid $20 for participation in one session lasting about 2 h. Additional background information for these participants is given in Table 1. The older adults had lower reading and computation spans than did the younger adults, although as in Experiment 1, an age difference in vocabulary was not found. Data from an additional 1 young and 4 older adults were excluded from analysis. The young adult had unusually low performance on our vocabulary test. Two older adults did not complete the plurality discrimination task; 1 made an excessive number of false alarms to new items, and 1 was an outlier on one of the secondary tasks.

Materials. The words used in Experiment 1 were also used in Experiment 2. For practice, the participants studied 20 words, 10 presented once (weak nouns), and 10 presented five times (strong nouns). Following study, the participants received a confidence-rated recognition test that consisted of 10 same words (5 weak and 5 strong), 10 plurality-reversed words (5 weak and 5 strong), and 10 new words. After the practice session, the participants studied two lists of 44 words. Twenty of the words were presented once (weak items), 20 were presented five times (strong items), and 4 served as buffers (2 at each end of the list). After each study list, the participants received a recognition test that consisted of 20 same words (10 weak, 10 strong), 20 plurality-reversed words (10 weak, 10 strong), and 20 new words.

Procedure. The procedure in Experiment 2 was identical to that in Experiment 1, except that responses were self-paced and testing required only one session that lasted for about 2 h.

Results and Discussion

Hits increased with repetition for both the young and the older adults (see Figure 2). False alarms on plurality-reversed lures decreased with repetition for the young adults. However, the older adults had numerically higher
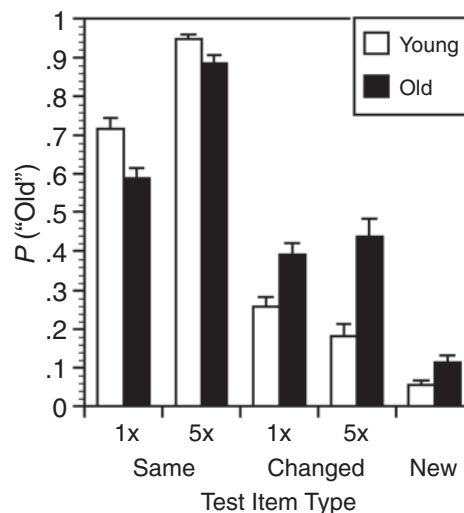

Figure 2. Mean proportions of *old* responses (+*SE*) as a function of age, item type, and repetition in Experiment 2.

false alarms on strong than on weak plurality-reversed lures. Separate ANOVAs that confirmed these visual impressions were performed on hits to same-plurality words, false alarms to plurality-reversed lures, and false alarms to new words.

A 2 (age: young vs. older) × 2 (strength: one vs. five repetitions) ANOVA conducted on hits yielded significant main effects of strength [$F(1,56) = 221.84$, $p < .001$; $\eta^2 = .80$] and age [$F(1,56) = 10.31$, $p < .01$; $\eta^2 = .16$]. Across age groups, there were more hits for strong words ($M = .92$) than for weak words ($M = .64$). For the young adults, the mean hit rates were .71 and .95, for weak and strong words, respectively. For the older adults, these means were .58 and .89, respectively.

A 2 (age) × 2 (strength) ANOVA on false alarms to plurality-reversed nouns revealed a significant main effect of age [$F(1,56) = 18.37$, $p < .001$; $\eta^2 = .25$]. The older adults made more false alarms on changed words ($M = .41$) than did the young adults ($M = .22$). However, this main effect must be interpreted in the context of an interaction between age and strength [$F(1,62) = 5.44$, $p < .05$; $\eta^2 = .09$]. The young adults had significantly fewer false alarms on plurality-reversed nouns studied five times ($M = .19$) than on those studied once ($M = .26$) [$t(28) = 2.55$, $p < .05$]. The older adults, however, had more false alarms on plurality-reversed nouns studied five times ($M = .43$) than on those studied once ($M = .38$), although this difference was not significant [$t(28) = -1.07$, $p = .29$].

