The effect of warnings on false memories
in young and older adults

DAVID P. McCABE and ANDERSON D. SMITH
Georgia Institute of Technology, Atlanta, Georgia

In the present experiments , we examined adult age differences in the ability to suppress false memories, using the Deese–Roediger–McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995). Participants studied lists of words (e.g., *bed, rest, awake,* etc.), each related to a nonpresented critical lure word (e.g., *sleep*). Typically, recognition tests reveal false alarms to critical lures at rates comparable to those for hits for studied words. In two experiments, separate groups of young and older adults were unwarned about the false memory effect, warned before studying the lists, or warned after study and before test. Lists were presented at either a slow rate (4 sec/word) or a faster rate (2 sec/word). Young adults were better able to discriminate between studied words and critical lures when warned about the DRM effect either before study or after study but before retrieval, and their performance improved with a slower presentation rate. Older adults were able to discriminate between studied words and critical lures when given warnings before study, but not when given warnings after study but before retrieval. Performance on a working memory capacity measure predicted false recognition following study and retrieval warnings. The results suggest that effective use of warnings to reduce false memories is contingent on the quality and type of encoded information, as well as on whether that information is accessed at retrieval. Furthermore, discriminating between similar sources of activation is dependent on working memory capacity, which declines with advancing age.

False memories, the mistaken belief that an event has previously occurred, have long been of interest to psychologists (see Roediger, 1996). These memory errors are typically consistent with the gist of previous experience and often go unnoticed, but these errors can have deleterious consequences in certain situations. Eyewitnesses, for example, may falsely identify suspects in criminal investigations (Loftus, 1993), or patients may misremember the source of important medical information (Skurnik, Park, & Schwarz, 2000). In the present investigation, we examined some of the factors that affect the veracity of memory performance, using a paradigm that has generated a great deal of recent research: the Deese–Roediger–McDermott (DRM) converging associates paradigm (Deese, 1959; Roediger & McDermott, 1995). The typical DRM procedure involves having participants listen to lists of words (e.g., *bed, rest, awake,* etc.) that are semantically associated with a nonpresented critical lure word (e.g., *sleep*).

Subsequent recall and/or recognition tests reveal that participants remember the critical lure at rates comparable to those for list items (Roediger & McDermott, 1995).

Several theoretical explanations have been offered to explain these robust false memory effects. One possible explanation, suggested by Roediger and McDermott (1995), is that the critical lure is repeatedly activated by the studied words through implicit associative responses (Underwood, 1965), resulting in the critical lure's coming to mind (either consciously or unconsciously) during the study episode. According to this account, the subsequent false memories can be understood as reality-monitoring errors in which a participant cannot recollect whether the critical lure was simply imagined during study or was actually perceived. This explanation is consistent with the source-monitoring framework proposed by Johnson and colleagues (Johnson, Hashtroudi, & Lindsay, 1993; Mather, Henkel, & Johnson, 1997), which suggests that remembering is a decision-making process driven both by information stored at encoding and by the context in which the information is retrieved.

The source-monitoring framework specifies two general types of decision-making processes involved in remembering: heuristic and systematic. Heuristic decision making involves using simple rules in deciding whether an event has occurred and is usually done with little deliberation, whereas systematic decision making involves the use of more contemplative, controlled processing. Whether heuristic or systematic decision making is used at retrieval depends on many factors, including the quantity and quality of the information available from the study

episode, as well as strategic processes. In the words of Johnson et al. (1993), source-monitoring decisions rely "fundamentally on the quality of the information recorded about events initially. These memory records are the product of specific perceptual and reflective processes engaged during the initial experience" (p. 5).

Warnings in the DRM Paradigm

One research strategy, the aim of which is to study strategic monitoring processes at encoding and retrieval, is to use explicit warnings about the false memory effect in order to examine whether participants can reduce or avoid false memories (Gallo, Roberts, & Seamon, 1997; McDermott & Roediger, 1998). Research to date has indicated that giving explicit warnings before presentation of the lists reduces, but does not eliminate, false recognition (Gallo et al., 1997; Gallo, Roediger, & McDermott, 2001; McDermott & Roediger, 1998). Furthermore, the strategy used most often by the participants who were warned about the false memory effect, and the strategy that was most effective, involved attempting to identify the critical lure during presentation of the lists. Research in which the question of whether warnings at retrieval can reduce false recognition has been examined has led to mixed results. Anastasi, Rhodes, and Burns (2000) and Gallo et al. (2001) found small but significant reductions in false recognition between participants who were warned before test and those who were not, but Neuschatz, Payne, Lampinen, and Toglia (2001) failed to find a significant difference. These findings suggest that once the lists are encoded, decision processes at retrieval are of little help in reducing false recognition.

Although none of the warning experiments led to elimination of false recognition, the possibility remains that warnings might be more effectively used to reduce false recognition under more favorable conditions. Alternatively, if DRM false memories are memory illusions in the same sense as perceptual illusions, warnings would not eliminate the effect, which should be obtained as a normal by-product of cognition. Of particular relevance to the present investigation, the presentation rates in previous warning studies were between 1 and 2 sec, allowing little time for effective processing of item-specific information. According to the implicit associative response hypothesis (Roediger & McDermott, 1995; Underwood, 1965), any occurrence of a word will automatically activate related associates, which would be expected to result in high levels of false recognition. However, if more time is available to study a word, more item-specific information can be encoded, and this elaborated information could be used to oppose the influence of automatically activated semantic information following warnings at retrieval. Stated differently, increasing the amount of study time will enhance one's knowledge of what *was* presented, and this information could be used to reject what *was not* presented.

Using a converging associates paradigm similar to the DRM paradigm, Arndt and Hirshman (1998) manipulated presentation rate as a means of testing predictions of the MINERVA2 model. Their results showed that veridical recognition increased to a greater degree than false recognition when words were presented at a slower presentation rate (e.g., 3 vs. 1 sec/word). More recently, McDermott and Watson (2001) found that veridical recall increased monotonically as a function of presentation rate. False recall initially increased as well (from 50 to 250 msec) but declined as presentation rates increased from 250 msec through 5,000 msec. In the present experiments, both a slow (4 sec/word) and a fast (2 sec/word) presentation rate were used, in conjunction with explicit warnings about the DRM false memory effect either before study or after study but before retrieval. Our basic prediction was that presenting words at a slower rate during study would result in participants' processing more item-specific information, which should result in enhanced veridical recognition. However, without knowledge of the false memory effect, presentation rate would have little or no effect on false recognition, because recognition decisions would be made in a heuristic, familiarity-based fashion. That is, participants who were unwarned about the false memory effect would use the highly salient semantic information to make their recognition decisions following either presentation rate. Conversely, the group warned at retrieval would use more systematic decision criteria on the recognition test and would be able to benefit from the enhanced item-specific processing that would occur during study following a slower presentation rate. Finally, when warnings were given at study, the slower presentation rate would allow the participants more time for strategic processing (i.e., identifying the critical lure), which would result in reductions of false recognition.

Age Differences in False Memories

Previous research has revealed greater rates of false recognition for older adults (Cohen & Faulkner, 1989; Smith, 1975). For example, Norman and Schacter (1997) found no differences in subjective ratings of semantic information for studied words and critical lures for either a young or an older adults group. Younger adults, however, were better able to discriminate between studied words and critical lures on the basis of both sound and list position (item-specific information), and they showed less false recognition.

