

Effects of repetition on memory for pragmatic inferences

KATHLEEN B. McDERMOTT and JASON C. K. CHAN
Washington University, St. Louis, Missouri

Social interaction requires active inferential processing on the part of the listener. Such inferences can affect memory. For example, after hearing *the karate champion hit the cinder block*, one might erroneously recollect having heard the verb *broke* (Brewer, 1977)—a reasonable inference, but one not logically necessitated. The mechanisms behind this type of erroneous recollection have not been much explored. Experiments in the present article assessed the influence of repetition, response deadline, and age (cf. Jacoby, 1999), in an effort to demonstrate the dual contributions of familiarity and recollection underlying this phenomenon. For older adults, repetition at encoding increased the later likelihood of erroneously recognizing pragmatic inferences. For younger adults, repetition exerted the opposite effect. Both age groups, however, benefited from a second study–test trial. Experiment 2 demonstrated a similar interaction on a cued recall test for younger adults, whereby repetition exerted different influences as a function of time permitted during retrieval. Implications for theories of memory and discourse processing are considered.

Although recollection for prior events often feels compelling and vivid, this richness of experience can be misleading. Remembering is constructive: We piece together fragments of remembered events to arrive at a seemingly coherent and complete representation. The goal of uncovering the mechanisms underlying memory construction has led researchers to devise various methods for inducing erroneous memories in the laboratory; these methods include introducing misleading postevent information (Loftus & Palmer, 1974), taking advantage of social influences (Roediger, Meade, & Bergman, 2001), and asking people to imagine events that they did not previously experience (Garry, Manning, Loftus, & Sherman, 1996; Goff & Roediger, 1998; Hyman, Husband, & Billings, 1995).

Even in instances in which no overt erroneous information is supplied, processes internal to the participant can lead to errors in recall. For example, after studying compound words (e.g., *blackbird* and *airmail*), people often recognize a lure that contains features of those studied words (e.g., *blackmail*, Reinitz, Lammers, & Cochran, 1992). Moreover, studying many words that are seman-

tically associated (e.g., *bed*, *rest*, *awake*, Deese, 1959; Roediger & McDermott, 1995) or phonologically associated (e.g., *weep*, *beep*, *slide*, Chan, McDermott, Watson, & Gallo, 2005; McDermott & Watson, 2001; Sommers & Lewis, 1999) to a single nonpresented word (e.g., *sleep*) can readily elicit memory for that nonpresented word.

Most relevant to the present article is the finding that sentences that encourage pragmatic inferences elicit later memory for those inferences (Brewer, 1977; Johnson, Bransford, & Solomon, 1973). For example, Brewer (1977) showed that after studying sentences such as *The karate champion hit the cinder block*; *The infant stayed awake all night*; and *The flimsy shelf weakened under the weight of the books*, people often remember that the karate champion *broke* the cinder block, the infant *cried* all night, and the flimsy shelf *collapsed* under the weight of the books. Importantly, the recall errors go beyond simple syntactical changes and can include inferences that do not necessarily follow from the statements. *Hitting*, *staying awake*, and *weakening* do not necessitate *breaking*, *crying*, and *collapsing*. These pragmatic inferences can be distinguished from meaning-preserving logical inferences (e.g., *John is taller than Jim* logically implies that *Jim is shorter than John*). Such logical inferences also affect memory in interesting ways (Bransford & Franks, 1971; Harris & Monaco, 1978) but are not the focus of the present research project.¹ Note that in some real-life situations, pragmatic inferences may be warranted, in that they capture the meaning the speaker intended to convey but did not portray with his/her specific choice of words. Nonetheless, we consider later recollection of the pragmatic inference to be a type of false memory, because a pragmatic inference changes the meaning of the originally presented information (and, as will be seen, in our ex-

We thank Bill Brewer for providing the materials from his 1977 article. We also thank Chris Broughton and April Clift for assistance in participant testing and Sean Kang, Roddy Roediger, Martha Storandt, Karl Szpunar, Jason Watson, and Franklin Zaromb for helpful discussion. The data reported here were presented at the 43rd Annual Meeting of the Psychonomic Society, Kansas City, MO, November 2002. This work was funded by a grant from NIMH to the first author. Correspondence should be addressed to K. B. McDermott, Department of Psychology, Campus Box 1125, Washington University, St. Louis, MO 63130 (e-mail: kathleen.mcdermott@wustl.edu).

Note—This article was accepted by the previous editorial team, when Colin M. MacLeod was Editor.

perimental procedures, participants were asked to remember precisely what had been previously encoded, without making any elaborations or word substitutions).

The present article reports two experiments designed to explore memory for pragmatic inferences, which have not received much recent attention in the false memory literature, despite their obvious relevance. Pragmatic inferences are particularly interesting, in that they are likely the origin of many everyday false memories and may offer a bridge between word-list studies of false memories and studies of discourse comprehension.

Our interest in pragmatic inferences arose from recent theories and empirical findings in the false memory literature. Specifically, dual-process theories of retrieval have been applied to the false memory literature quite successfully (see Yonelinas, 2002, for a comprehensive review). The two processes, generally denoted *recollection* and *familiarity*, represent controlled and relatively automatic processes, respectively. In the present article, we begin our investigation by applying well-established experimental manipulations rooted in the dual-process framework to the pragmatic inference paradigm. The goal was to begin to characterize the role of familiarity and recollection in leading to memory for pragmatic inferences.

Consider, first, a set of empirical findings. The influence of repetition during the encoding phase changes as a function of the amount of time given at retrieval (Jacoby, 1999). For example, using converging semantic associates to elicit false recognition of related but nonpresented associates [the Deese/Roediger–McDermott (DRM) paradigm], Benjamin (2001) showed that when younger adults were not rushed during retrieval, three presentations at study led to a higher hit rate and a lower false alarm rate than a single presentation (hit rates of .70 and .88 and false alarm rates of .53 and .43 for one and three presentations, respectively). When the participants were rushed during the retrieval phase, though, both the hit rate and the false alarm rate rose with repetition (hit rates of .49 and .65 and false alarm rates of .46 and .59 for one and three presentations, respectively).

Dual-process theories have been used to explain this set of findings in the following way. Repeated presentation during study leads to increases in recollection and familiarity (Yonelinas, 2002); such increases translate into an enhanced hit rate, because recollection and familiarity lead to the same (i.e., “old”) response for studied items. However, when a response deadline is imposed during the test phase, only the faster process (familiarity) can be used (Atkinson & Juola, 1973; McElree, Dolan, & Jacoby, 1999); there is not enough time for recollection to be used reliably during retrieval. Hence, the enhanced familiarity for the related nonpresented associates brought about by repetition during the study phase cannot be overridden by recollection and is therefore manifested as an increase in the false alarm rate.

