Learning to diminish the effects of proactive interference: Reducing false memory for young and older adults

LARRY L. JACOBY AND CHRISTOPHER N. WAHLHEIM
Washington University, St. Louis, Missouri

MATTHEW G. RHODES
Colorado State University, Fort Collins, Colorado

KAREN A. DANIELS
University of North Carolina, Wilmington, North Carolina

AND

CHAD S. ROGERS
Washington University, St. Louis, Missouri

Results from two experiments revealed that prior experience with proactive interference (PI) diminished PI's effects for both young and older adults. Participants were given two rounds of experience, with different materials, in a situation that produced PI. Comparisons with a control condition showed that the effects of PI on accuracy and on high-confidence intrusion errors (false memory) were reduced on the second round, as compared with those on the first. Also, the ability of confidence to diagnose accuracy of responding improved across rounds. Effects of prior experience with PI depended on feedback given at the time of test (Experiment 1). At least in part, the diminishment of PI resulted from participants' allocating more attention to interference items during study in the second round than in the first (Experiment 2). Implications of the results for interpreting age differences in PI and false memory are discussed.

Proactive interference (PI) refers to the reduction in memory performance for recently learned information resulting from the prior learning of related materials and has been shown to play an important role in forgetting (for a review, see Anderson & Neely, 1996). Investigations of PI have traditionally used a paired-associate learning procedure, in which interference is created by holding cues constant, with the responses being changed between two lists (A–B, A–D). Performance in this interference condition is compared with that in a control condition for which both cues and responses are changed between lists (A–B, C–D) or for which participants first "rest" and then learn cue–response pairs (rest, C–D). The magnitude of PI effects is indexed as the difference in cued-recall accuracy between the control and interference conditions on the final test.

Errors resulting from PI can take the form of a first-list response intruding when participants are asked to produce the response paired with a cue in the second list. Older adults are more susceptible to such interference than are young adults because of their lessened ability to avoid making erroneous responses that have been made highly probable by prior experience (see, e.g., Hasher, Zacks, &

May, 1999; Hay & Jacoby, 1999; May, Hasher, & Kane, 1999). Older adults are also more likely to produce intrusion errors that result from PI with high levels of confidence (for a review, see Dodson & Krueger, 2006; Jacoby, Bishara, Hessels, & Toth, 2005; Jacoby & Rhodes, 2006; Kelley & Sahakyan, 2003). Confidence in intrusion errors is important because older adults' high confidence in such errors makes them vulnerable to memory slips and memory scams (Hay & Jacoby, 1999; Jacoby et al., 2005). An intrusion error is likely to be acted on only if it is held with a high level of confidence. For example, memory for taking a daily medicine yesterday might intrude as a high-confidence memory of having taken the medicine today, a false memory that results in failure to take the medicine today.

The goal of our experiments was to determine whether PI could be reduced by providing participants with multiple experiences dealing with PI. Although the importance of PI effects is widely known, prior research has not examined whether people can adapt their processing in ways that allow them to avoid or, at least, diminish such effects. It is of interest to determine whether older (as well as younger)

L. L. Jacoby, lljacoby@artsci.wustl.edu

adults can do so. Older adults have been held to be deficient in inhibitory processes (e.g., Hasher et al., 1999), less likely to engage in deeper processing of the sort necessary for memory (e.g., Craik & Byrd, 1982), and less able to engage in recollection as a means of avoiding interference effects (e.g., Hay & Jacoby, 1999). Deficits of these sorts might result in older adults' being less able to benefit from prior experience with PI than are young adults.

Investigation of the effects of prior experience with interference on false memory has also been neglected. Although there have been many findings of false memory that originates from interference, it has not been determined whether such effects would persist across extended experience with the procedures used to produce them. We were particularly interested in determining whether older adults could learn to avoid false memories in the form of high-confidence intrusion errors that result from interference effects. Success in decreasing older adults' susceptibility to PI and in reducing false memories resulting from PI would increase understanding of the bases for PI, as well as of age-related differences in memory, and would potentially have import for applied purposes.

To investigate effects of experience, we examined PI across a pair of lists (A–B, A–D for interference items), presented a pair of two new lists (E–F, E–G), and tested again for PI. The effects of multiple experiences dealing with PI were examined by comparing both accuracy and confidence on the second encounter with PI to results from the first encounter with PI.

We employed a procedure that was introduced by Hay and Jacoby (1996). Specifically, the participants were first exposed to a list of semantically related word pairs (e.g., knee–bone), with each pair being presented three times. Following this training phase, the participants studied a list of word pairs. A third of the studied pairs (facilitation items) were identical to those presented during the training phase (e.g., knee–bone). For another third of the items (interference items), the right-hand member of a studied pair differed from that of a pair that had been presented during the training phase (e.g., knee–bend). These interference items were characterized by high levels of PI and are thus of primary interest. The remaining third of the studied pairs (control items) had not been presented during the training phase. Following the study phase, the participants were given a cued recall test and made a confidence judgment for each item. Test items consisted of a cue word and word fragment that could be completed with either the target or the alternate response that appeared during the training for interference items (e.g., knee–b_n_).