Finally, the older adults made somewhat more false alarms to new lures ($M = .10$) than did the young adults ($M = .06$) [$t(56) = -1.94$, $p < .06$]. Correcting the hit rates and false alarm rates on plurality-reversed lures by subtracting out the new item false alarm rate yielded the same effects as those already described: The young adults had a dual benefit of repetition in increased hit rates and reduced false alarm rates on plurality-reversed nouns, whereas the older adults had only the single benefit of repetition on hit rates.

GENERAL DISCUSSION

The questions addressed by the experiments reported here were whether young and older adults would benefit equally from increased study opportunities in the plurality discrimination task and whether the effects of repetitions would be moderated by imposition of a response deadline at test. In Experiment 1, both the young and the older adults endorsed previously studied nouns at a higher rate when these were presented five times rather than just once, and this was true for both long and short response deadlines. However, the young adults had different patterns of false alarms on plurality-reversed lures at long and short deadlines: an increase with greater study frequency of alternate forms at the short deadlines and no effect of strength at the longer deadline. The older adults had higher false alarms on plurality-reversed lures studied more often in their alternate forms at *both* deadlines. In Experiment 2, without time pressure to respond, hit rates for previously

studied nouns demonstrated a strength effect for both the young and the older adults. The young adults had lower false alarms on plurality-reversed lures whose alternate forms were studied five times, rather than once. The older adults continued to have higher false alarm rates for strong than for weak plurality-reversed lures, but without time pressure to respond, this effect was not significant. These results are readily explained by dual-process theories of recognition. Repetition of nouns at study increases item familiarity, and this process is relatively unaffected by normal aging. Recollection, too, is enhanced by repetition and works to oppose familiarity in young adults given adequate time to respond at test. Both aging and time pressure reduce recollection, so that the familiarity of strong plurality-reversed lures is not countered by retrieval of plurality information for studied items.[2]

The young adult findings are generally consistent with those observed with the plurality discrimination paradigm elsewhere (e.g., Hintzman & Curran, 1994; Hintzman et al., 1992). The decline in the young adults' false alarms on strong plurality-reversed lures when time pressure to respond on the test was eliminated, suggestive of registration *with* learning (Malmberg, Holden, & Shiffrin, 2004), may, however, seem unexpected, because some studies have shown little or no benefit of repetition for false alarms in this paradigm over a wide range of frequencies. Still, some studies have reported declining false alarms for lures whose alternates were studied more often (e.g., compare Experiments 2, 3, and 4 in Hintzman et al., 1992). When participants are instructed to remember noun plurality, *old/new* discrimination may improve with repetition, although the benefit appears to be confined to the first few presentations. Thus, Hintzman and Curran (1995) found a decrease in false alarms for plurality-reversed nouns whose complements had been studied three times, but this did not extend to nouns studied eight times. It is possible that our procedures, which alerted people to the importance of noun plurality, contributed to our findings for the young adults in Experiment 2.

Our results suggest that there is a gradient of response pressure, moving from the short- and long-deadline conditions in Experiment 1 to the no-deadline condition in Experiment 2, that is consistent with the recruitment of recall-to-reject processes over time in young adults. We are, however, somewhat loathe to overemphasize this finding, because we observed a different set of outcomes in an associative recognition study. That is, young adults in Light et al. (2004) had fewer false alarms for strong than for weak rearranged lures when tested under a long-deadline condition but similar levels of false alarms for these lures when no response deadline was imposed. It is unclear at this time whether these differences across experiments and across paradigms merely reflect noise or whether they reflect differing underlying processes; we therefore believe that it is premature to draw strong conclusions about these aspects of our findings.