Parallel findings are obtained when younger adults are compared with older adults in a no-deadline (or long-deadline) retrieval condition. Older adults, like younger adults who are rushed at retrieval, exhibit an increase in

both hit and false alarm rates after repetition (Benjamin, 2001). The explanation is the same as with the deadline manipulation: Older adults often have breakdowns in recollective processes, and when recollection fails during retrieval, the enhanced familiarity arising from repeated presentations leads to increases in hit rates and false alarm rates (Jacoby, 1999; Jennings & Jacoby, 1997; Postman, 1982). This set of findings has been referred to as the *ironic effect of repetition* (Jacoby, 1999), to highlight the observation that repetition at study can be a double-edged sword, by simultaneously helping and hurting memory accuracy.

Another set of repetition-related findings in the DRM false memory literature emerges when one examines the influence of multiple study–test trials on the likelihood of recalling or recognizing a nonpresented semantic associate. Whereas younger adults attenuate the probabilities of false memory across study–test trials (McDermott, 1996), older adults are consistently unable to do so (Budson, Daffner, Desikan, & Schacter, 2000; Kensinger & Schacter, 1999; Watson, McDermott, & Balota, 2004). The standard explanation for this difference in participant populations is similar to the idea considered previously: Older adults have deficits in controlled, recollective processing and are unable to use this set of processes to override the tendency for false recall or recognition.

The present experiments used both repetition during encoding (in Experiments 1 and 2) and repetition of study–test trials (Experiment 1) to examine the dual effects of familiarity and recollection in contributing to the erroneous recall and recognition of pragmatic inferences. The logic behind the experiments draws heavily from the aforementioned studies, but the findings from this work have implications for theories of text processing, too, as will be considered in the General Discussion.

EXPERIMENT 1

Three questions are asked in Experiment 1. First, would participants ever choose the pragmatic inference (e.g., *broke*) as having been studied, when the correct alternative was also shown during the test phase? Second, what effects do age and repetition during encoding have on false recognition of pragmatic inferences? On the tentative assumption that robust false recognition would be observed with this two-alternative forced-choice procedure, dual-process theories would lead to specific predictions regarding an interactive effect. Specifically, repetition should help younger adults, as evidenced by both an enhanced hit rate and a decreased false alarm rate. For older adults, however, the prediction is that repetition would elicit a small increase in hit rate and either a parallel increase (Benjamin, 2001; Jacoby, 1999) or no effect (Schacter, Koutstaal, Gross, Johnson, & Angell, 1997) on false alarm rate. The third question addressed by this experiment is whether older adults can reduce false memories of pragmatic inference through multiple study–test opportunities. Or would older adults fail to reduce such false memories, as in the DRM paradigm?

Method

Participants. Twenty-four younger adults (ages 18–22 years, mean = 20.0) participated in return for partial fulfillment of a course research requirement at Washington University. Twenty-four community-dwelling older adults (ages 68–91 years, mean = 77.5) were recruited from the participant pool established by the Aging and Development program at Washington University. They were paid \$10 each for their participation.

Materials. Forty-eight sentences were chosen from a pool of 70, which were normed in a pilot study. This pool of 70 sentences contained the most effective sentences in eliciting recall of pragmatic inferences (henceforth, false recall) from Brewer's (1977) report ($n = 35$) and other sentences constructed by the second author ($n = 35$). The resulting pool of 48 sentences consisted of the most effective sentences (at inducing false recall) on the basis of our pilot work. Data from the pilot work indicated that the probability of false recall for these sentences ranged from .21 to .91, with an average rate of .47 (see the Appendix). Six sentences *without* pragmatic implications were created as practice sentences.

For counterbalancing purposes, the 48 target sentences were divided into four sets of 12 sentences. Sentences in each set were, on average, equally effective in eliciting false recall. The practice sentences were always presented as the first six trials during the test phase, to provide the participants practice using the voice-activated relay.

Sentence fragments ($n = 54$) were created for the test phase. The fragments left out the critical information (e.g., *The karate champion _____ the cinder block*). In some cases ($n = 16$), the critical information was one word (as in the *karate champion* example, which required the participants to say *hit*); in other cases ($n = 32$), the critical information included multiple words (e.g., *The unskillful skateboarder lost his balance on his skateboard* required the participants to respond *lost his balance on*). Three response alternatives appeared below each sentence fragment. The three alternatives were (1) the correct answer, (2) the pragmatic inference, and (3) NEW. Sentence sets and spatial arrangement of the first two alternatives were counterbalanced across the participants; NEW was always the third alternative.

Procedure. Participants were tested one at a time. They were asked to study a series of sentences for an upcoming memory test. During encoding, younger adults saw the sentences for 2.5 sec each, whereas older adults viewed the sentences for 3.5 sec. We used a faster presentation duration for younger adults, with the intention of better equating their performance with that of older adults (Salthouse, 1996).² Half of the sentences were presented once; half were presented three times. The order of presentation was random, with the stipulation that the same sentence never appeared on consecutive trials. Practice sentences were randomly intermixed with the target sentences in the study phase. The participants saw 60 sentences at study (12 once-presented targets; 12 thrice-presented targets; 3 once-presented practice sentences; 3 thrice-presented practice sentences).

During the test, the participants were asked to articulate the correct answer into the microphone or to say "new" if the sentence had not been studied. Following the first test (Trial 1), the participants were informed that they would study and be tested on the same sentences in the same order for a second time (Trial 2).

Results and Discussion

Trial 1. Table 1 and the left panel of Figure 1 display the results from the first test trial. Both younger and older adults chose the (incorrect) pragmatic inference option with a high probability (.33 and .38, respectively), collapsed across one and three study presentations. Further, both younger and older adults enhanced their correct recognition probabilities with three (compared with one) study presentations. Younger adults took advantage of

repetition during encoding to attenuate false recognition of pragmatic inferences, but older adults did not. These impressions are quantified by a 2 (younger, older) \times 2 (correct, false) \times 2 (one, three study presentations) mixed ANOVA. The threshold for statistical significance was set at $p < .05$, two-tailed, unless otherwise noted. Partial eta squared (η_p^2) indicates effect size.

All inferential statistics are reported both on the basis of participant variance (F_1 , t_1) and item variance (F_2 , t_2). In the item analyses, we collapsed across participants instead of items/sentences (hence $N = 48$). In virtually every case, the results of the two approaches lead to the same conclusion; in the cases in which there is a slight discrepancy, it is likely due to a greater variability in the item data. Although discussion tends to focus on the participant analyses, none of the fundamental conclusions differ as a function of the analysis approach.