The test procedure of presenting a cue word along with a fragment of the response was meant to restrict responses to the target word and the alternate word that would serve to complete the fragment with an associatively related word. This was intended to increase PI and allow us to better examine effects of PI on intrusion errors. Including both facilitation and interference pairs within a list accords with experience outside the laboratory, in that one usually encounters cues requiring responses that are consistent with their prior responses intermixed with those for which prior experience serves as a source of PI. Inter-

mixing facilitation and interference pairs was expected to increase PI. Also, including facilitation items discourages a strategy of generating both responses to a cue and then producing the less familiar one as a means of avoiding interference effects. Adopting that strategy might enhance performance on interference pairs but would reduce accuracy for facilitation pairs.

In order to improve accuracy and reduce high-confidence intrusions, half of the participants were provided with feedback following each test trial (Experiment 1; cf. Rebok & Balcerak, 1989). The feedback was structured in a manner that depended on confidence judgments. Specifically, participants were instructed to treat each confidence judgment as a wager that they had correctly recalled the studied item. A running score was made visible on all trials. Participants were told that, for each correct answer reported, their point total would increase by a number equivalent to their confidence rating, whereas each incorrect answer would decrease their score by that same number. We expected that providing feedback in this way would serve to diminish the effects of interference and likewise reduce reported confidence when intrusions were produced. Participants were administered two rounds of the procedure, with different materials being used in each round, allowing a comparison of performance for the first versus second round of testing. Another group received the same procedure, with the exception that they did not receive feedback. Participants in the no-feedback condition were simply instructed to make confidence judgments without being told to treat those judgments as a wager. We expected the presence of feedback to be important, particularly for older adults. Older adults were expected to be more likely to produce high-confidence intrusion errors than were young adults and, therefore, to be more reliant on feedback to reveal their errors.

As was anticipated, the results of Experiment 1 revealed that prior experience with PI did indeed diminish its effects for both older and young adults. Having had prior experience with PI, older adults were dramatically less likely to produce false memories in the form of high-confidence intrusion errors.

We conducted Experiment 2 in order to gain insight into the means by which experience with PI allowed participants to diminish its effects.

EXPERIMENT 1

Method

Participants. Forty-eight young and 48 older adults were recruited from the Washington University Psychology Department subject pool. Participants in each age group were randomly assigned to the two feedback conditions (24 participants each). The mean age did not differ significantly between the feedback and no-feedback conditions for either older (74.58 vs. 75.58 years) or young adults (19.92 vs. 19.13 years). Participants were tested individually and received course credit (young adults only) or $10 per hour.

Design and Materials. A 2 (feedback: feedback vs. no-feedback) × 3 (item type: facilitation vs. control vs. interference) × 2 (round: 1 vs. 2) × age (young vs. older) mixed design was used. Item type and round were manipulated within subjects; feedback and age were between-subjects variables.

Materials consisted of 156 three-word sets. Each set contained a cue word (e.g., knee) and two related responses (e.g., bone, bend). Groups were created such that, across items, the average length and strength of association of each response with the cue was equated. Both responses could complete the same word fragment (e.g., b_n_). Six groups of 24 sets served as critical items; the remaining sets served as buffers. Each group was balanced for frequency and length of cues and responses. The groups were counterbalanced across conditions, so that each occurred equally often in each of the within-subjects conditions.

Item types were created by varying the relationship of the pairs in the training (List 1) and study (List 2) lists. For facilitation items, the cues and responses were the same in Lists 1 and 2. For the interference items, the cues were the same in both lists, but the responses changed from List 1 to List 2 (e.g., knee–bone, knee–bend). Control items were only presented in the study list. Test items consisted of a cue word and word fragment that could be completed with either the target or an alternate response that appeared during training for interference items (e.g., knee–b_n_).

On each round, List 1 consisted of 48 pairs presented three times each (144 total), and List 2 consisted of 24 items from each of the three item types (72 total). Both lists were presented in random order. An additional three pairs were presented at the beginning and end of the study list to serve as primacy and recency buffers. Test lists contained 24 items of each type (72 total) and were presented in a fixed random order with the restriction that no more than three items from the same condition were presented consecutively.

Procedure. The training phase (List 1) occurred first. The participants were told that they would see word pairs that would be repeatedly presented and that they were to read each pair aloud. Pairs were presented for 2 sec each, followed by a 1,000-msec interstimulus interval (ISI). After List 1 was completed, participants began the study phase (List 2). They were told that they would see word pairs that they would need to remember for a later memory test and were instructed to read each pair aloud. Each pair was presented for 2 sec, followed by a 1,000-msec ISI.

At the time of test, participants were told that a cue paired with a word fragment would be presented (e.g., knee–b_n_). They were instructed to complete the fragment with the word that was paired with the cue word in the study phase. Participants were given 10 sec to provide a response aloud. Next, they were asked to provide a confidence rating on a scale from 1 (*not very confident*) to 5 (*very confident*) that the item had been studied in List 2. Participants were encouraged to use the full range of the scale. Four practice trials were given prior to test. All responses were recorded by the experimenter.

For participants in the no-feedback condition, a new test item was presented, following each confidence rating. In contrast, participants in the feedback condition were given corrective feedback following each item in both rounds of the experiment. Participants given feedback were told to treat their confidence rating as a wager based on the likelihood that their answer was correct. Participants gained or lost points equal to their confidence rating for correct and incorrect responses, respectively. A message regarding the accuracy of each response was displayed following the confidence rating, and points were added or deducted from a running tally. Participants were encouraged to earn as many points as possible. After the first round, participants in the feedback condition were shown their cumulative score for that round. Both groups completed the entire procedure a second time with a new set of items.