What are novel and, we believe, of theoretical import here are our observations with respect to aging in the plurality discrimination task. As was noted earlier, in

the Deese/Roediger–McDermott paradigm, young adults show dual benefits of repetition in correct identification of previously studied words and in false alarms to strongly associated targets, regardless of whether repetition is varied within a single list followed by a test or whether multiple study–test cycles are used (Benjamin, 2001; Budson et al., 2000; Kensinger & Schacter, 1999; Watson et al., 2004). In contrast, older adults have shown an increase in false alarms to target lures (Benjamin, 2001), no change in false alarm rates to target lures with multiple presentations (Kensinger & Schacter, 1999; Watson et al., 2004), or a fluctuating false alarm rate to target lures over study–test cycles (Budson et al., 2000). Associative recognition studies (Light et al., 2004) also suggest that older adults are more likely to produce conjunction errors on rearranged lures than are young adults and not to benefit from repeated study of list pairs. Our findings of parallel effects in the plurality discrimination task suggest that it may prove more fruitful from a theoretical vantage to treat all of these phenomena as instances of reduced ability to differentiate between highly similar events even after many experiences with the study materials. Doing so highlights the fact that elevated rates of false alarms on similar lures can occur in older adults in a variety of paradigms involving associative recognition and memory not only for lists of strongly associated or phonologically related words (Sommers & Huff, 2003; Watson, Balota, & Sergent-Marshall, 2001) in the Deese/Roediger–McDermott paradigm, but also for lists of words that are ostensibly unrelated, when recognition of those words requires differentiation between old words and highly similar lures. Endorsement of target items as *old* in the Deese/Roediger–McDermott paradigm is generally treated as *false memory* or as evidence for the existence of *memory illusions*, to which older adults are seen as more susceptible. Noting commonalities across paradigms helps to focus attention on issues of whether there are common underlying mechanisms that produce the phenotypically similar results observed when recognition test lures are similar to studied targets in any of a number of different ways (see, e.g., Norman & O'Reilly, 2003).

We have interpreted our findings in terms of dual-process models of recognition memory, and as we have already noted, these have also been invoked to explain the influences of repetition, deadline, and aging in associative recognition (e.g., Cleary, Curran, & Greene, 2001; Jones & Jacoby, 2001; Kelley & Wixted, 2001; Light et al., 2004), as well as in the Deese/Roediger–McDermott paradigm (e.g., Benjamin, 2001; Kensinger & Schacter, 1999; Seamon et al., 2002; Watson et al., 2004). Alternative single-process models for plurality discrimination have also been proposed by Malmberg, Holden, and Shiffrin (2004). One of these models assumes a two-step retrieval and decision process. In the first step, the plurality of the test item is not considered when the test item is matched to items in memory, because plurality is not diagnostic of whether the item is old or new. Only items exceeding a criterial level of familiarity are considered in a second step in which plurality of the test item is matched to po-

tential list items in memory.[3] Poor discrimination between studied targets and similar lures could arise because short response deadlines compromise formation of a specific memory probe or limit execution of the second phase of the retrieval and decision process. Aging effects could also be attributed to truncation of the retrieval and decision process on at least some proportion of test trials because utilization of specific test probes is effortful (Shiffrin & Steyvers, 1997). Two-step models of associative recognition that consider item familiarity in the initial step and associative familiarity in a second step may also be capable of explaining age, repetition, and deadline effects in that task (Light et al., 2004) and could, in principle, be applied to the Deese/Roediger–McDermott paradigm as well, but the details of such models have not, to our knowledge, been worked out, and doing so is beyond the scope of this article.

It is also possible that, instead of or in addition to less effective retrieval, deficits in encoding plurality could lead to less complete or less accurate memory traces in older adults. Malmberg, Zeelenberg, and Shiffrin (2004) have recently demonstrated that some phenomena thought to require recollection can be explained in terms of encoding deficits. Naveh-Benjamin and his colleagues (Naveh-Benjamin, 2000; Naveh-Benjamin, Hussain, Guez, & Bar-On, 2003) have argued that older adults have a deficit in remembering associative information even when matched in performance with young adults on item memory. One way of thinking about the present results is that in the context of experiments such as this one, participants encode each study word as a singular noun, along with a tag that represents the presence or absence of a plural morpheme, rendering the recognition test functionally equivalent to a source memory test. Within such a framework, older adults' increased confusion of studied nouns with their plurality-reversed complements could be interpreted as a failure to encode or to retrieve information needed for source identification and could be construed as a problem in binding morphemes representing nouns with the plural morpheme. Such a conceptualization would be most compatible with a lexical-semantic system in which only the singular forms of nouns are stored and pluralization takes place via rule application; there is currently debate about whether this is the case or whether both singular and plural forms are stored in a connectionist network (e.g., Marchman, Plunkett, & Goodman, 1997). From the latter perspective, our findings would suggest that older adults have impaired encoding of the details of experienced episodes, leading to a deficit in item memory that is most easily detected when test lures must be differentiated from highly similar studied items (see also Gallo & Roediger, 2003).[4]