The effect of warnings on age differences in the DRM paradigm may shed light on the mechanisms underlying the effect. Balota et al. (1999) suggested that warnings about the DRM effect might benefit older adults more than young adults, since older adults might not spontaneously use source-specifying (item-specific) information that is nevertheless available. This hypothesis is based on previous research in which age differences in false fame were examined (Dywan & Jacoby, 1990), indicating that older adults can reduce false memories to a greater degree than can young adults when encouraged to monitor source at retrieval (Multhaup, 1995). The contention is that older adults may have item-specific information available at retrieval but that they simply do not use it

under conditions that promote lax decision criteria, whereas young adults often spontaneously use this information. Thus, whereas research in which the DRM paradigm has been used indicated that older adults do not use item-specific information to reduce false recognition, the incorporation of explicit warnings at retrieval in the present experiments would determine whether they *can* actively suppress false recognition.

Balota et al. (1999) explained age differences in the discriminability of veridical and false memories as resulting from age-related declines in attentional control of relevant processing pathways. The spreading activation assumed to give rise to false memories in the DRM paradigm is relatively automatic and is not expected to decline with normal aging, but the processes that support veridical recollection, including discriminating similar sources of activation, rely on frontal lobe functioning that is known to decline in old age (see West, 1996). Although Balota and colleagues did not find a correlation between frontal lobe measures and false recall in their older adult sample, they suggested that this result may have been due to a poorly targeted battery of frontal lobe tests (the tests used were the Boston Naming Test, word fluency, and Mental Control, which were administered as part of a longitudinal study).

In the present experiments, we tested whether a measure of working memory capacity would correlate with false recognition. If the Balota et al. (1999) hypothesis is correct and increases in false memories result from source confusions caused by decreased frontal lobe functioning, working memory capacity should correlate with false memories. This conclusion follows from theories that view the central executive component of working memory as synonymous with frontal lobe functioning (e.g., Kane & Engle, in press). Toward this end, we have developed a task that we call *Stroop span* to examine the relation between false recognition suppression and working memory capacity. In this task, Stroop color words (e.g., the word *green* in red font; Stroop, 1935) appear on the screen at a 1-sec rate, and participants must remember the color of each color word. After all of the color words for a given trial have appeared (trials range from two to seven color words in length), participants must recall the color of the words in their correct serial order. This measure is consistent with the conceptualization of working memory capacity as the ability to concurrently process and maintain information (Daneman & Carpenter, 1980). That is, requiring participants to resolve Stroop interference (i.e., suppressing the word-reading response) while actively maintaining the names of previously identified colors should tax working memory capacity.

It is important to note that we expected working memory capacity to be related to false recognition only in the conditions in which the participants were actively attempting to reduce false recognition. These conditions should also show age differences in false recognition, since previous research has indicated that working memory capacity decreases with increased age (Park et al., 1996). We expected age differences to be small or nonexistent in the unwarned group, because most of the participants in both age groups would use a heuristic decision criterion, so that words that were consistent with the theme of previously studied lists would be endorsed as *old*. When the participants were warned about the DRM effect prior to study, they would likely use a strategy of attempting to identify the critical lure during list presentation, a strategy that essentially transforms the study episode into a dual-task situation, a situation that requires working memory capacity for effective performance. Therefore, we were expecting age differences in this condition, and we expected an inverse relation between Stroop span performance and false recognition. When the participants were warned after study but prior to retrieval, they were faced with a very different task, but it should have required working memory capacity nevertheless. In this condition, the participants were being asked to compare directly two highly activated, similar sources of information. Specifically, when the participants encountered a critical lure on the recognition test, semantic information associated with the studied list would be activated, and the participants were being asked to discriminate directly between these highly similar sources of activation.

To summarize, the ability of young and older adults to suppress false recognition in the DRM paradigm was tested by providing explicit warnings before study or after study but before retrieval. Age differences were expected to be largest in the warning conditions, which should tax working memory. In Experiment 1, we examined presentation rate within groups, and in Experiment 2, we examined presentation rate between groups for young adults and examined only the slower presentation rate for older adults. False recognition was predicted to decline with a slow (4 sec) presentation rate, relative to a fast presentation rate (2 sec), but only for young adults.

EXPERIMENT 1

In Experiment 1, young and older adults studied four 10-word DRM lists at a 4-sec presentation rate and four 10-word DRM lists at a 2-sec presentation rate. The participants were unwarned about the false memory effect (the control condition), were warned prior to studying the lists, or were warned after study but before the recognition test. The slower lists were presented before the faster lists because we suspected that older adults would choose a nonoptimal rehearsal strategy (e.g., repetition) if they encountered the faster lists first. In addition, the study warning group was encouraged to use a strategy of identifying the critical lure during list presentation, since we wanted to tax working memory during study in this group.

Method

Participants. The participants were 72 young adults with an age range of 18–22 years ($M = 19.75$, $SD = 1.32$) and 72 older adults with an age range of 60–80 years ($M = 71.36$, $SD = 5.21$). The young adults were undergraduate students at Georgia Institute of Technology, and the older adults were recruited from newspaper advertisements in the Atlanta area. The young adults received course

credit, and the older adults received monetary compensation for their participation. Individuals with recent stroke or head injury were excluded. The same experimenter (the first author) tested all of the participants in the experiment.

Design. A 2 (age: young or old) × 3 (instruction group: unwarned, retrieval warning, or study warning) × 2 (presentation rate: slow or fast) × 4 (word type: studied targets, critical lures, unstudied targets, or distractor lures) mixed design was used. The first two variables were manipulated between participants, with repeated measures on the last two variables.

Materials. The word lists used were 12 of the 18 lists that led to the highest false recognition rates in the Stadler, Roediger, and McDermott (1999) norms. The first 10 associates from each of these lists were used (the normed lists are 15 words in length). List words for each critical lure were ordered from highest to lowest associative strength (i.e., the probability of the critical lure's eliciting the list words in free association norms). The 12 lists were used to create three 40-word list sets for counterbalancing purposes. The average normed false recognition rates for these three sets were equal ($M = 80\%$). Each participant heard two of these three 40-word lists: one presented at a slow rate (4 sec/word), and one at a faster rate (2 sec/word). Each of the three 40-word lists was recorded at both a slow and a fast rate, and they were counterbalanced so that each list was presented an equal number of times at each rate across the experiment. The list that was not presented provided distractors for the recognition test. The word lists were digitally recorded by the experimenter and played on a Macintosh computer using Labtec speakers (Model LC-1030). The recognition test consisted of 48 words presented one at a time on a computer monitor in 72-point Arial font. Test items included 3 words from each of the studied lists (studied words), the nonpresented lure word for each of these lists (critical lures), and the corresponding words for each distractor list (unstudied distractors and distractor lures, respectively). The 3 list words were taken from the first, fourth, and eighth list positions from each of the lists.

Procedure. All the participants were told that they would be listening to eight 10-word lists in immediate succession and that they should attempt to remember the words for a memory test that they would be given a few minutes afterwards. Each 10-word list was preceded by a "camera click" sound to indicate that a new list was beginning. The participants were informed that they would be hearing the first four lists presented at a rate of 1 word every 4 sec, followed by the last four lists at a rate of 1 word every 2 sec. The study warning group was also given an explicit warning regarding the nature of the lists, and an example list (corresponding to the lure word *needle*) was read to them. These participants were instructed that trying to figure out what the nonpresented critical lure word was while the list was presented would be an effective strategy. They were further instructed that some of the words on the recognition test would be critical lures and that they should avoid falsely recognizing them. After any questions were answered, and the experimenter confirmed that the instructions were understood, the study lists were presented.