Main effects for number of study presentations [$F_1(1,46) = 75.70$, $\eta_p^2 = .62$; $F_2(1,94) = 34.23$, $\eta_p^2 = .27$]; response type (correct, false) [$F_1(1,46) = 33.09$, $\eta_p^2 = .42$; $F_2(1,94) = 18.56$, $\eta_p^2 = .17$]; and age group [$F_1(1,46) = 9.49$, $\eta_p^2 = .17$; $F_2(1,94) = 10.80$, $\eta_p^2 = .10$] were significant. The interaction between response type and age group was significant in the participant-based analysis [$F_1(1,46) = 6.54$, $\eta_p^2 = .12$] but only marginally significant in the item-based analysis [$F_2(1,94) = 2.80$, $p = .10$]. The interaction between number of study presentations and response type [$F_1(1,46) = 22.75$, $\eta_p^2 = .33$; $F_2(1,94) = 22.72$, $\eta_p^2 = .20$] was also significant. The interaction between number of study presentations and age group was marginally significant by participants [$F_1(1,46) = 3.15$, $p = .08$] and did not reach significance by items [$F_2(1,94) = 2.30$, $p = .13$]. Most importantly, the reliable three-way interaction suggested that the number of study presentations exerted different effects on accurate and false recognition, depending on age group [$F_1(1,46) = 10.94$, $\eta_p^2 = .19$; $F_2(1,94) = 9.83$, $\eta_p^2 = .10$]. To better understand this three-way interaction, we conducted separate ANOVAs for correct and false (inference) recognition probabilities.

For correct recognition rates (presented in the left half of the left panel of Figure 1), number of study presentations exerted quantitatively (but not qualitatively) different effects for younger and older adults [$F_1(1,46) = 6.19$, $\eta_p^2 = .12$; $F_2(1,94) = 4.58$, $\eta_p^2 = .05$]. Specifically, younger adults benefited more from the additional study presentations than did older adults. That is, younger adults' correct recognition probability rose from .48 to .70 [$t_1(23) = 6.60$, $SEM = .03$; $t_2(47) = 6.57$, $SEM = .03$] after three study presentations, whereas older adults showed a more modest benefit (correct recognition probability rose from .43 to .54) [$t_1(23) = 3.65$, $SEM = .03$; $t_2(47) = 2.51$, $SEM = .04$].

For false recognition rates (presented in the right half of the left panel of Figure 1), number of study presentations exerted different effects on younger and older adults [$F_1(1,46) = 14.07$, $\eta_p^2 = .23$; $F_2(1,94) = 14.95$, $\eta_p^2 = .14$]. This interaction is of a very different nature than that

Table 1
Mean Recognition Probabilities in Experiment 1 (Test Trial 1)

Age Group	Study Presentations	Targets				Practice Correct
		Correct	False/ Inference	New	Spoiled*	
Younger	One	.48	.39	.11	.02	.71
	Three	.70	.28	.01	.01	.90
	Zero	.04**	.06	.89	.01	
Old	One	.43	.36	.18	.04	.65
	Three	.53	.41	.02	.03	.76
	Zero	.01**	.02	.95	.02	

**Spoiled* refers to accidental voice-activated relay trigger trials (e.g., coughing, changing answer in the middle of a response, etc.). **These values refer to the probabilities that when presented with two versions of a nonstudied sentence, participants chose the version that had been studied by other participants. This is one type of false alarm. For example, this type of response happened when a person had not studied the karate champion sentence but subsequently claimed that the word "hit" had been presented with that nonstudied sentence.

for correct recognition. Whereas the interaction between study presentations and age group on accurate recognition was of the quantitative kind, the interaction between these two variables on false recognition was qualitative. Although younger adults were able to reduce false recognition rates after three study presentations (.28) compared with one study presentation (.39) [$t_1(23) = 3.7$, $SEM = .03$; $t_2(47) = 4.28$, $SEM = .03$], older adults were unable to do so. In fact, their false recognition probability *rose* from .36 to .40 after three (compared with one) study presentations, although this comparison was not significant in the conventional sense [$t_1(23) = 1.55$, $p = .07$ (one-tailed); $t_2(47) = 1.04$, $p = .15$ (one-tailed)].

To summarize, both the participant and item analyses revealed a significant three-way interaction between response type, number of study presentations, and age group. Moreover, results from separate ANOVAs for accurate and false recognition showed that the number of study presentations exerted different effects as a function of age group (both for correct recognition and false recognition).

Trial 2. Younger and older adults both exhibited improvement in their recognition probabilities in the second test trial compared with the first trial, as can be seen in Figure 1. Younger adults exhibited similar patterns on Trial 2 as they did on Trial 1, but with higher correct and lower false recognition probabilities. Unlike in the Trial 1

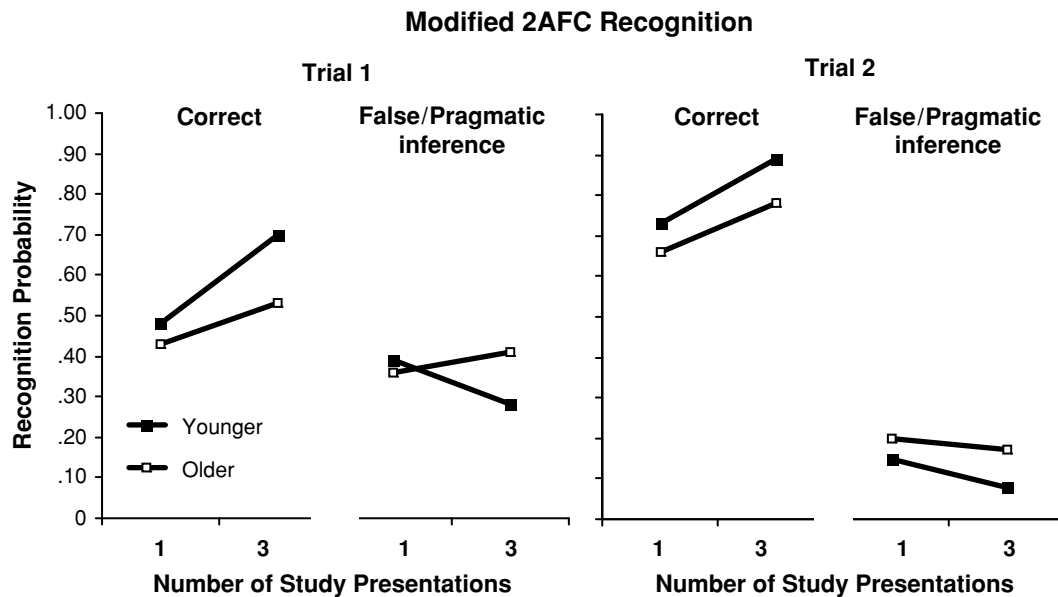


Figure 1. Mean probabilities of correct and false (pragmatic inference) recognition as a function of number of study presentations and age for the first (left panel) and second trial on the modified two-alternative forced-choice recognition test (Experiment 1).

data, older adults showed no signs of ironic effects of repetition. In fact, their performance was very similar to that of younger participants.