It remains for us to explain, from an encoding viewpoint, why older adults do not have a dual benefit of repetition, whereas young adults do. Let us assume that on each trial, the amount of information stored or its accuracy is greater for young than for older adults (e.g., Malmberg, Zeelenberg, & Shiffrin, 2004; see Li, 2002, for a neurocognitive model). Let us assume further that each presen-

tation of a noun after the first results not only in registration of that event, but also in a reminder of the previous study trial(s) for that event (Hintzman, 2004). If recursive reminding is an effortful recollective process, rather than a simple matching operation, it would not be surprising if older adults store less new information about repeated episodes on later presentations than do young adults or if their later encodings tend to perpetuate faulty memories, rather than correcting storage errors. Failure to use recursive reminding or more error-prone initial encoding, followed by failure to correct repetitions on later occasions, could also explain why young, but not older, adults show a benefit from spaced repetitions on an exclusion task (Benjamin & Craik, 2001). Although impoverished encoding may contribute to the effects we observed in older adults, it is unlikely that an encoding account can handle all of our results without additional machinery. Finding that repetition has different effects at long and short deadlines in young adults, for instance, is hard to reconcile with any model that considers only encoding and does not include dynamic processes operative at retrieval (Jones & Jacoby, 2001).

Our data do not permit us to resolve questions about the locus of the different ways in which repetition and deadline affect performance in young and older adults on the plurality discrimination task. They do, however, suggest some potentially useful ways to investigate the contributions of encoding and retrieval processes in this task.

REFERENCES

ARNDT, J., & REDER, L. M. (2002). Word frequency and receiver operating characteristic curves in recognition memory: Evidence for a dual-process interpretation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 830-842.

ATKINSON, R. C., & JUOLA, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology: Vol. 1. Learning, memory, and thinking* (pp. 243-293). San Francisco: Freeman.

BARTLETT, J. C., STRATER, L., & FULTON, A. (1991). False recency and false fame of faces in young adulthood and old age. *Memory & Cognition*, **19**, 177-188.

BENJAMIN, A. S. (2001). On the dual effects of repetition on false recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 941-947.

BENJAMIN, A. S., & CRAIK, F. I. M. (2001). Parallel effects of aging and time pressure on memory for source: Evidence from the spacing effect. *Memory & Cognition*, **29**, 691-697.

BENTON, A. L., & HAMSHER, K. D. (1976). *Multilingual aphasia exam.* Iowa City: University of Iowa.

BUDSON, A. E., DAFFNER, K. R., DESIKAN, R., & SCHACTER, D. L. (2000). When false recognition is unopposed by true recognition: Gist-based memory distortion in Alzheimer's disease. *Neuropsychology*, **14**, 277-287.

CLEARY, A. M., CURRAN, T., & GREENE, R. L. (2001). Memory for detail in item versus associative memory. *Memory & Cognition*, **29**, 413-423.

COLTHEART, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, **33A**, 497-505. Available online at www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm.

CURRAN, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, **28**, 923-938.

DILLER, D. E., NOBEL, P. A., & SHIFFRIN, R. M. (2001). An ARC–REM model for accuracy and response time in recognition and recall. *Jour-*

nal of Experimental Psychology: Learning, Memory, & Cognition, **27**, 414-435.

GALLO, D. A., & ROEDIGER, H. L., III (2003). The effects of associations and aging on illusory recollection. *Memory & Cognition*, **31**, 1036-1044.