The retrieval warning group received the explicit warning about the nature of the lists after study but before test. To equate the time between study and test for all groups, the unwarned and study warning groups completed the pattern comparison test (described below) as a filler task. On the recognition test, the participants were instructed that they would have to make one of four responses for each word on the recognition test. The participants indicated *sure old*, *probably old*, *probably new*, or *sure new* by pressing the 4, 5, 6, or + key on the computer number pad that were labeled "SO," "PO," "PN," and "SN," respectively. All the participants were told that there was no time limit and that they should be as accurate as possible. After the recognition test, the participants were asked what strategy they had used during study. For the study warning group, this included whether they had attempted to identify the lure word during study. During the debriefing, the unwarned and retrieval warning participants were asked whether they had become aware of the purpose of the experiment during the study phase (i.e., to evoke false

alarms to the critical lures on the recognition test). Because they became aware of the purpose of the experiment, data from 4 young and 6 older adults were replaced.

Several other measures were administered to each participant as well; some were used as filler tasks, and others were used in the individual difference analysis. All of these measures were administered after the recognition test, except for the letter comparison task, which occurred before the study phase for all the groups, and the pattern comparison test, which served as a filler task for the unwarned and the study warning groups between study and test (for reasons already described). The retrieval warning group was given the pattern comparison task after the strategy questionnaire.

The *letter comparison* and *pattern comparison* tasks are perceptual speed tests in which participants compare strings of letters or simple line drawings (respectively) and decide whether they are the same or different as quickly as possible. These were both filler tasks in the context of the present study.

Stroop span was used as a measure of working memory capacity. The participants were seated at a comfortable viewing distance from the computer screen (approximately 24 in.), and all stimuli were presented in the center of the screen. An answer sheet consisted of columns of answer blanks, which the participants used to record their responses. The task began with the participants briefly practicing naming the color of four congruent (e.g., the word *blue* in blue font) and four incongruent (e.g., the word *blue* in red font) color words. After this practice, the participants were told that they would see color words of both types presented on the screen one at a time and that they were to remember the *color* of each of the words presented. After several (between two and seven) had been presented, they would see a screen with the word RECALL prompting them to recall the first letter of each of the colors they had been remembering, in order. Words were presented for 1 sec each.

There were two 2-word practice trials presented before the beginning of the actual task. Trials proceeded from shortest to longest. There were three 2-word trials, five 3-word, 4-word, 5-word, and 6-word trials, and three 7-word trials. The stimuli used were the words *blue*, *green*, *red*, and *yellow* presented on the computer screen in lower case 100-point Arial Black font (note that "Black" refers to the name of the font, not its color). Each word was presented as a congruent color word and an incongruent color word (in each of the incongruent colors) approximately an equal number of times across the task. For trials with even numbers of words (i.e., two, four, and six), there were equal numbers of congruent and incongruent words on each trial; for trials with an odd number of words (i.e., three, five, and seven), there was one extra incongruent word for the majority of the trials and one extra congruent word for the remaining trials. The same two color words never appeared consecutively, nor did the same two colors (with one exception, owing to a programming error).

The participants pressed the space bar to begin each trial, which began with a blank screen appearing for 1 sec, followed by the color words. A recall screen displayed the word RECALL in capital letters after the words had been presented; after 2,500 msec, this screen was replaced by a screen with the instruction, "Please press any key to continue," which the participants pressed when they were ready to begin the next trial. All the stimuli were presented on a black background to maximize contrast. The total number of trials on which all the colors were correctly recalled was used as a measure of working memory capacity.

After completing Stroop span, the *Shipley Institute of Living Vocabulary*, a standardized multiple-choice synonym vocabulary test (Zachary, 1986), was administered. On this test, the participants viewed a capitalized word and were asked to choose a synonym from four possible answer choices.

Results

All results reported as significant were significant at $p < .05$.

Table 1
Descriptive Statistics for Experiments 1 and 2

Measure	Young Adults		Older Adults	
	M	SD	M	SD
Experiment 1				
Chronological age	19.75	1.32	71.36	5.21
Education (in years)	13.67	1.30	15.78	2.59
Vocabulary	30.90	3.25	35.88	3.77
Stroop span*	15.93	4.20	10.81	5.05
Experiment 2				
Chronological age	20.70	1.31	70.27	5.31
Education (in years)	14.78	2.52	14.07	1.31
Vocabulary	31.94	2.96	34.56	3.95
Stroop span	23.74	3.72	15.53	5.10

Note—Stroop span had 26 trials in Experiment 1 and 36 trials in Experiment 2.

Descriptive statistics. Descriptive statistics are displayed in Table 1. Older adults reported more years of education [$t(142) = 6.18$] and had higher vocabulary scores [$t(142) = 8.49$]. Older adults had slower response times on the pattern comparison test [$t(136) = 12.17$] and on the letter comparison test as well [$t(141) = 9.16$]. Young adults correctly recalled more trials on computation span [$t(136) = -9.88$] and Stroop span [$t(139) = -6.56$]. There were no significant age differences in self-reported health [$t(142) = 1.80$]. These results are consistent with those typically found in the literature.

Overall recognition performance. The proportion of responses called *sure old* or *probably old* on the recognition test are presented for each word type in Table 2. To correct for baseline rates of false alarms, corrected means for each participant were calculated by subtracting the hit rates for unstudied targets from those for studied targets, giving a measure of veridical recognition, and those for distractor lures from those for critical lures, giving a measure of false recognition. These corrected scores are used throughout. As a measure of the ability to discriminate presented from nonpresented items, we also calculated the difference between veridical and false recognition and included these results in the last set of rows in the table.

We began by conducting a 2 (age: young or older) × 3 (instruction group: unwarned, study warning, or retrieval warning) × 2 (item type: veridical or false) analysis of variance (ANOVA), with recognition scores collapsed across presentation rate. This analysis yielded a main effect of item type [$F(1,138) = 77.16, MS_e = 4.38$], an age × item type interaction [$F(2,138) = 4.60, MS_e = 0.26$], and an instruction group × item type interaction [$F(3,138) = 9.04, MS_e = 0.51$]. The three-way interaction was not significant [$F(4,138) = 0.35, MS_e = 0.02$]. We followed up by considering age and item type separately. Examining the effect of instruction group (unwarned, study warning, or retrieval warning) on veridical recognition for young adults revealed no significant differences between the three instruction groups for young adults [$F(2,69) = 0.71, MS_e = 0.03$]. Thus, warnings did not affect veridical recognition, and planned comparisons between the unwarned and each warning group separately did not yield any significant differences (all p values $> .10$). Turning to false recognition, there was a significant effect of instruction group for young adults [$F(2,69) = 6.17, MS_e = 0.48$]. Planned com-

Table 2
Experiment 1: Proportion of *Old* Responses as a Function of Item Type, Instruction Group, Age Group, and Presentation Rate