The biggest difference between the data from the first and second trials is that on Trial 2, age group was no longer the potent variable that it had been on Trial 1. For example, age group no longer produced a main effect in the participant analysis [$F_1(1,46) = .87, p = .36$], although it did produce a small effect in the item analysis [$F_2(1,94) = 5.09, \eta_p^2 = .05$]. Age group also did not interact with number of study presentations (both F_1 and $F_2 < 1$). There also was no three-way interaction among the variables [$F_1(1,46) = 1.38, p = .25$; $F_2(1,94) = 2.16, p = .14$].

Age group did, however, produce an interaction with response type (correct, false) [$F_1(1,46) = 10.88, \eta_p^2 = .19$; $F_2(1,94) = 9.28, \eta_p^2 = .09$]. This interaction indicated that, when collapsed across number of study presentations, correct recognition probabilities were lower for older (.72) than for younger adults (.81), but false recognition probabilities were higher for older (.19) than for younger adults (.12). Number of study presentations [$F_1(1,46) = 28.71, \eta_p^2 = .38$; $F_2(1,94) = 25.85, \eta_p^2 = .22$] and response type [$F_1(1,46) = 634.07, \eta_p^2 = .93$; $F_2(1,94) = 534.70, \eta_p^2 = .85$] also resulted in significant main effects, and these two variables interacted with each other [$F_1(1,46) = 32.81, \eta_p^2 = .42$; $F_2(1,94) = 29.37, \eta_p^2 = .24$]. This interaction indicated that the difference between correct and false recognition was smaller for once-presented sentences (.69 - .18 = .51) than for thrice-presented sentences (.84 - .13 = .71).

In summary, Experiment 1 demonstrates several interesting findings. First, the false memory effect in this paradigm is quite robust: Even when presented with the correctly worded sentence, the participants often chose the incorrect sentence (containing the pragmatic inference) as having been studied. Second, the ironic effect of repetition was observed on Trial 1 for older but not for younger adults. Third, older and younger adults benefited from a second study-test opportunity: Their hit rates were higher and false alarm rates were lower relative to Trial 1. These findings are considered more fully in the General Discussion.

EXPERIMENT 2

In Experiment 2, we sought to replicate and extend the findings from Trial 1 of Experiment 1. Specifically, we changed from a recognition test to a cued recall test. Although the dual-process framework originated in recognition, it is a useful framework for considering other memory tests as well. As Jacoby (1991) has argued, memory tests should not be considered process-pure. Familiarity-like processes play a significant role in performance in cued recall and even free recall (Hamilton & Rajaram, 2003; McDermott, 2006; Tulving, 1985). Therefore, it is important to see whether our results from Experiment 1 (using recognition) would generalize. Further, we replaced the participant variable (age) with a manipulation of response speed at retrieval (cf. Jacoby, 1999). Hence, we tested only

younger adults under a speeded and an unspeeded condition. On the basis of prior studies, we hypothesized that an interaction would emerge between response speed and repetition that was similar to the interaction between age and repetition in Trial 1 of Experiment 1.

Method

Participants. Eighty younger adults (ages 18–22 years, mean = 19.24) from Washington University participated.

Design. A 2 (instruction: speeded, unspeeded) \times 3 (number of study presentations: 0, 1, 3) design was used. Instruction was manipulated between participants (40 participants per condition), number of presentations was manipulated within participants, and materials were counterbalanced across participants.

Materials. Materials for the cued recall test consisted of the same 54 sentence fragments used in Experiment 1. Half of the sentence fragments corresponded to sentences that the participants had seen during the study phase (old); the remaining were new sentence fragments. The participants were instructed to recall the blank part of the sentence fragments.

Procedure. The participants were tested one at a time; all of the participants studied 24 target sentences for 3.5 sec each. All other presentation protocols were the same as in Experiment 1.

The test instructions informed the participants that sentence fragments would be presented on the computer screen and that they should recall the missing part word for word. They were asked not to guess and to articulate their answer into the microphone or to say “new.” The participants were further instructed to say “forget” in the cases in which they knew they had studied the sentence but felt they could not remember the exact wording of the studied sentence.

The participants in the speeded condition were told to respond within 3 sec of the onset of the sentence fragment. If their response time exceeded the 3-sec limit, the words “TOO SLOW!” appeared on the computer screen. The participants in the unspeeded condition were informed that there was no time limit for responding. The logic for choosing a 3-sec window instead of some shorter duration (e.g., the response window that Benjamin, 2001, used was 750 msec) was to take into account the time the participants needed to read the sentence fragment.

We opted for a speeded/unspeeded manipulation over a deadline manipulation to avoid possible trial selection problems. All responses (even if they occurred after 3 sec in the speeded condition) were scored. (The same pattern of results emerged, however, when only responses given within the deadline were considered.) In the speeded condition, the slowest response recorded for each individual participant (excluding forget, new, and accidental voice-activated relay trigger trials) ranged from 2,228 msec (slowest response of the fastest participant) to 6,242 msec (slowest response of the slowest participant), and the average was 3,686 msec. This number, the average slowest response for participants in the speeded condition, was still faster than 4,400 msec, the overall average response time in the unspeeded condition.

Results and Discussion

Scoring procedures separated responses into five categories. A response could be *correct* (when recall was verbatim, although synonyms, tense, and singular-plural form switching were acceptable), *false* (when recall fit the a priori defined pragmatic inference), *other* (when the recalled information did not fit either the correct or prespecified inference category), *forget*, or *new*. To keep scoring criteria consistent across all sentences, we identified a priori specific responses that would be accepted as inferences for each sentence. For example, responses that

would be classified as inferences for *The karate champion hit the cinder block* were *broke*, *split*, or *smashed*. All three responses include the idea of *breaking*, an outcome that is not logically necessitated by *hit*. Therefore, if a participant recalled *kicked*, which did not indicate clearly that the karate champion had broken apart the cinder block, the trial would be scored as *other*. Indeed, even if a participant recalled *demolished*, which could also be considered a pragmatic inference, we scored the response as incorrect, because it was not defined a priori. We note, though, that responses such as this rarely occurred, and an alternate scoring approach, in which we scored such responses as pragmatic inferences, did not change any of the reported patterns of results. As can be seen under the "Other" column in Table 2, the probabilities of participants reporting other responses not in our predefined list were very low.

As can be seen in Figure 2 and Table 2, when participants were not speeded, three study presentations (relative to a single study presentation) led to a higher probability of accurate recall and a lower probability of false recall. In the speeded condition, however, a different pattern emerged: Although three study presentations enhanced the likelihood of accurate recall (relative to a single presentation), it also led to a higher probability of false recall (at least numerically).

Statistical analyses are in line with the above description. As predicted, a 2 (test instructions: speeded, unspeeded) \times 2 (response type: correct, false) \times 2 (number of study presentations: 1, 3) repeated measures ANOVA exhibited a three-way interaction by participants [$F_1(1,78) = 4.12$, $\eta_p^2 = .05$] and a marginally significant interaction by items [$F_2(1,94) = 3.65$, $p = .06$, $\eta_p^2 = .04$]. A $2 \times 2 \times 2$ ANOVA was conducted instead of a $2 \times 2 \times 3$ (number of study presentations: 0, 1, 3) ANOVA because the prob-

abilities of false recall to the new (0-study presentation) sentence fragments were (as predicted) near floor in both the speeded (.01 false recall) and the unspeeded (.01 false recall) conditions, and this 0-presentation condition was included primarily for counterbalancing and base-rate purposes. The following analyses consider the correct and false (inference) recall separately, to better understand the three-way interaction.