Results and Discussion

Preliminary inspection of the results revealed that the probability of correctly responding to items in the facilitation condition was near ceiling, so we analyzed the results from that condition separately from performance in other conditions. Our primary interest was in whether prior experience with PI would diminish its effects, particularly for false memory (intrusion errors at the highest level of confidence).

Diminishing effects of PI were evidenced by a significant interaction between rounds (first vs. second encounter with PI) and item type (control vs. interference) in the probability of intrusion errors. Having found an interaction of that sort, we analyzed effects separately for control and interference items. The probability of producing either the target or its alternate was extremely high for facilitation (.97), control (.92), and interference items (.97), so errors other than producing an alternate response as an intrusion error did not enter into the analyses. We do not report significant main effects of variables when an interaction involving the variables was significant. The significance level for all tests was set at $p < .05$.

An analysis of the number of correct responses in the facilitation condition revealed only a significant effect of age [$F(1,92) = 12.85, MS_e = .17, \eta_p^2 = .12$]. The probability of producing a correct response for facilitation pairs was higher for young adults (.93) than for older adults (.87). An additional analysis of performance in the facilitation condition examined effects on the joint probability of producing a correct response and reporting the highest level of confidence (5) in its accuracy. Analysis of those results (Table 1) revealed only a significant effect of feedback [$F(1,92) = 6.14, MS_e = .41, \eta_p^2 = .06$]. Providing feedback increased highest confidence, accurate responding for both young and older adults. Neither the main effect of round nor any other main effect or interaction approached significance. The lack of a significant effect of round on responding to facilitation items is important in that it suggests that participants did not reduce their attention to the training list in the second round as a means of reducing PI. Had they done so, one would expect to find a reduction in high-confidence, accurate responding in the facilitation condition across rounds. More important, the results reveal the importance of feedback for increasing confidence in correct responses. As will be seen, feedback was also important for reducing confidence in erroneous responses.

Prior experience with PI diminished its effects on intrusion errors, as evidenced by a highly significant interaction of item type (control vs. interference) and round [$F(1,92) = 9.46, MS_e = .09, \eta_p^2 = .10$]. The results from a separate analysis of performance in the control condition

Table 1
Probability of Correct Recall (PCR) Held in Highest Confidence for Facilitation Items As a Function of Round, Feedback, and Age

	Young				Older			
	Feedback		No Feedback		Feedback		No Feedback	
	PCR	SEM	PCR	SEM	PCR	SEM	PCR	SEM
				Experiment 1				
Round 1	.67	.04	.62	.04	.72	.04	.62	.04
Round 2	.69	.04	.55	.04	.71	.04	.62	.04
				Experiment 2				
Round 1	.55	.05	–		.61	.05	–	
Round 2	.58	.06	–		.57	.06	–	

Table 2
Probability of List 1 Intrusions (PLI) on Interference Items
As a Function of Age, Round, and Feedback

| | Young | | | | Older | | | |
| | Feedback | | No Feedback | | Feedback | | No Feedback | |
	PLI	SEM	PLI	SEM	PLI	SEM	PLI	SEM
	Experiment 1							
Round 1	.33	.03	.34	.02	.54	.03	.55	.03
Round 2	.27	.03	.26	.03	.43	.03	.51	.03
	Experiment 2							
Round 1	.29	.04	–		.39	.04	–	
Round 2	.20	.03	–		.27	.03	–	

revealed only a significant interaction of age and feedback [$F(1,92) = 7.10$, $MS_e = .08$, $\eta_p^2 = .07$]. This interaction was produced by the age difference in the probability of an intrusion error being smaller in the feedback condition (.15 and .20 for young vs. older adults, respectively) than in the no-feedback condition (.12 and .25 for young vs. older adults, respectively). Neither the effect of round nor any interaction with round approached significance in the analysis of performance on control items.

Our primary interest was in intrusion errors in the interference condition (Table 2). Older adults produced more intrusion errors on interference items (.51) than did young adults (.30) [$F(1,92) = 60.88$, $MS_e = 2.09$, $\eta_p^2 = .40$]. In contrast to results for the control condition, the effect of round was highly significant for the interference condition [$F(1,92) = 20.86$, $MS_e = .24$, $\eta_p^2 = .19$], with the probability of an intrusion error being lower in the second round (.37) than in the first round (.44). Inspection of the results in Table 2 suggests that, for older adults, the effect of prior experience with PI did more to diminish its effects in the feedback condition than in the no-feedback condition. However, the relevant interaction did not approach significance.

The effects of prior experience on false-memory errors were of particular interest because people are likely to act on such errors. False-memory errors are measured as the joint probability of producing an intrusion error and holding the highest level of confidence (5) in the accuracy of the erroneous response. For purposes of comparison, errors held at the highest level of confidence were analyzed for control pairs. Responding to the cue from a control pair with the alternate to the target (the response that, for an interference pair, was presented in List 1 and served as a source of PI) was counted as an intrusion error. The interaction of type of pair (control vs. interference) with round was highly significant [$F(1,92) = 24.35$, $MS_e = .10$, $\eta_p^2 = .21$].

A separate analysis of performance on control pairs revealed that high-confidence intrusions were rare because the alternate response had not been presented in List 1 but occurred with a higher probability for older adults (.056) than for young adults (.012) [$F(1,92) = 20.94$, $MS_e = .09$, $\eta_p^2 = .19$]. The reduction from Round 1 (.038) to Round 2 (.030) in the probability of a highest confidence intrusion error to control pairs was exceedingly small but approached significance [$F(1,92) = 3.36$, $MS_e = .003$, $p =$

.07, $\eta_p^2 = .04$]. The effect of prior experience on false-memory errors for interference pairs was much larger.