Item Type	Instruction Group	Young Adults				Older Adults			
		4 sec		2 sec		4 sec		2 sec	
		M	SD	M	SD	M	SD	M	SD
Studied targets	unwarned	.91	.11	.81	.13	.82	.14	.69	.19
	study warning	.82	.11	.74	.18	.79	.19	.75	.15
	retrieval warning	.86	.09	.72	.18	.75	.16	.69	.18
Critical lures	unwarned	.79	.23	.77	.23	.81	.24	.73	.27
	study warning	.45	.29	.40	.24	.52	.29	.58	.30
	retrieval warning	.56	.27	.63	.23	.74	.21	.74	.21
Unstudied targets	unwarned	.17	.18	.17	.18	.16	.15	.16	.15
	study warning	.14	.18	.14	.18	.18	.13	.18	.13
	retrieval warning	.16	.12	.16	.12	.23	.18	.23	.18
Unstudied lures	unwarned	.31	.22	.31	.22	.29	.25	.29	.25
	study warning	.23	.25	.23	.25	.34	.28	.34	.28
	retrieval warning	.21	.24	.21	.24	.30	.24	.30	.24
Corrected veridical recognition	unwarned	.74	.22	.65	.23	.66	.19	.53	.22
	study warning	.68	.20	.60	.25	.61	.21	.57	.18
	retrieval warning	.70	.16	.56	.19	.53	.22	.47	.24
Corrected false recognition	unwarned	.48	.27	.46	.34	.52	.30	.44	.35
	study warning	.22	.32	.17	.28	.18	.34	.24	.42
	retrieval warning	.35	.40	.42	.27	.44	.31	.44	.28
Veridical minus false recognition	unwarned	.26	.31	.19	.26	.14	.37	.09	.40
	study warning	.47	.37	.44	.31	.43	.45	.33	.46
	retrieval warning	.35	.40	.14	.34	.09	.36	.03	.37

Note—Unstudied targets and unstudied lures do not differ at each presentation rate, because there was only one recognition test.

parisons revealed a significant reduction in false recognition for the study warning group [$F(1,46) = 12.10, MS_e = 0.91$], but not for the retrieval warning group [$F(1,46) = 1.06, MS_e = 0.08$], when each was compared with the unwarned group. Nearly all (23 of 24) of the participants in the study warning group reported using the strategy of identifying the critical lure during list presentation. So, although warnings did not significantly reduce veridical memory, false memory was reduced when warnings occurred before, but not after, the lists were studied.

Examining the memory performance of older adults led to a somewhat different pattern of results. There was no significant effect of the instruction group variable on veridical recognition [$F(2,69) = 2.00, MS_e = 0.07$]. However, planned comparisons of each warning group with the unwarned group separately did indicate a marginally significant drop in veridical recognition following the retrieval warning [$F(1,46) = 3.18, MS_e = 0.12, p = .081$]. Thus, if anything, warning older adults after study but before retrieval actually hurt their veridical recognition slightly. Consistent with the data from young adults, false recognition for older adults revealed a significant effect of instruction group [$F(2,69) = 5.46, MS_e = 0.51$]. Planned comparisons of each warning group to the unwarned group indicated that false recognition was significantly reduced in the study warning group [$F(1,46) = 8.29, MS_e = 0.88$], but not in the retrieval warning group [$F(1,46) = 0.26, MS_e = 0.02$]. This pattern of results indicates that, like young adults, older adults benefited from warnings given prior to the study episode. Furthermore, as was the case for young adults, the majority of older participants (21 of 24) reported using the strategy of identifying the critical lure during presentation of the study lists. Because there were so few data points for the participants who did not attempt to identify the critical lure, we could not conduct reliable statistical tests to compare them with those who had. Although both age groups were able to reduce false recognition following study warnings, neither age group significantly reduced false recognition when warnings were given after study but prior to retrieval. If anything, older adults' veridical recognition actually suffered when a warning was given at retrieval, although this comparison was only marginally significant.

The effect of presentation rate on recognition. Next, we used a 3 (instruction group: unwarned, study warning, or retrieval warning) $\times$ 2 (presentation rate: 4 sec or 2 sec) mixed model ANOVA to examine whether presentation rate affected the pattern of results in veridical and false recognition differently. Examining veridical recognition in the young adult group, we found a significant effect of presentation rate, with the slower rate leading to higher hit rates [$F(1,69) = 26.10, MS_e = 0.42$], but no instruction group $\times$ presentation rate interaction [$F(2,69) = 1.01, MS_e = 0.02$]. Thus, a slower presentation rate resulted in enhanced veridical recognition for all three instruction groups. The same instruction group $\times$ presentation rate ANOVA, in which false recognition for young adults was examined, revealed no difference as a function of presentation rate [$F(1,69) = 0.01, MS_e = 0.0001$] and no instruction group $\times$ presentation rate interaction [$F(2,69) = 0.97, MS_e = 0.04$]. It should be noted, however, that the difference between Veridical and false recognition was significantly greater at the slower presentation rate for the retrieval warning group ($.70 - .35 = .35$) than it was at the fast presentation rate ($.56 - .42 = .14$). Thus, despite no absolute differences in false recognition as a function of presentation rate for the young retrieval warning group, they did display better memory performance with a slower presentation rate. This was confirmed by a 2 (presentation rate: slow or fast) $\times$ 2 (recognition: veridical or false) repeated measures ANOVA for the retrieval warning group [$F(1,23) = 7.05, MS_e = 0.27$]. This same comparison did not reach significance for either the unwarned group [$F(1,23) = 1.44, MS_e = 0.03$] or the study warning group [$F(1,23) = 0.24, MS_e = 0.005$].

For older adults, a 3 (instruction group: unwarned, study warning, or retrieval warning) $\times$ 2 (presentation rate: 4 or 2 sec) mixed model ANOVA in which veridical recognition was examined yielded a significant effect of presentation rate, with the slower rate leading to higher rates of veridical recognition [$F(1,69) = 13.01, MS_e = 0.21$] but no instruction group $\times$ presentation rate interaction [$F(2,69) = 1.57, MS_e = 0.02$]. Thus, as with young adults, a slower presentation rate resulted in enhanced veridical recognition for all three instruction groups. Also consistent with young adults, an instruction group $\times$ presentation rate ANOVA in which false recognition was examined revealed no difference in false recognition as a function of presentation rate [$F(1,69) = 0.04, MS_e = 0.002$] and no instruction group $\times$ presentation rate interaction [$F(2,69) = 1.58, MS_e = 0.06$]. Contrary to the data for young adults, however, the difference between corrected veridical and false recognition was not significantly greater at the 4-sec rate ($.53 - .44 = .09$) than at the 2-sec rate ($.47 - .44 = .03$) for older adults. This was confirmed by a 2 (presentation rate: 4 or 2 sec) $\times$ 2 (recognition: veridical or false) repeated measures ANOVA conducted separately for the unwarned [$F(1,23) = 1.11, MS_e = 0.01$], study warning [$F(1,23) = 2.14, MS_e = 0.06$], and retrieval warning [$F(1,23) = 0.85, MS_e = 0.02$] groups.

Finally, as was expected, a 2 (age) $\times$ 3 (instruction group) ANOVA revealed that, overall, young adults had higher rates of veridical recognition [$F(1,138) = 8.81, MS_e = 0.32$] but did not differ from older adults in false recognition [$F(1,138) = 0.29, MS_e = 0.02$]. The age $\times$ presentation rate interaction was not significant [$F(1,138) = 0.20, MS_e = 0.07$].

Individual difference measures. We postpone an examination of the individual difference measures for Experiment 1 until after the second experiment has been discussed.

Discussion

Both young and older adults were able to use warnings given prior to the study episode to reduce false recognition, as compared with an unwarned control group, and

this reduction came without a cost to veridical recognition. Furthermore, the vast majority of the participants in both age groups reported using the strategy suggested by the experimenter—namely, identifying the critical lure word during list presentation. This lack of an age difference in false recognition is somewhat at odds with our working memory capacity hypothesis, which predicted that older adults would be less able to use this strategy, since it required concurrent processing and storage of information. This result will be considered further in the regression analyses following Experiment 2.