Correct recall. Both number of study presentations (1, 3) and response speed (speeded, unspeeded) produced reliable main effects [$F_1(1,78) = 142.69$, $\eta_p^2 = .65$; $F_2(1,94) = 98.43$, $\eta_p^2 = .51$ for number of presentations] and [$F_1(1,78) = 9.59$, $\eta_p^2 = .11$; $F_2(1,94) = 10.36$, $\eta_p^2 = .10$ for response speed]. However, these two variables did not interact ($F_1 < 1$ and $F_2 = 1.0$). In other words, number of study presentations and response speed produced similar effects on correct recall probability. Specifically, three presentations led to greater probabilities of correct recall than a single presentation under both the unspeeded response condition [$t_1(39) = 9.38$, $SEM = .03$; $t_2(47) = 6.61$, $SEM = .03$] and the speeded response condition [$t_1(39) = 7.59$, $SEM = .03$; $t_2(47) = 7.40$, $SEM = .03$].

False recall of pragmatic inferences. Unlike results from accurate recall, number of presentations and response speed exerted different effects on false (inference) recall [$F_1(1,78) = 7.49$, $\eta_p^2 = .09$; $F_2(1,94) = 5.10$, $\eta_p^2 = .05$]. When participants were not speeded at the retrieval phase, three study presentations led to a lower likelihood of false recall ($M = .27$) relative to a single presentation ($M = .34$) [$t_1(39) = 3.18$, $SEM = .02$; $t_2(47) = 2.08$, $SEM = .03$]. When participants were speeded at the retrieval phase, although the probability of false recall for three study presentations (.37) was numerically greater than that for one study presentation (.33), the difference was not reliable [$t_1(39) = 1.10$, $p = .28$; $t_2(47) = 1.13$, $p = .27$]. This unreliable difference might have been a result of weak statistical power (observed power = .19 for the participant analysis and .20 for the item analysis).

In sum, Experiment 2 conceptually replicated the findings of test Trial 1 in Experiment 1 by demonstrating the ironic effects of repetition in false recall of pragmatic inferences.

GENERAL DISCUSSION

The experiments reported here demonstrate several key findings with respect to memory for pragmatic inferences. First, both familiarity and recollection play a role in the pragmatic inference effect. This conclusion emerges from the ironic effects of repetition, which were demonstrated in older adults (Experiment 1) and in younger adults who were speeded during retrieval (Experiment 2). Second, participants in both age groups were able to take advantage of a second study-test trial, both in terms of an enhanced hit rate and a decreased false alarm rate (relative to Trial 1). Third, even when faced with the original sentence and the pragmatic inference on a forced-choice recognition test, participants chose the inference option almost as often as the correct option; this observation speaks volumes about the robustness of the inference that par-

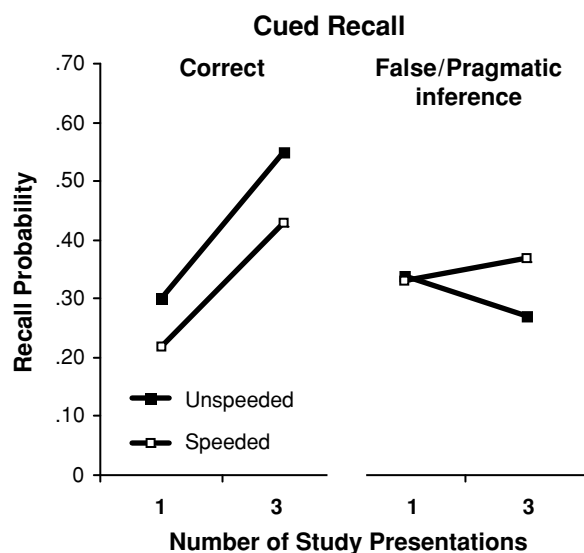


Figure 2. Mean probabilities of correct and false (pragmatic inference) cued recall for younger adults in Experiment 2 as a function of number of study presentations (1 or 3) and test conditions (speeded or unspeeded).

Table 2
Mean Recall Probabilities in Experiment 2 (Younger Adults)

Test Condition	Presentations	Targets						Practice Correct
		Correct	False/ Inference	Other	Forget	New	Spoiled*	
Unspeeded	One	.30	.34	.02	.27	.07	.00	.67
	Three	.55	.27	.04	.13	.01	.00	.75
	Zero		.01		.04	.95		
Speeded	One	.22	.33	.04	.28	.11	.02	.47
	Three	.43	.37	.04	.13	.02	.01	.59
	Zero		.01		.08	.91		

**Spoiled* refers to accidental voice-activated relay trigger trials (e.g., coughing, changing answer in the middle of a response, etc.).

ticipants generated. We now consider these three principal findings in relation to the literature.

Ironic Effects of Repetition

Study repetition, in combination with other variables, such as response deadline or participant's age, has been used in clever ways to tease apart the relative contributions of recollection- and familiarity-based processes in memory performance (Bartlett, Strater, & Fulton, 1991; Benjamin, 2001; Jacoby, 1999; Jacoby, Jones, & Dolan, 1998; Jones & Jacoby, 2001; McDermott, 2006; McElree et al., 1999). Specifically, when recollective processes are allowed to work at retrieval, as when younger adults are given ample time to respond, such processes can be used to override the enhanced familiarity brought about by study repetition to reduce false memories. When recollective processes are relatively unavailable during retrieval, as when younger adults are speeded or when older adults are tested, level of familiarity drives level of false memories.

The finding of ironic effects of repetition in the present experiments therefore suggests that recollection for the meaning-preserving aspects of a sentence can be used to override the tendency for recall or recognition of the pragmatic inferences. However, when recollective processes are compromised, the tendency to recall or recognize the pragmatic inference increases, because participants have to respond based largely on familiarity (Toth, 1996; Yonelinas & Jacoby, 1994). The manipulation of repetition with response deadline or with variation in participants' age, therefore, offers insight into the processes underlying memory for pragmatic inferences.

Another noteworthy finding is that even when older adults were tested or when younger adults were placed under a speeded response condition, repeated presentations at encoding still led to higher probabilities of accurate recall and recognition, compared with a single presentation. There are several possible reasons for this finding. The first possibility is that fast responding and cognitive aging do not eliminate recollection, a possibility that has been suggested by several researchers (Jennings & Jacoby, 2003; Toth, 1996). The second possible explanation is that familiarity can contribute to accurate responding even for the sentences in the present paradigm (Chan & McDermott, 2006). Some sentences may fail to

elicit pragmatic inferences for some participants at encoding (after all, none of our sentences led to 100% average false recall in the pilot study). Moreover, a pragmatic inference generated at encoding may be lost during the retention interval or not accessed at the retrieval phase. Finally, as we have discussed, even when the pragmatic inference is retrieved, it may be overridden by recollection of the accurate response.