An analysis of false-memory errors for interference pairs (Table 3) revealed that older adults were much more likely to make such errors than were young adults [$F(1,92) = 95.15$, $MS_e = 2.46$, $\eta_p^2 = .51$] and also revealed a large decrease in the probability of false-memory errors across rounds [$F(1,92) = 32.47$, $MS_e = .24$, $\eta_p^2 = .26$]. The triple interaction of round, feedback, and age approached significance [$F(1,92) = 3.76$, $MS_e = .03$, $p = .056$, $\eta_p^2 = .04$]. Further analyses examined results separately for young and older adults. The analysis of results for young adults showed that they were more likely to produce false-memory errors when feedback was provided [$F(1,46) = 7.92$, $MS_e = .07$, $\eta_p^2 = .15$]. More important, the round × feedback interaction was significant for older adults [$F(1,46) = 4.88$, $MS_e = .06$, $\eta_p^2 = .10$] but did not approach significance for young adults ($F < 1$). As is shown in Table 3, for older adults, the decrease in false-memory errors across rounds occurred only in the condition that received feedback, whereas the provision of feedback was unimportant for the effect of round on the performance of young adults.

The decrease in the probability of high-confidence intrusion errors across rounds might reflect an overall improvement in the ability to use confidence judgments to discriminate between correct responses and errors. Additional analyses examined the extent to which confidence judgments differed between correct recalls and errors, separately for control and interference items. For control items, young adults were more confident in their correct recall of target items (4.05) than in their errors (2.40), as were older adults (4.14 vs. 3.09). However, the difference in confidence between correct recalls and errors was greater for young adults than for older adults [$F(1,86) = 17.94$, $MS_e = 8.29$, $\eta_p^2 = .17$], largely because of older adults' greater confidence in their errors. Confidence in responses was higher in the first round (3.49) than in the second round (3.34) [$F(1,86) = 7.62$, $MS_e = 2.04$, $\eta_p^2 = .08$], but the interaction of round with confidence in correct recalls versus errors did not approach significance ($F < 1$). That is, the effects of prior experience did not increase the usefulness of confidence judgments for discriminating between correct recall and errors on control items.

Table 3
Probability of List 1 Intrusions (PLI) Held at the Highest Level
of Confidence (False Memories) As a Function
of Age, Round, and Feedback

| | Young | | | | Older | | | |
| | Feedback | | No Feedback | | Feedback | | No Feedback | |
	PLI	SEM	PLI	SEM	PLI	SEM	PLI	SEM
	Experiment 1							
Round 1	.15	.03	.10	.03	.36	.03	.35	.03
Round 2	.09	.03	.03	.03	.24	.03	.32	.03
	Experiment 2							
Round 1	.08	.03	–		.20	.03	–	
Round 2	.03	.02	–		.11	.02	–	

Table 4
Confidence Ratings for Correct Recalls and Intrusion Errors on Interference Items As a Function of Age, Round, and Feedback

	Feedback								No Feedback							
	Young				Older				Young				Older			
	Correct		Error		Correct		Error		Correct		Error		Correct		Error	
	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM
Experiment 1																
Round 1	4.41	0.10	3.78	0.13	4.31	0.09	4.27	0.12	4.41	0.10	3.52	0.12	4.30	0.09	4.32	0.12
Round 2	4.34	0.15	3.14	0.16	4.30	0.13	3.84	0.14	4.22	0.14	3.40	0.15	4.15	0.13	4.20	0.14
Experiment 2																
Round 1	.84	.03	.69	.04	.82	.03	.74	.04	–		–		–		–	
Round 2	.86	.03	.54	.05	.81	.03	.67	.04	–		–		–		–	

Not surprising, confidence judgments for interference items (Table 4) discriminated less well between correct recalls (4.29) and intrusion errors (3.82) than did confidence judgments for control items (4.08 vs. 2.76) [$F(1,82) = 126.58$, $MS_e = 31.20$, $\eta_p^2 = .61$]. As was found for control items, older adults' confidence judgments for interference items discriminated less well between correct recalls and errors than did those of young adults [$F(1,86) = 53.15$, $MS_e = 13.58$, $\eta_p^2 = .38$]. More important, for interference items, prior experience with PI improved the usefulness of confidence judgments for distinguishing between correct recalls and errors, but only when feedback was given. The interaction of confidence in correct recall vs. intrusion errors with round and feedback for performance on interference items was significant [$F(1,86) = 6.81$, $MS_e = 1.66$, $\eta_p^2 = .07$]. An analysis of performance in the feedback condition revealed that the difference between confidence in correct recalls and intrusion errors was greater in the second round than in the first round, due to a reduction in confidence for intrusion errors [$F(1,42) = 11.24$, $MS_e = 2.69$, $\eta_p^2 = .21$]. For the no-feedback condition, in contrast, the corresponding interaction did not approach significance ($F < 1$). Rather, round only had the effect of equally reducing confidence in correct recalls and intrusion errors [$F(1,44) = 4.27$, $MS_e = .92$, $\eta_p^2 = .09$].