Neither age group was able to reduce false recognition when warnings were given after study but prior to retrieval. If anything, this warning reduced the veridical recognition of older adults. However, there was some evidence that a slower presentation rate led to better memory performance. Specifically, veridical recognition was greater for both young and older adults at a slower presentation rate, meaning that discrimination between veridical and false recognition was greater following the slower presentation rate. Furthermore, there was some indication that a slower rate of presentation improved the performance of young adults following retrieval warnings, and there was a trend toward a reduction in false recognition overall (i.e., an 8% reduction in corrected false recognition, relative to the unwarned participants). It should also be noted that, consistent with our hypotheses, for young adults retrieval warnings reduced the hit rate for critical lures for the retrieval warning group (.59), as compared with the unwarned group [.78; $F(1,46) = 13.07, MS_e = 0.42$]. This same comparison was not significant for older adults [$F(1,46) = 0.31, MS_e = 0.01$], with the retrieval warning group and the unwarned group having similar hit rates for critical lures (.77 and .74, respectively).

EXPERIMENT 2

Our second experiment was conducted to address several methodological shortcomings of Experiment 1. The primary issue was that presentation rate was manipulated within subjects and, consequently, there were only four critical lures associated with presented lists at each presentation rate (and four distractor lures to correct by), giving us very little power to detect differences in false recognition. Another issue complicating our interpretation of the results is that there were differences in hits for studied words between the two presentation rates but there was only one false alarm rate for unstudied distractors. It is possible that false alarm rates would have differed for these groups had presentation rate been manipulated between lists (e.g., higher hit rates for studied targets at the 4-sec rate likely would have resulted in lower false alarm rates to unstudied targets). We also did not counterbalance the order of the lists, instead electing to present all the slow lists, followed by all the faster lists, for each participant. As was mentioned in the introduction, this was done because older adults likely would have used a less optimal strategy to rehearse the lists had the faster lists been pre-

sented first. Nevertheless, the order of presentation for these lists may have interacted with age or warnings in unforeseen ways. To address the above-mentioned concerns, we manipulated presentation rate between groups for the young adults and included the slower rate only for the older adults (presentation rate did not interact with any other measures for the older adults in the first experiment, so we felt confident that including only the slower rate would be sufficient to address most of our concerns). Manipulating presentation rate between groups allowed us to calculate false alarms to unstudied distractors separately for each group, gave us more power to detect significant differences between groups, and eliminated the confound of presentation order.

Another change made in the second experiment was that we used the instructions of Gallo et al. (2001), except that we did not include an explicit instruction to the participants in the study warning group to attempt to identify the critical lure during list presentation. Instead, the instructions simply explained the nature of the lists and gave an example list. We had encouraged the strategy of identifying the critical lure in the first experiment because of our interest in taxing the working memory capacity of the participants in this group, but including this instruction makes it difficult to make comparisons with other published studies using similar warnings, but without explicit identification instructions.

Method

Participants. The participants were 108 young adults with an age range of 18–22 years ($M = 19.75$, $SD = 1.32$) and 54 older adults with an age range of 60–80 years ($M = 71.36$, $SD = 5.21$).

Design. A 2 (age: young or old) $\times$ 3 (instruction group: unwarned, retrieval warning, or study warning) $\times$ 2 (presentation rate: slow or fast) $\times$ 4 (word type: studied words, critical lures, unstudied distractors, or distractor lures) mixed design was used. The first two variables were manipulated between groups. Presentation rate was a nested between-groups variable for the young group, and word type was manipulated within subjects.

Materials. The materials used in Experiment 2 were the same as those in Experiment 1, with a few exceptions. The number of 10-word DRM lists was increased from 12 to 16, creating two sets of 8 lists for counterbalancing purposes. The set that was not presented provided distractors for the recognition test, which resulted in 64 words on the recognition test, taken from the same list positions as those in Experiment 1.

Procedure. The instructions for the memory test were the same, except that the study warning group was not encouraged to try to figure out what the nonpresented critical lure word was during list presentation. Also, in an effort to discourage the use of a liberal criterion by the unwarned group, all the groups were explicitly told that false alarms to distractors would be deducted from their hit rate for studied words. The vocabulary and Stroop span measures were the same as those in Experiment 1, except that the number of trials on the Stroop span task was increased to six trials at each set size (between two and seven) and the presentation order of trials was randomized.

Results

Descriptive statistics. Descriptive statistics for Experiment 2 are displayed in Table 1, and all of the relevant comparisons were consistent with those of Experiment 1.
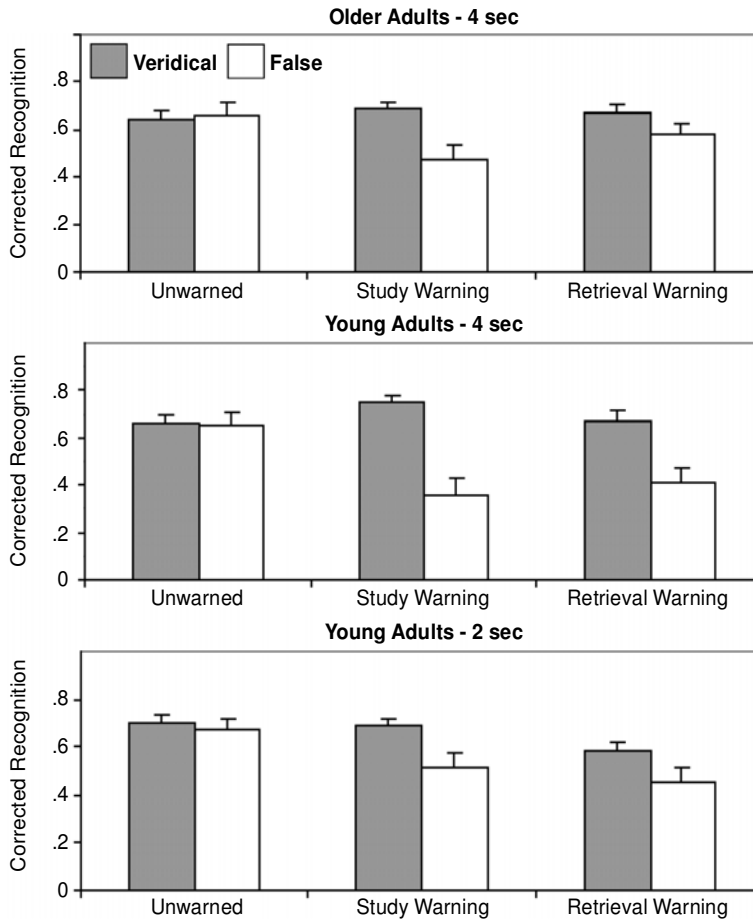
Figure 1. Experiment 2: corrected veridical and false recognition for older and young adults as a function of presentation rate and instruction group.