It is possible that familiarity can lead to multiple traces for a given encoding episode. For example, after reading *the karate champion hit the cinder block*, concepts such as *broke* and *hit* can both be activated and registered through familiarity; response competition between these concepts can lead to either the correct or the pragmatic inference response. It is likely unrealistic to think of familiarity as a process that can only lead to one response (or even one class of response) for each encoding episode. Future research is needed to further explore this hypothesis. For the time being, we conclude that it is crucial to consider recollection and familiarity as flexible processes that can lead to different (and perhaps even multiple) responses, depending on the encoding and retrieval conditions.

It is important to note that the two processes implicated in the dual-process framework most fundamentally represent a slower process that is subject to cognitive control (i.e., recollection) and a faster process that is relatively automatic (i.e., familiarity). The term *familiarity*, as used in the present article, can easily be replaced by other terms, such as *gist processing* (Reyna & Brainerd, 1995).

Benefits of Multitrial Learning

Both younger and older adults were able to diminish the propensity for false recognition of pragmatic inferences when given multiple study–test trials. In this way, our findings are somewhat inconsistent with findings from the DRM paradigm, which indicate that older adults are consistently *unable* to take advantage of multiple study–test trials, even when given as many as five iterations (Kensinger & Schacter, 1999; Watson et al., 2004). We offer a possible explanation for this discrepancy.

In Experiment 1, unlike in the DRM studies, we used a modified two-alternative forced-choice recognition task. During the test phase, participants were presented with sentences that differed only on the critical word(s)—e.g.,

the karate champion hit the cinder block or the karate champion broke the cinder block (with the alternative to say “new,” for nonstudied sentences). Hence, in the subsequent study phase, participants were fully informed with respect to exactly what portion of the sentences to devote attention to for the upcoming test. This feature differs from the DRM multitrial experiments, in which participants are not given any information at test regarding the experimental design or potential lures; instead, they simply perform a free recall or a free choice recognition test, with no clues given with respect to potential false memories.

The idea that people can selectively focus their attentional resources on specific information during multitrial learning is not new. One theoretical explanation that is particularly relevant to the present argument is the *knowledge of recallability hypothesis* (Thompson, Wenger, & Bartling, 1978). This hypothesis suggests that in a multitrial learning paradigm, participants can use their performance on the first recall trial to evaluate the recallability of an item, and when subsequent encoding trials are presented, they can use this knowledge of recallability to allocate different resources to the well-encoded and poorly encoded items. Applying this logic to the present study, older adults could pay special attention to the sentences that they did not remember on the first test during the second study trial. In fact, they could focus their encoding resources on the very information (i.e., the blanks in the test) that is required for recall. Indeed, the DRM literature has suggested that older adults can take advantage of instructions designed to focus attention on avoiding false memories (Watson et al., 2004).

Relation to Discourse Research

Sentences such as those used in the present study have been the focus of many studies in the discourse literature. Many such studies have focused on delineating types of inferences and identifying which types of inferences are routinely generated by readers online during the reading process. One example that is particularly relevant to the present study is predictive inference, in which the reader may infer a plausible outcome (e.g., that the karate champion *broke* the cinder block after reading that the champion *hit* the cinder block).

Considerable debate has arisen with respect to whether predictive inferences are generated online. Some researchers have argued that such inferences are generated only minimally, if at all (McKoon & Ratcliff, 1986, 1992; Potts, Keenan, & Golding, 1988; Singer & Ferreira, 1983), and others conclude that the inferences are indeed generated but that they quickly dissipate due to decay or interference (see, e.g., Keefe & McDaniel, 1993). One theme emerging from the literature is that online inference generation for predictive inferences is tenuous, seems to depend on the specific experimental conditions used, and (if it occurs at all) dissipates quickly, over a few seconds or intervening items. Most studies examining this question have focused on reading time or priming measures as dependent variables.

Our report offers a departure from the aforementioned approaches, in that it examines the extent to which such

inferences are carried forward over considerable delays, and shows that they are robust enough to exert strong influence on later retrieval (see also Brewer, 1977; Johnson et al., 1973), even to the extent that participants choose the inference over the studied option on a recognition test. It is possible that the effects we observed on our long-term memory task emerge as a combination of factors from the encoding, storage, and retrieval phases (cf. Harris & Monaco, 1978; Johnson et al., 1973; Singer & Ferreira, 1983), but the specific pattern of results (as a function of repetition, age, and response deadline) suggests that the reading/encoding phase indeed plays a role in the later memory for the inferences. It appears that the momentary activation and fast dissipation of the inference is not the end of its influence (see, e.g., Klin, Guzmán, & Levine, 1999).

Although a distinction has been made between forming an inference and other fast-acting processes tied to associative or semantic activation (Klin et al., 1999), the idea that fast-acting processes may exert a long-lasting influence on memory is similar to a puzzle that has been considered recently in the DRM false memory literature. Here, variables that are well known to modulate levels of semantic activation have been shown to also modulate the likelihood of false recognition and false recall, long after associative activation is typically thought to have faded (Chan et al., 2005; McDermott & Watson, 2001; Roediger, Balota, & Watson, 2001). One possibility is that associative activation is not as short-lived as commonly thought (Balota & Paul, 1996) and that the current conceptualization of activation may be incomplete. At this point, we simply want to note the parallels between the two literatures. The suggestion emerging from the commonalities in these two literatures is that seemingly fleeting processes (associative activation and inference formation) may assert profound influences on memory retrieval over long delays. In this case, these processes lead to memory errors, but see Chan, McDermott, and Roediger (in press) for an example in which such processes benefit memory.

In addition to considering empirical findings with respect to predictive inferences (as considered above), we can also consider broad-based theoretical perspectives on discourse processing. Within this literature, text is said to be represented in memory in three levels: the surface level, the text base, and the situation model (van Dijk & Kintsch, 1983; see Fletcher, 1994, for a review). The surface level of representation contains the exact words and phrases that participants studied. The text base does not contain the exact wording, but it preserves the semantic content described by the text. The situation model, on the other hand, contains the content that is integrated into the existing knowledge base of the reader but is detached from the text structure itself. A pragmatic inference would be part of the situation model but not the surface level or the text base.