In sum, the results revealed that prior experience with PI diminished its effects for both young and older adults. For older adults, there was a tendency toward the reduction in intrusion errors being larger when feedback was provided. The probability of false-memory intrusion errors held with the highest level of confidence was greater for older than for young adults and was reduced by prior experience with PI. However, for older adults, this reduction in false memory occurred only when feedback was given. For both young and older adults, prior experience with PI increased the ability of confidence judgments to distinguish between correct recalls and intrusion errors in the interference condition, but only when feedback was given.

It might be argued that the reduced PI on the second round occurred because participants paid less attention to the list that served as a source of interference on the second round as compared with on the first round. Against that possibility, reduced attention to the first list would be expected to reduce performance on facilitation items as well as reducing interference. An effect of round was not found for facilitation items, even when performance was examined for correct responses given with highest confidence, which were far from a ceiling level of performance. Furthermore, for older adults, a reduction in intrusion errors produced at the highest level of confidence occurred only when feedback was given. Also, the effects of prior experience with PI on the ability of confidence judgments to distinguish between correct recalls and intrusions was restricted to interference items and occurred only when feedback was given. The specificity of these effects argues against the possibility that the diminishment in PI resulted from a general reduction in attention to the list that served as the source of PI. Experiment 2 provides evidence that allows further specification of how experience with PI serves to diminish its effects.

EXPERIMENT 2

In Experiment 2, we sought to replicate the finding that prior experience with PI diminishes its effects. Given the importance of feedback found in Experiment 1, all of the participants received feedback in Experiment 2. In contrast to Experiment 1's design, the participants in Experiment 2 were informed about the makeup of the study list, being told that the list would include interference and facilitation pairs as well as control pairs. Despite their having been warned about the presence of interference pairs, we did not expect participants to be aware of the greater difficulty of such pairs until Round 2, after prior experience with PI.

Having become aware of the difficulties that interference pairs produce for later memory performance as a result of a prior encounter with PI, participants might pay special attention to interference pairs during study in the second round, so as to diminish the effects of PI. We examined this possibility by employing a self-allocated study-time procedure. Prior research has shown that learners generally allocate more study time to difficult items than to easy items when there are no time constraints on the study episode (for a review, see Son & Metcalfe, 2000). Item difficulty has typically been manipulated by varying the strength of association between a cue and its response, whereas our interest was in differences in difficulty produced by interference effects. Study time served as an index of the extent to which participants monitored item difficulty across item types (see, e.g., Koriat, Ma'ayan, &

Nussinson, 2006). We expected that study time would not differ for the interference and control conditions in the first round. However, in the second round, as a result of feedback in the first round, participants were expected to monitor their study in ways that revealed interference pairs as being more difficult than control pairs and, therefore, to devote more study time to interference pairs. Greater attention to interference pairs during study might serve to better support later memory performance in ways that diminish the effects of PI.

Experiment 2 was also designed to determine whether effects of prior experience with PI on confidence judgments would affect participants' ability to withhold intrusion errors by not responding when given the opportunity to do so. After being forced to respond to each test item, participants gave a confidence judgment and were then given the option to volunteer or withhold their response (cf. Koriat & Goldsmith, 1996). This option provided an opportunity for participants to withhold incorrect responses to interference items and further diminish, or even eliminate, PI effects. For volunteered responses, participants were awarded 5 points for correct responses and penalized 15 points for incorrect responses. For responses that were withheld, no feedback was given, and points were neither awarded nor penalized. We expected that the option to withhold responses would allow participants to further diminish the effects of PI, particularly in Round 2.

Method

Participants. Participants were 24 young adults (mean age = 20.25 years) and 24 older adults (mean age = 74.50 years) recruited from the Washington University Psychology Department subject pool. Participants were tested individually and received course credit (young adults only) or $10 per hour.

Design and Materials. The design and materials in Experiment 2 were identical to those in Experiment 1, but with the exceptions described below. A 2 (age: young vs. older) × 3 (pair type: facilitation vs. control vs. interference) × 2 (round: 1 vs. 2) mixed design was used. Age was a between-subjects factor, and pair type and round were manipulated within subjects. The number of pairs was reduced to shorten the length of the experiment in order to accommodate the procedural changes. Materials consisted of 132 word triples. Six sets of 20 triples were rotated through conditions, with the remaining triples being used as buffers. Training lists comprised 20 pairs of each type. Buffers were used for practice tests that preceded the actual tests in each round.

Procedure. The procedure was the same as in Experiment 1, but with the following exceptions. The training phase (List 1) in each round included fewer pairs (40 critical, 4 buffers = 132 total presentations). Prior to the study phase, the participants were informed about the nature of the study pairs in relation to those presented in the training phase. They were told to study each pair until it had been learned completely and to click on a box labeled "Next" displayed below each pair in order to move on, once they had finished studying. All ISIs were set to 500 msec.

At the time of test, the participants were given cue–fragment pairs in the same manner as in Experiment 1 (e.g., knee–b_n_) and were instructed to complete the fragment with the response presented in the study phase. The participants were then asked to rate their confidence, on a scale from 0 (*wild guess*) to 100 (*certain correct*), that their response matched what they had studied. The participants were encouraged to use the full range of the scale.

Following their confidence judgments, the participants were given the option to report or withhold their responses. For reported responses, they were awarded 5 points for each correct response and penalized 15 points for each incorrect response. No points were gained or lost for withheld responses. Participants were encouraged to maximize their point total by reporting only responses for which they were sufficiently confident of being correct. A running score was displayed in the upper right-hand corner of the screen. Corrective feedback was not provided for withheld responses.