Recognition: 4-sec presentation rate. The recognition results for Experiment 2 are presented in Table 3, and the corrected recognition scores are graphed in Figure 1. We first will consider the recognition performance at the slower presentation rate and examine age differences therein, followed by a consideration of the faster presentation rate for young adults. A 2 (age: young or older) $\times$ 3 (instruction group: unwarned, study warning, or retrieval warning) $\times$ 2 (item type: veridical or false) ANOVA revealed a significant main effect of item type [$F(1,102) = 43.83, MS_e = 1.33$], a significant age $\times$ item type interaction [$F(2,102) = 7.16, MS_e = 0.22$], and a significant age $\times$ instruction group interaction [$F(3,102) = 14.30, MS_e = 0.43$]. The three-way interaction was not significant [$F(4,102) = 1.07, MS_e = 0.03$]. We followed up this analysis by examining the effect of instruction group (unwarned, study warning, or retrieval warning) on corrected veridical recognition for young adults, which was not significant [$F(2,51) = 1.90, MS_e = 0.05$]. Planned comparisons revealed a marginally significant difference between the unwarned and the study warning groups [$F(2,51) = 3.90, MS_e = 0.08, p = .056$], so that the study warning group

actually improved their veridical recognition, relative to the unwarned group. There was no difference between the unwarned and the retrieval warning groups [$F(2,51) = 0.001, MS_e = 0.0001$]. There was a significant effect of instruction group for corrected false recognition for young adults [$F(2,51) = 5.27, MS_e = 0.44$]. Planned comparisons revealed a significant reduction in corrected false recognition for the study warning group, as compared with the unwarned group [$F(1,34) = 11.67, MS_e = 0.77$], as well as between the unwarned and the retrieval warning groups [$F(1,34) = 9.13, MS_e = 0.53$]. We should also mention that all of the participants (18 of 18) reported using the strategy of identifying the critical lure during list presentation, despite no instruction to use this strategy. In summary, young adults were able to increase their veridical recognition following study warnings and to decrease their false recognition following either study or retrieval warnings.

For older adults, there was no significant effect of the instruction group variable on corrected veridical recognition [$F(2,51) = 0.44, MS_e = 0.01$], and planned comparisons revealed no significant differences between any of

the instruction groups. There was a significant effect of instruction group on corrected false recognition $[F(2,51) = 3.20, MS_e = 0.16]$. Planned comparisons of the unwarned group with the study warning group indicated a significant reduction in false recognition following study warning $[F(1,34) = 5.46, MS_e = 0.32]$. The majority of older adults in the study warning group (12 of 18) reported using the strategy of identifying the critical lure during study. Comparing those who attempted to identify the critical lure (*ID lure*) with those who did not (*no ID lure*) revealed no significant differences for veridical recognition (means of .70 and .66, respectively) $[t(16) = -0.71]$, but the ID lure group was significantly less likely to false recognize critical lures ($M = 0.38$) than was the no ID lure group $[M = 0.67; t(16) = 2.49]$. Older adults failed to reduce false recognition when warnings were given before study if they did not attempt to identify the critical lure. Finally, false recognition was not reduced for older adults in the retrieval warning group $[F(1,34) = 1.37, MS_e = 0.05]$.

Recognition: 2-sec presentation rate. Next, we will turn our attention to presentation rate for young adults (recall that older adults were not included in this condition). A 2 (presentation rate: 4 sec or 2 sec) × 3 (instruction group: unwarned, study warning, or retrieval warning) × 2 (item type: veridical or false) ANOVA indicated a significant main effect of item type $[F(1,102) = 45.94, MS_e = 1.47]$, a significant presentation rate × item type interaction $[F(2,102) = 5.15, MS_e = 0.16]$, and a significant presentation rate × instruction group interaction $[F(3,102) = 10.64, MS_e = 0.34]$. The three-way interaction was not significant $[F(4,102) = 1.86, MS_e = 0.06]$. Because the data for the young adults at the slow presentation rate have already been presented, we will begin by presenting the data for the fast presentation rate first, followed by comparisons between the two presentation rates. First, a between-groups ANOVA on corrected veridical recognition for instruction group revealed a significant effect $[F(1,51) = 4.02, MS_e = 0.08]$. This was due to a decrease in corrected veridical recognition for the retrieval warning group, as compared with the unwarned group $[F(1,34) = 4.99, MS_e = 0.13]$. There was no difference between the unwarned and the study warning groups $[F(1,34) = 0.003, MS_e = 0.00001]$.

A between-groups ANOVA in which the effect of instruction group on corrected false recognition was examined indicated a significant effect $[F(1,51) = 4.33, MS_e = 0.24]$. The unwarned group had higher rates of corrected false recognition than did both the study warning $[F(1,34) = 4.62, MS_e = 0.23]$ and the retrieval warning $[F(1,34) = 9.08, MS_e = 0.44]$ groups. Thus, despite a faster presentation rate, the young adults were able to reduce false recognition following warnings before or after study and before retrieval. However, the retrieval warning also reduced veridical recognition, suggesting that there was a cost associated with reducing false recognition in this group. Recall that at the slower presentation rate, the young adults were able to significantly reduce false recognition, with no differences in veridical recognition. Thus, part of the reason for a reduction of false recognition for the retrieval warning group at the fast presentation rate was that they were more conservative in their responses on the recognition test.

Discussion

Consistent with Experiment 1, there were no age differences in false recognition in the unwarned group, but both

Table 3
Experiment 2: Proportion of *Old* Responses as a Function of Item Type, Instruction Group, Age Group, and Presentation Rate

| | | Young Adults | | | | Older Adults | |
| | | 4 sec | | 2 sec | | 4 sec | |
Item Type	Instruction Group	M	SD	M	SD	M	SD
Studied targets	unwarned	.86	.06	.86	.11	.76	.17
	study warning	.84	.12	.81	.10	.81	.11
	retrieval warning	.80	.12	.79	.12	.77	.13
Critical lures	unwarned	.86	.11	.90	.13	.83	.19
	study warning	.40	.30	.65	.30	.61	.29
	retrieval warning	.61	.22	.70	.18	.75	.22
Unstudied targets	unwarned	.19	.13	.16	.12	.12	.14
	study warning	.08	.07	.11	.08	.13	.09
	retrieval warning	.13	.17	.22	.16	.10	.08
Unstudied lures	unwarned	.21	.16	.22	.16	.17	.16
	study warning	.04	.06	.14	.13	.14	.13
	retrieval warning	.20	.19	.25	.17	.17	.12
Corrected veridical recognition	unwarned	.66	.15	.70	.16	.64	.18
	study warning	.75	.13	.69	.11	.69	.11
	retrieval warning	.67	.19	.58	.17	.67	.14
Corrected false recognition	unwarned	.65	.22	.67	.18	.66	.21
	study warning	.36	.29	.51	.26	.47	.27
	retrieval warning	.41	.26	.45	.25	.58	.19
Veridical minus false recognition	unwarned	.01	.16	.03	.21	−.02	.20
	study warning	.39	.34	.18	.31	.21	.31

Table 4
Partial Correlation for Each Individual Difference
Variable With Veridical and False Recognition

	Veridical Recognition (Studied Targets Called *Old*)			False Recognition (Critical Lures Called *Old*)		
Measure	UW	SW	RW	UW	SW	RW
Chronological age	−.27*	.05	−.26*	−.12	.29**	.21
Stroop span	−.03	.12	−.11	−.11	−.22*	−.29*
Vocabulary	.00	.03	.15	−.06	−.32**	−.15

Note—UW, unwarned; SW, study warning; RW, retrieval warning.   *$p <$ .05.   **$p <$ .01.

young and older adults were able to use warnings given prior to the study episode to reduce false recognition, and this reduction came without a cost to veridical recognition. Also consistent with Experiment 1, the majority of the participants in both age groups reported using the strategy of attempting to identify the critical lure during study as a means of reducing false recognition. This occurred in Experiment 2 despite no encouragement on the part of the experimenter. However, one third of the older adults did not use this strategy, and this subset of the participants was not able to reduce false recognition. Again, the implication of this result for our working memory capacity hypothesis will be considered further in the section on regression analyses. Unlike in Experiment 1, only the young adults were able to reduce false recognition when warnings were given after study but prior to retrieval, and they were somewhat better able to do this when the studied lists had been presented at a slower presentation rate. When the studied lists were presented at a faster presentation rate, young adults were able to reduce false recognition in the retrieval warning group, but veridical recognition suffered.