Consistent with the present study, findings in discourse research have revealed age invariance in the formation of inferences, provided that working memory capacity in older adults is not taxed (Hasher & Zacks, 1988; Light & Albertson, 1993; Radvansky, Copeland, & Zwaan, 2003;

Radvansky, Gerard, Zacks, & Hasher, 1990; Radvansky, Zwaan, Curiel, & Copeland, 2001). Specifically, older adults and younger adults show equivalent retrieval of the situation model, but older adults are impaired in the retrieval of surface and text base levels (Radvansky et al., 2001). Hence, older adults may rely heavily on information available from the situation model. Applying this idea to the present results, older adults may show the ironic effect of repetition because they fail to remember text base information even after multiple study opportunities, but the availability of information from the situation model has increased with study repetitions, which is evidenced as a higher likelihood of false recognition.

The idea that older adults are less efficient at remembering surface-based materials (e.g., surface level and text base) is in line with memory-based theories, which generally suggest that older adults have trouble remembering verbatim or perceptual information (Johnson, Hashtroudi, & Lindsay, 1993; Yonelinas, 2002). It is unclear, however, how discourse theories would handle the data from Experiment 2, in which only younger adults were tested and a response deadline manipulation was implemented. Future research may investigate the effect of variables, such as speed of responding, on the usage of information from the various levels of text/speech representation.

Conclusion

In conclusion, the present experiments use prose materials to study erroneous memory for pragmatic inferences. Our results demonstrate that the erroneous retrieval of pragmatic inferences is a robust phenomenon that can be overridden, to some extent, under conditions in which recollective processes can be fully utilized. More generally, the pragmatic inference paradigm offers fertile ground for enhancing our understanding of how everyday types of false memories can arise.

REFERENCES

- ATKINSON, R. C., & JUOLA, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 583-612). New York: Academic Press.
- BALOTA, D. A., & PAUL, S. T. (1996). Summation of activation: Evidence from multiple primes that converge and diverge within semantic memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 827-845.
- 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.
- BRANSFORD, J. D., & FRANKS, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, **2**, 331-350.
- BREWER, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition*, **5**, 673-678.
- 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.
- CHAN, J. C. K., McDERMOTT, K. B. (2006). Remembering pragmatic inferences. *Applied Cognitive Psychology*, **20**, 633-639.
- CHAN, J. C. K., McDERMOTT, K. B., & ROEDIGER, H. L. III (in press). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General*.
- CHAN, J. C. K., McDERMOTT, K. B., WATSON, J. M., & GALLO, D. A. (2005). The importance of material-processing interactions in inducing false memories. *Memory & Cognition*, **33**, 389-395.
- DEESE, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58**, 17-22.
- FLETCHER, C. R. (1994). Levels of representation in memory for discourse. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 589-607). San Diego: Academic Press.
- GARRY, M., MANNING, C. G., LOFTUS, E. F., & SHERMAN, S. J. (1996). Imagination inflation: Imagining a childhood event inflates confidence that it occurred. *Psychonomic Bulletin & Review*, **3**, 208-214.
- GOFF, L. M., & ROEDIGER, H. L., III (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recollections. *Memory & Cognition*, **26**, 20-33.
- HAMILTON, M., & RAJARAM, S. (2003). States of awareness across multiple memory tasks: Obtaining a "pure" measure of conscious recollection. *Acta Psychologica*, **112**, 43-69.
- HARRIS, R. J., & MONACO, G. E. (1978). Psychology of pragmatic implication: Information processing between the lines. *Journal of Experimental Psychology: General*, **107**, 1-22.
- 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* (pp. 193-225). San Diego: Academic Press.
- HYMAN, I. E., HUSBAND, T. H., & BILLINGS, F. J. (1995). False memories of childhood experiences. *Applied Cognitive Psychology*, **9**, 181-197.
- JACOBY, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, **30**, 513-541.
- JACOBY, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 3-22.
- JACOBY, L. L., JONES, T. C., & DOLAN, P. O. (1998). Two effects of repetition: Support for a dual-process model of knowledge judgments and exclusion errors. *Psychonomic Bulletin & Review*, **5**, 705-709.
- JENNINGS, J. M., & JACOBY, L. L. (1997). An opposition procedure for detecting age-related deficits in recollection: Telling effects of repetition. *Psychology & Aging*, **12**, 352-361.
- JENNINGS, J. M., & JACOBY, L. L. (2003). Improving memory in older adults: Training recollection. *Neuropsychological Rehabilitation*, **13**, 417-440.
- JOHNSON, M. K., BRANSFORD, J. D., & SOLOMON, S. K. (1973). Memory for tacit implications of sentences. *Journal of Experimental Psychology*, **98**, 203-205.
- JOHNSON, M. K., HASHTROUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28.
- 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.
- KEEFE, D. E., & McDANIEL, M. A. (1993). The time course and durability of predictive inferences. *Journal of Memory & Language*, **32**, 446-463.
- KENSINGER, E. A., & SCHACTER, D. L. (1999). When true memories suppress false memories: Effects of ageing. *Cognitive Neuropsychology*, **16**, 399-415.
- KLIN, C. M., GUZMÁN, A. E., & LEVINE, W. H. (1999). Prevalence and persistence of predictive inferences. *Journal of Memory & Language*, **40**, 593-604.
- LIGHT, L. L., & ALBERTSON, S. A. (1993). Comprehension of pragmatic implications in young and older adults. In L. L. Light & D. M. Burke (Eds.), *Language, memory, and aging* (pp. 133-153). New York: Cambridge University Press.
- LOFTUS, E. F., & PALMER, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning & Verbal Behavior*, **13**, 585-589.
- McDERMOTT, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory & Language*, **35**, 212-230.