Results and Discussion

We begin by reporting the results that replicate those found in Experiment 1. Next, we examine the effects of prior experience with PI on the ability to withhold intrusion errors so as to increase the accuracy of responding. Finally, we report effects on the allocation of study time to show that participants increase the amount of study time devoted to interference pairs following prior experience with PI. As in Experiment 1, the probability of producing either the target or its alternate was extremely high for facilitation (.98), control (.96), and interference items (.97), so errors other than producing an alternate response did not enter into the analyses.

Analysis of performance in the facilitation condition did not yield any significant effects. Performance was near ceiling for both young (.93) and older adults (.91). An additional analysis of performance in the facilitation condition examined effects on the joint probability of producing a correct response and reporting the highest level of confidence in its accuracy (Table 1, bottom rows). That analysis did not reveal any significant effects. As in Experiment 1, round did not influence responding on facilitation items, suggesting that attention to List 1 was not reduced across rounds. This is important for dismissing the possibility that diminished effects of PI across rounds result from reduced attention to List 1, the source of PI. A reduction in attention of that sort would be expected to reduce performance on facilitation pairs, as well as to reduce interference for interference pairs.

More important, the analysis of intrusion errors in the control and interference conditions revealed a significant interaction between item type (control vs. interference) and round [$F(1,46) = 8.17$, $MS_e = .08$, $\eta_p^2 = .15$], showing that prior experience with PI diminished its effects. For items in the control condition, only the effect of age was significant [$F(1,46) = 5.49$, $MS_e = .06$, $\eta_p^2 = .11$]. The probability of an intrusion error on control items was higher for older adults (.13) than for young adults (.08). Neither the effect of round nor any interaction with round approached significance.

Performance on interference items was of greater interest (Table 2, bottom row). Fewer intrusion errors were produced for interference items in the second round than in the first [$F(1,46) = 20.04$, $MS_e = .28$, $\eta_p^2 = .30$]. Also, young adults produced fewer intrusion errors than did older adults [$F(1,46) = 4.78$, $MS_e = .19$, $\eta_p^2 = .09$].

As in Experiment 1, our primary interest was in whether experience with PI would reduce the probability of false memory, defined as an intrusion error accompanied by 100% confidence. False memory was significantly reduced by prior experience, as evidenced by a highly significant interaction between round and item

type (control vs. interference) [$F(1,46) = 14.75$, $MS_e = .04$, $\eta_p^2 = .24$]. An analysis of performance on control items revealed that the probability of highest confidence errors was greater for older adults (.029) than for young adults (.001) [$F(1,46) = 10.54$, $MS_e = .02$, $\eta_p^2 = .19$]. The effect of round was also significant, although the probability of highest confidence errors for control pairs differed little between the first and second rounds (.022 vs. .008) [$F(1,46) = 5.59$, $MS_e = .004$, $\eta_p^2 = .11$]. As is shown in the bottom row of Table 3, the probability of false memory for interference pairs was much higher for older (.15) than for young adults (.06) [$F(1,46) = 10.34$, $MS_e = .24$, $\eta_p^2 = .18$]. Also, for interference pairs, the probability of false memory was greatly reduced on the second round (.07), as compared with the first round (.14) for both young and older adults [$F(1,46) = 28.34$, $MS_e = .13$, $\eta_p^2 = .38$].

A further analysis of confidence judgments examined their ability to distinguish between correct recall of target items and intrusion errors. An analysis of confidence judgments for responses to control items revealed only that participants' confidence in their correct recalls (.72) was higher than that in their intrusion errors (.42) [$F(1,25) = 61.29$, $MS_e = 2.42$, $\eta_p^2 = .71$]. Neither the effect of age nor that of round approached significance ($Fs < 1$). In contrast, an analysis of confidence judgments for responses to interference items (Table 4, bottom row) showed that experience with PI increased the ability of confidence judgments to distinguish between correct recalls and intrusion errors for both young and older adults, as evidenced by a significant interaction between response (correct recall vs. intrusion error) and round [$F(1,43) = 9.59$, $MS_e = .13$, $\eta_p^2 = .18$]. However, older adults' confidence judgments for interference items discriminated less well between correct recall and intrusion errors than did those of young adults [$F(1,43) = 7.11$, $MS_e = .19$, $\eta_p^2 = .14$], largely because of the older adults' higher confidence in their intrusion errors.

When given the opportunity to withhold responses, the probability of a response being withheld, regardless of whether or not it was correct, was lower for facilitation items (.11) than for either control (.22) or interference items (.16) [$F(2,92) = 30.02$, $MS_e = .28$, $\eta_p^2 = .40$]. An analysis that included only control and interference items revealed a significant interaction between age and item type [$F(1,46) = 4.93$, $MS_e = .05$, $\eta_p^2 = .10$]. Young adults differed little from older adults in their probability of withholding a response to an interference item (.16 vs. .17) but were much less likely to withhold a response to a control item (.18) than were older adults (.25).

Did prior experience with PI increase the probability of withholding intrusion errors to interference items? It did so for young adults, but not for older adults, as evidenced by a significant interaction of age and round [$F(1,43) = 5.93$, $MS_e = .35$, $\eta_p^2 = .12$]. For young adults, the probability of withholding an intrusion error for interference items was higher on the second round (.46) than on the first round (.22). In contrast, the probability of older adults withholding an intrusion error to interference items was identical on the two rounds (.21).