HEALY, M. R., LIGHT, L. L., & CHUNG, C. (2005). Dual-process models of associative recognition in young and older adults: Evidence from receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 768-788.

HINTZMAN, D. L. (2004). Judgment of frequency versus recognition confidence: Repetition and recursive reminding. *Memory & Cognition*, **32**, 336-350.

HINTZMAN, D. L., & CURRAN, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory & Language*, **33**, 1-18.

HINTZMAN, D. L., & CURRAN, T. (1995). When encoding fails: Instructions, feedback, and registration without learning. *Memory & Cognition*, **23**, 213-226.

HINTZMAN, D. L., CURRAN, T., & OPPY, B. (1992). Effects of similarity and repetition on memory: Registration without learning? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 667-680.

JACOBY, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 3-22.

JONES, T. C., & JACOBY, L. L. (2001). Feature and conjunction errors in recognition memory: Evidence for dual-process theory. *Journal of Memory & Language*, **45**, 82-102.

KELLEY, R., & WIXTED, J. T. (2001). On the nature of associative information in recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 701-722.

KENSINGER, E. A., & SCHACTER, D. L. (1999). When true memories suppress false memories: Effects of ageing. *Cognitive Neuropsychology*, **16**, 399-415.

LI, S.-C. (2002). Connecting the many levels and facets of cognitive aging. *Current Directions in Psychological Science*, **11**, 38-43.

LIGHT, L. L., PATTERSON, M. M., CHUNG, C., & HEALY, M. R. (2004). Effects of repetition and response deadline on associative recognition in young and older adults. *Memory & Cognition*, **32**, 1182-1193.

LIGHT, L. L., PRULL, M. W., LA VOIE, D. J., & HEALY, M. R. (2000). Dual-process theories of memory in old age. In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive aging* (pp. 238-300). New York: Oxford University Press.

MALMBERG, K. J., HOLDEN, J. E., & SHIFFRIN, R. M. (2004). Modeling the effects of repetitions, similarity, and normative word frequency on old–new recognition and judgments of frequency. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 319-331.

MALMBERG, K. J., ZEELENBERG, R., & SHIFFRIN, R. M. (2004). Turning up the noise or turning down the volume? On the nature of the impairment of episodic recognition memory by Midazolam. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 540-549.

MANDLER, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, **87**, 252-271.

MARCHMAN, V. A., PLUNKETT, K., & GOODMAN, J. (1997). Overregularization in English plural and past tense inflectional morphology: A response to Marcus (1995). *Journal of Child Language*, **24**, 767-779.

NAVEH-BENJAMIN, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 1170-1187.

NAVEH-BENJAMIN, M., HUSSAIN, Z., GUEZ, J., & BAR-ON, M. (2003). Adult age differences in episodic memory: Further support for an associative-deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 826-837.

NELSON, M. J., & DENNY, E. C. (1961). *Examiner's manual: The Nelson–Denny Reading Test.* Boston: Houghton Mifflin.

NORMAN, K. A. (2002). Differential effects of list strength on recollection and familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 1083-1094.

NORMAN, K. A., & O'REILLY, R. C. (2003). Modeling hippocampal and

neocortical contributions to recognition memory: A complimentary-learning-systems approach. *Psychological Review*, **110**, 611-646.

REDER, L. M., NHOUYVANISVONG, A., SCHUNN, C. D., AYERS, M. S., ANGSTADT, P., & HIRAKI, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember–know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 294-320.

REITAN, R. M. (1992). *Trail Making Test: Manual for administration and scoring*. South Tucson, AZ: Reitan Neuropsychology Laboratory.

ROTELLO, C. M., MACMILLAN, N. A., & VAN TASSEL, G. (2000). Recall-to-reject in recognition: Evidence from ROC curves. *Journal of Memory & Language*, **43**, 67-88.

SALTHOUSE, T. A. (1992). Influence of processing speed on adult age differences in working memory. *Acta Psychologica*, **79**, 155-170.