- McDERMOTT, K. B. (2006). Paradoxical effects of testing: Repeated retrieval attempts enhance the likelihood of later accurate and false recall. *Memory & Cognition*, **34**, 261-267.
- 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.
- McELREE, B., DOLAN, P. O., & JACOBY, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 563-582.
- McKOOK, G., & RATCLIFF, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 82-91.
- McKOOK, G., & RATCLIFF, R. (1992). Inference during reading. *Psychological Review*, **99**, 440-466.
- POSTMAN, L. (1982). An examination of practice effects in recognition. *Memory & Cognition*, **10**, 333-340.
- POTTS, G. R., KEENAN, J. M., & GOLDING, J. M. (1988). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory & Language*, **27**, 399-415.
- RADVANSKY, G. A., COPELAND, D. E., & ZWAAN, R. A. (2003). Aging and functional spatial relations in comprehension and memory. *Psychology & Aging*, **18**, 161-165.
- RADVANSKY, G. A., GERARD, L. D., ZACKS, R. T., & HASHER, L. (1990). Younger and older adults' use of mental models as representations for text materials. *Psychology & Aging*, **5**, 209-214.
- RADVANSKY, G. A., ZWAAN, R. A., CUIEL, J. M., & COPELAND, D. E. (2001). Situation models and aging. *Psychology & Aging*, **16**, 145-160.
- REINITZ, M. T., LAMMERS, W. J., & COCHRAN, B. P. (1992). Memory-conjunction errors: Miscombination of stored stimulus features can produce illusions of memory. *Memory & Cognition*, **20**, 1-11.
- REYNA, V. F., & BRAINERD, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning & Individual Differences*, **7**, 1-75.
- ROEDIGER, H. L., III, BALOTA, D. A., & WATSON, J. M. (2001). Spreading activation and arousal of false memories. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95-115). Washington, D.C.: American Psychological Association.
- 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, MEADE, M. L., & BERGMAN, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review*, **8**, 365-371.
- SALTHOUSE, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, **103**, 403-428.
- SCHACTER, D. L., KOUTSTAAL, W., GROSS, M. S., JOHNSON, M. K., & ANGELL, K. E. (1997). False recollection induced by photographs: A comparison of older and younger adults. *Psychology & Aging*, **12**, 203-215.
- SINGER, M., & FERREIRA, F. (1983). Inferring consequences in story comprehension. *Journal of Verbal Learning & Verbal Behavior*, **22**, 437-448.
- SOMMERS, M. S., & LEWIS, B. P. (1999). Who really lives next door: Creating false memories with phonological neighbors. *Journal of Memory & Language*, **40**, 83-108.
- THOMPSON, C. P., WENGER, S. K., & BARTLING, C. A. (1978). How recall facilitates subsequent recall: A reappraisal. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 210-221.
- TOTH, J. P. (1996). Conceptual automaticity in recognition memory: Levels-of-processing effects on familiarity. *Canadian Journal of Experimental Psychology*, **50**, 123-138.
- TULVING, E. (1985). Memory and consciousness. *Canadian Psychology*, **26**, 1-12.
- VAN DIJK, T. A., & KINTSCH, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- 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.
- YONELINAS, A. P., & JACOBY, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology*, **48**, 516-534.

NOTES

1. For the sake of simplicity, we have adopted Brewer's terms, pragmatic and logical inferences, to refer to these two broad ranges of inference. This terminology does not distinguish among many other possible types of inference, such as predictive inference, anaphoric inference, causal antecedent, superordinate and subordinate goals, and so on, although many of the sentences used as materials could be classified as invoking predictive inferences.

2. Variation of presentation duration during study affects recollection and familiarity similarly (Yonelinas, 2002); therefore, a shorter study presentation duration should simply reduce the overall recognition rates relative to a longer study presentation duration. In other words, reducing the presentation duration at study from 3.5 sec to 2.5 sec should not change any patterns of interaction in this experiment. In contrast, as we have mentioned earlier, variation of response speed at retrieval affects recollection disproportionately, while leaving familiarity relatively unaffected (Yonelinas, 2002). In fact, we have collected data on an additional 18 younger adults who viewed the sentences for 3.5 sec each, and they exhibited the same pattern of results as reported here, in Experiment 1. Specifically, for younger adults in this condition, accurate recognition rose from .53 to .70, whereas false recognition dropped from .33 to .26, when comparing three- to one-study presentation.

APPENDIX Stimuli Used in Experiments 1 and 2

Study Sentence	Inference Accepted	Normed False Recall Probability
The track star was the first runner to pass the line in the race.	the finish/goal line	.91
After dropping off her kids at the school, the mother picked up bread.	bought	.89
The charming prince gently put his lips towards Snow White's cheek.	kissed	.86
*The Sherman tank hit the flimsy roadblock.	broke/went through	.80
*The children's snowman vanished when the temperature reached 80.	melted	.73
*The flimsy shelf weakened under the weight of the books.	collapsed	.73
The unskillful skateboarder lost his balance on the skateboard.	fell off	.69
The ugly sisters asked Cinderella to mop the floor.	told/forced	.69
*The absent-minded professor didn't have his car keys.	forgot/lost	.67
The friendly bartender got extra money from the customer.	tips	.64
*The angry rioter threw the rock at the window.	through the window	.62
*The popular radio station picked only hard rock music.	played	.62
*The karate champion hit the cinder block.	broke/smashed/split	.60
*Dennis the Menace sat in Santa's chair and asked for an elephant.	lap	.57
*The painter knocked over the bucket of black paint.	spilled	.57
*The clumsy chemist had acid on his coat.	spilled acid	.57
*The hungry python caught the mouse.	ate	.50
The hungry squirrel was biting the grass.	was eating	.50
*The safe-cracker put the match to the fuse.	lit	.50
The audiophile turned his stereo system to maximum.	stereo system to maximum volume/up loud	.50
The agile cat reached the fish with its claw.	caught	.50
The politician was making an untrue statement about his financial crisis.	lying	.46
The flimsy water gate leaked under the rainstorm.	collapsed	.46
The rat was stopped by the mousetrap.	was caught	.44
The birthday boy blew at the candles.	blew out	.43
*The night watchman took some coffee from his thermos.	poured	.43
*King Kong stood on the Empire State Building.	stood on top of	.42
*The narcotics officer pushed the door bell.	rang	.40
*The new baby stayed awake all night.	cried	.38
The noisy guard dog growled at the trespasser.	barked	.38
The racecar driver hit the wall in the middle of the race.	crashed into the wall	.36
*The little brat stuck the balloon with a pin.	popped the balloon	.36
After suffering the worst injury in her career, the ballet dancer had to quit.	retire	.36
*The barnacles were growing on the ship.	on the side/bottom of the ship	.33
The hypnotist put his fingers together and awakened his client.	snapped his fingers	.33
*The submarine captain thought the submarine should submerge.	ordered the submarine to submerge	.33
*The cat burglar climbed up the apartment building.	climbed up the side of	.31
The Christian closes his eyes and reflects before every meal.	closes his eyes and prays	.30
*The beautiful blonde walked into the dark room and flipped the switch.	flipped the light switch/turned on the light	.27
The studious girl turned the pages of her textbook.	was reading	.27
The suspicious man went into a 7-11 and took a chocolate bar.	stole	.27

APPENDIX (Continued)

Study Sentence	Inference Accepted	Normed False Recall Probability
*The sleepy husband hit the mosquito.	killed	.25
The bully put his fist to the geek.	beat up/hit	.23
*The 100 yard dash began when the referee pulled the trigger.	fired the gun	.23
The cantankerous granny talked to her neighbor.	yelled/shouted at	.22
After chasing the thief for three blocks, the cop finally caught up to the thief.	caught	.22
The cop shouted at the suspect and fired his gun.	shot him/her	.22
The gymnast made a big mistake and might not win the gold medal.	lost/did not win	.21
The painter tore his sketch apart.		Practice
The bus driver witnessed a car accident.		Practice
The Sumo wrestler ate everything in the restaurant.		Practice
The housebuyer signed the contract.		Practice
*The Titanic hit the iceberg and sank to the bottom of the ocean.		Practice
Martin left his suitcase at the airport.		Practice

Note—The underlined sections represent the segments to be retrieved during the test phase. *Sentences selected from Brewer (1977).

(Manuscript received October 4, 2004;

revision accepted for publication June 2, 2005.)