As compared with forced report, allowing participants to withhold responses (free report) increased the accuracy of recall in the control and interference conditions, and there was a tendency toward this increase being larger for young adults than for older adults [$F(1,46) = 3.64$, $MS_e = .01$, $p = .063$, $\eta_p^2 = .07$]. Across conditions, the probability of an error was higher for young adults when recall was forced rather than free (.16 vs. .12), whereas the corresponding difference for older adults was somewhat smaller (.23 vs. .21). The pattern was the same when the results were analyzed separately for interference items. The probability of an intrusion error for young adults was higher when recall was forced rather than when it was free (.24 vs. .20), and there was a tendency toward that difference being smaller for older adults (.33 vs. .32), but the interaction only approached significance [$F(1,46) = 3.10$, $MS_e = .01$, $p = .09$, $\eta_p^2 = .06$]. No other interactions involving report option, including the interaction with round, approached significance.

As in Experiment 1, prior experience with PI diminished its effects. Examination of participants' allocation of study time is revealing, with regard to the basis for this increased resistance to PI. As is shown in Table 5, participants spent less time studying facilitation pairs than they did either control or interference pairs, and the amount of study time devoted to facilitation pairs decreased across rounds. More important, the interaction of round with item type (control vs. interference) was significant [$F(2,92) = 6.95$, $MS_e = 2,845,578.22$, $\eta_p^2 = .13$]. The significant interaction arose from study time allocated to control items decreasing across rounds, whereas study time allocated to interference items increased across rounds [$F(1,46) = 7.72$, $MS_e = 1,922,400.75$, $\eta_p^2 = .15$]. There was also a tendency for older adults to devote more study time to items in all conditions than did young adults (5,482 vs. 5,081 msec), but, because of the high variability of study time, neither that difference nor any interaction with age approached significance ($Fs < 1$). The differential effects of round suggest that both young and older adults became aware of the greater difficulty of interference items, so they increased the amount of study time devoted to those items in the second round. It is likely that their doing so was at least partially responsible for the diminished effects of PI.

Table 5
Study-Time Allocation (in Milliseconds) As a Function of Round, Item Type, and Age Group in Experiment 2

| | Item Type | | | | | |
| | Facilitation | | Control | | Interference | |
	M	SEM	M	SEM	M	SEM
	Young Adults					
Round 1	4,844	792	5,606	882	5,454	757
Round 2	3,985	402	5,078	706	5,522	631
	Older Adults					
Round 1	4,939	509	5,543	588	5,796	627
Round 2	4,824	550	5,667	638	6,124	632

Note—Observations exceeding 2.5 *SD*s above or below the mean in each within-participants condition were trimmed prior to analysis. Less than 2% of all observations were excluded.

To measure the relation between self-allocated study time and accuracy of responding, gamma correlations between study time and accuracy were computed for each participant (see Nelson, 1984) and then analyzed by means of an ANOVA. The analysis of those correlations revealed only a marginally significant effect of item type [$F(2,38) = 2.78$, $MS_e = .43$, $p = .07$, $\eta_p^2 = .13$]. There was an inverse correlation between study time and accuracy for facilitation pairs ($-.14$), whereas the correlation was not significant and slightly positive for control ($.05$) and interference ($.04$) pairs. The overall correlation was near zero. The inverse correlation for facilitation pairs suggests that participants were able to identify pairs that would be easily recalled and devoted little study time to those pairs.

The generally low correlation between study time and accuracy might be surprising. Instead, one might expect a positive correlation between study time and accuracy because experimenter-controlled increases in study time typically increase accuracy. However, the lack of correlation is not surprising if one realizes that, ideally, the amount of study time devoted to an item should depend on its difficulty, with study time devoted to an item's being greater the higher its judged difficulty for recall. Allocating study time in this way would reduce the correlation between study time and accuracy and produce a zero correlation if there were differences in item difficulty and participants were sufficiently able to recognize and overcome those item differences by means of their allocation of study time (Koriat et al., 2006; cf. Nelson & Leonesio, 1988). Due to the complexities created by item differences, along with the possibility of qualitative changes in study, it would be difficult to detect changes in the relation between study time and accuracy across rounds.

GENERAL DISCUSSION

Experience with PI diminishes its effects. Both experiments revealed a significant interaction with round, such that intrusion errors to interference items decreased across rounds, whereas performance on control items remained unchanged. Older adults were generally more reliant on feedback than were young adults, which is not surprising, since older adults were much more prone to produce high-confidence intrusion errors. There was a trend toward the reduction in intrusion errors for older adults' being larger in Experiment 1 when feedback was provided, and older adults' reduction in false memory—defined as intrusion errors held at the level of highest confidence—was observed only when feedback was provided.

For both young and older adults, the increase across rounds in the extent to which confidence judgments were diagnostic of correct responding occurred only when feedback was provided. Feedback was provided for all conditions in Experiment 2, and results from that experiment replicated those of Experiment 1 by showing a decrease in false memory as a result of prior experience with PI, along with an increase in the extent to which confidence judgments were diagnostic of correct responding. The

decrease in false memory gained from prior experience with PI was substantial. For older adults, the probability of false memory was almost halved from the first round to the second round (.20 vs. .11) in Experiment 2.