SEAMON, J. G., LUO, C. R., SCHWARTZ, M. A., JONES, K. J., LEE, D. M., & JONES, S. J. (2002). Repetition can have similar or different effects on accurate and false recognition. *Journal of Memory & Language*, **46**, 323-340.

SHEFFERT, S. M., & SHIFFRIN, R. M. (2003). Auditory registration without learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 10-21.

SHIFFRIN, R. M., & STEYVERS, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, **4**, 145-166.

SOMMERS, M. S., & HUFF, L. M. (2003). The effects of age and dementia of the Alzheimer's type on phonological false memories. *Psychology & Aging*, **18**, 791-806.

WATSON, J. M., BALOTA, D. A., & SERGENT-MARSHALL, S. D. (2001). Semantic, phonological, and hybrid veridical and false memories in healthy older adults and in individuals with dementia of the Alzheimer type. *Neuropsychology*, **15**, 254-267.

WATSON, J. M., MCDERMOTT, K. B., & BALOTA, D. A. (2004). Attempting to avoid false memories in the Deese/Roediger–McDermott paradigm: Assessing the combined influence of practice and warnings in young and old adults. *Memory & Cognition*, **32**, 135-141.

YONELINAS, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language*, **46**, 441-517.

NOTES

1. We are grateful to Kenneth Malmberg for noting this point.

2. Both Rotello et al. (2000) and Arndt and Reder (2002) have reported fairly linear receiver operating characteristics (ROCs) relating hits on same-plurality nouns to false alarms on reversed-plurality nouns in probability space. Such ROCs have been taken as evidence for pure threshold models of recognition. We examined the functions relating hits to false alarms for plurality-reversed lures for each combination of age, repetition, and deadline in our two experiments. In every case, the r^2 for the linear function in *p*-space was at least .91, and a linear function described the data better than did a quadratic for 11 of 12 conditions. Calculation of *y*- and upper *x*-intercepts revealed that, in general, these behaved in a way that would be predicted if the young adults had better recall-to-accept and better recall-to-reject than did the older adults; repetition produced sensible changes in these (higher *y*- and lower *x*-intercepts), and short deadlines also yielded lower *y*- and higher upper *x*-intercepts. All of this is compatible with a pure high-threshold model of recognition in the plurality discrimination task. Unfortunately, in 11 of 12 cases, the fit of the ROC was improved by adding a quadratic component, consistent with dual-process models that include a familiarity process. Our findings should, therefore, not be interpreted as ruling out a role for familiarity in this task, but without explicit modeling of the data (a task beyond the scope of the present project), we cannot comment on this issue in more detail.

3. This model did not capture all aspects of the data in Malmberg, Holden, and Shiffrin's (2004) study (e.g., an initial increase in false alarms to plurality-reversed lures with repetition, followed by a flattening of the function relating repetition to false alarms for these similar lures). Because we included only two levels of repetition in our study, we cannot comment on this issue, except to note that the relationship between repetition and false alarms to similar lures varied with age group and deadline.

4. Whether such models can actually accommodate the present results is uncertain, and predictions depend on specific assumptions about cognitive architecture and encoding processes. For instance, in the source of activation confusion (SAC) model (Reder et al., 2000), memory consists of word nodes, context nodes, and links between them. Encoding leads to formation of an associative link between a word node and a context node. If singular and plural words are represented by separate nodes, repetitions might strengthen associations between only the node corresponding to the studied word and its associative link with a context node or between both singular and plural word nodes and their associations to context. In the former case, it is unclear why similar targets and dissimilar targets should lead to different false alarm rates (because neither should be affected by repeated presentations). In the latter case, it is unclear how participants could discriminate between similar distractors and studied targets. Our experiments suggest that even older adults benefit from repeated presentations (inasmuch as the difference in *old* judgments for studied words and similar lures increases with presentation frequency). Such findings might be accommodated within the SAC model by arguing that older adults sometimes encode the wrong plurality on a given presentation or that, for this age group, repetitions increase the strength of both singular and plural word nodes and/or their associations with context. Additional theoretical machinery, however, would be needed to explain why young adults show different patterns of false alarms for repeated items when tested with and without time pressure.