Regression analyses (Experiments 1 and 2). The Spearman–Brown split-half reliabilities for all tests were between .88 and .95. To increase power in the analysis of the resource measures on false recognition, we combined the measures from both experiments. We did not use the recognition scores for the 2-sec presentation rate from Experiment 2, because only the young adults had been included in this condition. Also, raw hit rates for studied words and critical lures were used, because the error of measurement in the corrected scores was dramatic.

Table 4 shows the partial regression of each individual difference measure (including chronological age) on veridical and false recognition. These partial correlations reveal the relation between each individual difference measure and the memory measures, holding all of the other measures constant. Thus, partial correlations provide an estimate of the nonredundant variance between two variables of interest.

Chronological age was negatively related to veridical recognition in both the unwarned and the retrieval warning groups, consistent with the group differences already reported. Of greater interest in the present context, false recognition was not predicted by any of the individual dif-

ference measures in the unwarned group, consistent with our predictions. Also consistent with our predictions, Stroop span, a measure of working memory capacity, predicted false recognition in both the study warning group (partial $r = -.22$), and the retrieval warning group (partial $r = -.29$). This indicates that higher levels of working memory capacity were associated with lower levels of false recognition. Interestingly, in the study warning group, chronological age was still associated with higher levels of false recognition (partial $r = .29$), even after removing variance associated with the working memory capacity and word knowledge measures. Finally, vocabulary was inversely related to false recognition in the study warning group (partial $r = -.32$). This last result helps to explain why the older adults, who scored higher on measures of vocabulary in both experiments, were able to reduce false recognition in the study warning group, despite lower levels of working memory capacity.

Thus, our original hypothesis predicting a relation between working memory capacity and false recognition in the study and retrieval warning groups was confirmed, and we were able to explain the null effect of age group following study warnings as being partly due to age differences in word knowledge. However, working memory capacity and word knowledge did not eliminate age as a significant predictor in the study warning group, suggesting a possible role for other processes in reducing false recognition as well. One limitation of the present experiments is that only one measure of each construct was used. Future research using factor analytic techniques, which make use of multiple indicators of each theoretical construct, could be helpful in discovering the processes involved in the creation and maintenance of false memories.

GENERAL DISCUSSION

The primary goal of the present experiments was to examine the ability of young and older adults to suppress false recognition. To achieve this goal, warnings were given to young and older adults at study or after study and before retrieval, and studied words were presented at either a fast (2 sec/word) or a slow (4 sec/word) rate. As was expected, older adults were less able to actively suppress false recognition, especially when given warnings at retrieval. Surprisingly, older adults were nearly as effective as young adults when warnings occurred at study, indicating that both age groups were able to use strategic processing that allowed them to suppress false recognition in this condition. In both experiments, using a within-subjects (Experiment 1) or between-subjects (Experiment 2) design, presenting words at a slower presentation rate benefited young adults in terms of discriminating events that had or had not occurred, primarily when the participants were warned at retrieval. A slower presentation rate did not benefit older adults in terms of discriminating events that had or had not occurred (Experiment 1). Our individual difference analysis indicated that, as was predicted, a measure

of working memory capacity (Stroop span) was related to false recognition when the participants were warned before or after study, but not when the participants were unwarned about the DRM effect. Also, when the participants were warned at study, age was associated with higher levels of false recognition, even after controlling for working memory capacity and word knowledge. Finally, vocabulary was a significant predictor of reduced false recognition in the study warning group, even after controlling for working memory capacity and age, suggesting that word knowledge plays a role in strategic processing in this condition.

In terms of the source-monitoring framework (Johnson et al., 1993), the present results suggest that the quality of both the encoded information and the decision-making processes at retrieval gives rise to the creation and maintenance of verbal false memories. For example, the study phase did not differ for the group that was unwarned and the group that was warned at retrieval, yet the group warned at retrieval was able to reduce false recognition. Thus, the unwarned group could have reduced their false recognition, but they made their recognition decisions by using a familiarity-based heuristic, so that words matching the theme of the list, including critical lures that were not presented, were endorsed as having been studied. Both age groups likely used this strategy, since they had high levels of false recognition that did not differ in either experiment. The lack of a relation between working memory capacity and false recognition in this group supports the idea that heuristic decision criteria, which do not rely on controlled attention, were used. In terms of fuzzy-trace theory (Brainerd, Wright, Reyna, & Mojardin, 2001), another prevalent explanation for false memories, the participants relied on gist (or meaning) when making memorial decisions, rather than accessing the specific, verbatim details.

By contrast, only young adults were able to use warnings at retrieval, when more systematic decision processes were required for effective performance, and these warnings were more effective when study lists had been presented at a slower rate. Generally speaking, veridical recognition was enhanced for both age groups when words were presented at a slower presentation rate, a result consistent with previous research in which presentation rate was varied (Arndt & Hirshman, 1998). However, the difference between veridical and false recognition was greater following the slower presentation rate for young adults warned at retrieval.

The differences between corrected veridical and false recognition following our fast presentation rate were modest (14% for Experiment 1 and 13% for Experiment 2), as compared with this difference following the slower presentation rate (35% for Experiment 1 and 26% for Experiment 2). These results suggest that as the quality of the encoded information improves, young adults are able to use this information to discriminate between events that did or did not occur, but only when they are encouraged to

use this information via explicit warnings. We should also note that this argument is not consistent with the notion that false memories in the DRM paradigm arise because of an excessively liberal response criterion, as others have suggested (Miller & Wolford, 1999). Rather, the poorer discrimination with faster presentation rates suggests that experiments using presentation rates of 2 sec or less do not allow enough time to engage in the degree of item-specific processing necessary for retrieval warnings to be effective.

An alternative explanation for the present results would be that the young adults simply used some kind of strategy during the recognition test to identify the words most likely to be critical lures and rejected these words on this basis. However, this account does not readily explain why different presentation rates during study would have such a dramatic effect on the ability to discriminate studied words and critical lures. Also, it would not explain why the older adults, who apparently were able to identify the critical lures during study when given a warning, would not be able to use this same strategy at test.

One limitation of the present experiments is that there is no way to differentiate whether the age differences in the retrieval warning group are a result of poorer encoding of item-specific information during study or of deficits in accessing or using item-specific information to discriminate between sources of activation at retrieval. Research suggests that older adults are likely to process word lists in a more general fashion (Rabinowitz, Craik, & Ackerman, 1982). Moreover, Glisky, Rubin, and Davidson (2001) have convincingly demonstrated that frontal lobe functioning, as assessed by standardized psychometric tests, is related to deficits in encoding perceptual information during study, as well as in accessing source-specifying information at retrieval. More work will be required to differentiate the locus of age differences in the present paradigm.

It is notable that the participants with greater working memory capacity were better able to suppress false recognition following retrieval warnings, consistent with the idea that discriminating between similar sources of activation at retrieval requires attentional control (Balota et al., 1999). This also explains why older adults were unable to achieve accurate memory performance in the retrieval warning group.

Perhaps the most surprising finding in the present study, at least given our initial hypotheses, is the absence, in both experiments, of an age difference in the group warned at study. As we had expected, both age groups primarily used a strategy of identifying the critical lure during the study phase, and working memory capacity was related to reduced rates of false recognition. However, contrary to our expectations, older adults were able to use these warnings effectively, despite the processing demands required by this strategy. The finding that word knowledge (as measured by a synonym vocabulary test) was inversely related to false recognition helps to explain why no age group differences were found for this group in the present experi-

ments, despite age-related deficits in working memory capacity. Libby and Neisser (2001) suggested that the use of higher order structural knowledge (i.e., knowledge of the specific relation between studied targets and critical lures) places constraints on how gist-based information is used during a memory test, allowing gist-based processing to support accurate memory performance. Research in which age differences in the use of situation models was examined has been consistent with this notion, revealing no age differences in the processing of higher level structural information, despite age differences in the processing of verbatim information (Radvansky, Zwaan, Curiel, & Copeland, 2001).

Research by Watson, McDermott, and Balota (2000) is also consistent with this explanation. They examined false recall in young and older adults, using a combination of study warnings and repeated study–test trials. Although levels of false recall were reduced to the same extent for young and older adults on the first test (relative to unwarned participants), only young adults were able to further reduce false recall over repeated study–test trials. These results are consistent with our suggestion that multiple processes, which show differential age-related declines, are involved in reducing false memories in the study warning condition.

Mechanisms of False Recognition Suppression

Returning again to the account of false memories described by Balota et al. (1999), the processes involved in avoiding false memories appear to be similar to those involved in the traditional Stroop interference task. By this account, the highly activated semantic information in the DRM task is analogous to the automatically activated but irrelevant word information in the Stroop task, whereas the item-specific information in the DRM task is analogous to the color information in the Stroop task. The critical difference seems to be that, in the Stroop task, the need to use controlled attention to resolve the interference is readily apparent, since participants are asked to name the color rather than to read the word, but in the typical DRM experiment, participants are unaware of the need to use attentional control to avoid the influence of the inappropriate source of activation. Consequently, providing warnings increases the similarity between the processing demands of false recognition suppression and the Stroop task, consistent with both the experimental and the correlational results.

Recently, an *activation-monitoring* framework for understanding DRM false memories has been proposed that may help in understanding the present results (McDermott & Watson, 2001; Roediger, Watson, McDermott, & Gallo, 2001). According to this framework, spreading activation in semantic networks gives rise to activation of critical lures, and monitoring processes at encoding and retrieval can be used to reduce false remembering. This explanation is consistent with the other models based on spreading activation already discussed (Balota et al., 1999) and predicts age differences consistent with those found in the present experiments.

CONCLUSIONS

In conclusion, suppressing false recognition in the DRM paradigm is a difficult task that involves discriminating between highly similar sources of activation (i.e., studied words and critical lures). Our results provide evidence for theoretical frameworks suggesting that individuals who are more susceptible to interference in working memory, including older adults, are less likely to accurately discriminate highly similar sources (Balota et al., 1999; Hasher & Zacks, 1988; McDermott & Watson, 2001). However, in all conditions, false recognition remained robust, suggesting that DRM false recognition may be an unavoidable memory illusion (Roediger, 1996).

REFERENCES

Anastasi, J. S., Rhodes, M. G., & Burns, M. C. (2000). Distinguishing between memory illusions and actual memories using phenomenological measurements and explicit warnings. *American Journal of Psychology*, **113**, 1-26.

Arndt, J., & Hirshman, E. (1998). True and false recognition in MINERVA2: Explanations from a global matching perspective. *Journal of Memory & Language*, **39**, 371-391.

Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger, H. L., III, McDermott, K. B., & Yerys, B. E. (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology*, **16**, 361-384.

Brainerd, C. J., Wright, R., Reyna, V. F., & Mojardin, A. H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 307-327.

Cohen, G., & Faulkner, D. (1989). Age differences in source forgetting: Effects of reality monitoring and eyewitness testimony. *Psychology & Aging*, **4**, 10-17.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, **19**, 450-466.

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

Dywan, J., & Jacoby, L. [L.] (1990). Effects of aging on source monitoring: Differences in susceptibility to false fame. *Psychology & Aging*, **5**, 379-387.

Gallo, D. A., Roberts, M. J., & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review*, **4**, 271-276.

Gallo, D. A., Roediger, H. L., III, & McDermott, K. B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin & Review*, **8**, 579-586.

Glisky, E. L., Rubin, S. R., & Davidson, P. S. R. (2001). Source memory in older adults: An encoding or retrieval problem? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 1131-1146.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193-225). San Diego: Academic Press.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28.

Kane, M. J., & Engle, R. W. (in press). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*.

Libby, L. K., & Neisser, U. (2001). Structure and strategy in the associative false memory paradigm. *Memory*, **9**, 145-163.

Loftus, E. F. (1993). Psychologists in the eyewitness world. *American Psychologist*, **48**, 550-552.

Mather, M., Henkel, L. A., & Johnson, M. K. (1997). Evaluating characteristics of false memories: Remember/know judgments and

memory characteristics questionnaire compared. *Memory & Cognition*, **25**, 826-837.

McDermott, K. B., & Roediger, H. L., III (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory & Language*, **39**, 508-520.

McDermott, K. B., & Watson, J. M. (2001). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory & Language*, **45**, 160-176.

Miller, M. B., & Wolford, G. L. (1999). Theoretical commentary: The role of criterion shift in false memory. *Psychological Review*, **106**, 398-405.

Multhaup, K. S. (1995). Aging, source, and decision criteria: When false fame errors do and do not occur. *Psychology & Aging*, **10**, 492-497.

Neuschatz, J. S., Payne, D. G., Lampinen, J. M., & Toglia, M. P. (2001). Assessing the effectiveness of warnings and the phenomenological characteristics of false memories. *Memory*, **9**, 53-71.

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition*, **25**, 838-848.

Park, D. C., Smith, A. D., Lautenschlager, G., Earles, J. L., Frieske, D., Zwahr, M., & Gaines, C. L. (1996). Mediators of long-term memory performance across the life span. *Psychology & Aging*, **11**, 621-637.

Rabinowitz, J. C., Craik, F. I. M., & Ackerman, B. P. (1982). A processing resource account of age differences in recall. *Canadian Journal of Psychology*, **36**, 325-344.

Radvansky, G. A., Zwaan, R. A., Curiel, J. M., & Copeland, D. E. (2001). Situation models and aging. *Psychology & Aging*, **16**, 145-160.

Roediger, H. L. III. (1996). Memory illusions. *Journal of Memory & Language*, **35**, 76-100.

Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 803-814.

Roediger, H. L., III, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, **8**, 385-407.

Skurnik, I., Park, D. C., & Schwarz, N. (2000, April). *Repeated warnings about false medical information can make it seem true: A paradoxical age difference*. Paper presented at the biennial Cognitive Aging Conference, Atlanta.

Smith, A. D. (1975). Partial learning and recognition memory in the aged. *International Journal of Aging & Human Development*, **6**, 359-365.

Stadler, M. A., Roediger, H. L., III, & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, **27**, 494-500.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, **70**, 122-129.

Watson, J. M., McDermott, K. B., & Balota, D. A. (2000, April). *Veridical and false memories in young and older adults using repeated study/test trials and warning instructions*. Poster presented at the Cognitive Aging Conference, Atlanta.

West, R. L. (1996). An application of prefrontal cortex theory to cognitive aging. *Psychological Bulletin*, **120**, 272-292.

Zachary, R. A. (1986). *Shipley Institute of Living Scale, Revised Manual*. Los Angeles: Western Psychological Services.