When given the option to withhold responses in Experiment 2, young adults were more likely to withhold false recalls on the second round than on the first round; older adults were not. Neither young nor older adults substantially increased the accuracy of their responding under conditions of free, as compared with forced, responding. In contrast, Kelley and Sahakyan (2003) found that both young and older adults were able to greatly increase their accuracy under conditions of free responding. There are a number of differences between our experiments and theirs. Most important, perhaps, are the payoffs for correct and false recall that were employed. We awarded 5 points for each correct response that was volunteered, and we penalized by subtracting 15 points for each error that was volunteered, whereas Kelley and Sahakyan awarded 25 cents for each correct response and penalized $2.50 for each incorrect response. Perhaps our awards and penalties were simply not extreme enough to result in participants' imposing a criterion for volunteering responses that would substantially increase their accuracy of responding. In this vein, confidence judgments, particularly for young adults, appeared to be sufficiently diagnostic of correct responding on the second round to allow them to greatly increase their accuracy by withholding responses if they had adopted a more stringent criterion for responding.

We believe that our experiments are the first to show that experience with PI can diminish its effects—including false memory. We find it striking that minimal training can have a substantial impact. Much of the work on memory accuracy and aging has primarily drawn conclusions from a single session of performance. The data from the present study suggest that, if older adults are given feedback (cf. Balota, Duchek, Sergent-Marshall, & Roediger, 2006; Rebok & Balcerak, 1989) and another opportunity to be tested, they can demonstrate improvements in performance. The implications of this are important in showing that older adults' accuracy and confidence are both open to remediation and in providing further evidence that older adults' susceptibility to false memories can be ameliorated (cf. McCabe & Smith, 2002). It is important that the probability of false memory that originates from PI can be reduced, because PI is a common source of false memories.

How did experience with PI reduce false memory? One might suggest that reductions in confidence for intrusion errors to interference items occurred because participants became more risk averse as a consequence of high-confidence errors. A risk-aversion hypothesis would posit a generalized reduction in confidence, with participants reducing their confidence for all items across rounds. However, confidence for correct responses to facilitation pairs did not decrease across rounds. Also, the extent to which differences in confidence judgments were diagnostic of the accuracy of responses did not change across rounds for control items. Rather, those effects of prior

experience with PI were selective to interference items. Thus, a risk-aversion hypothesis would have to rather implausibly assume that participants became more conservative on only a subset of the items (i.e., interference items) presented randomly throughout the test. Furthermore, the risk-aversion hypothesis would not account for the overall reduction in intrusion errors across rounds that was observed when responding was forced, as it was in Experiment 1 and in the forced-recall condition of Experiment 2. Yet another account would hold that the reduction of PI across rounds resulted from reduced attention to the training list, which served as a source of PI, on the second round. As discussed in conjunction with Experiment 1, the specificity of effects of prior experience on PI weighs against that account.

Effects of prior experience with PI on the self-allocation of study time observed in Experiment 2 suggest that prior experience with PI resulted in a change in the encoding of interference items across rounds. Although informed of the relationship between lists, participants were apparently unaware of the greater difficulty of interference items until after they had experienced PI.

In a similar vein, Benjamin (2003) showed that participants gave higher judgments of learning for recognition of high- than of low-frequency words in an initial study session, predicting a pattern of performance that was opposite to that observed. However, after experience with a recognition test in the first session, participants correctly predicted that their recognition-memory performance would be better for low-frequency words in a second session.

For PI, awareness of the greater difficulty of interference items resulted in the amount of study time devoted to interference pairs increasing across rounds, with the result that substantially more study time was devoted to interference pairs than to control pairs on the second round. Furthermore, there may have been both qualitative and quantitative differences in encoding processes across rounds that were selective to interference pairs. As a result of prior experience with PI, participants may have adapted their encoding of interference pairs in ways that provided greater support for later recollection. Similarly, participants in Experiment 1 may have devoted more attention to interference pairs in the second round and adapted their encoding processes to better deal with PI, although they were unable to devote more time to studying those pairs.

Participants' responding on the first round may have relied heavily on a fluency heuristic (Jacoby et al., 2005; Jacoby & Dallas, 1981), reporting items that came to mind most easily. This would be reasonable for facilitation items, but it would also have the consequence of leading participants to report the most accessible and, therefore, incorrect answer for interference items. However, having gained experience with PI, participants may have modified their encoding of interference pairs in ways that made them better able to shift from heavy reliance on fluency to other, more diagnostic bases for memory, such as reinstating prior encoding context to recollect prior details from study. Thus, the basis for participants' confidence in their responses may have undergone a qualitative shift from a heavy reliance on judged fluency in the first round to greater reliance on recollection in the second round, resulting in increased accuracy and a reduction in confidence for errors. That is, the diminished effects of PI can be seen as reflecting a qualitative shift toward greater reliance on recollection as a basis for memory and confidence.

It is noteworthy that older adults were as able as young adults to diminish effects of PI as a result of prior experience with PI. Experiments by Hay and Jacoby (1999) and by Jacoby, Debner, and Hay (2001) used a process-dissociation procedure to show that older adults' greater vulnerability to PI resulted from their lessened ability to recollect the occurrence of particular events. Hay and Jacoby (1999) showed that, given supportive conditions, older adults were able to benefit from distinctive contextual information as a means of enhancing recollection. Results from the present experiments suggest that older adults' ability to recollect can also be enhanced by means of training aimed at diminishing the effects of PI. The possibility that older adults' ability to recollect particular events was improved by training to diminish the effects of PI is consistent with results from other experiments (e.g., Jennings & Jacoby, 2003; Jennings, Webster, Kleykamp, & Dagenbach, 2005) that have enhanced older adults' ability to recollect by training under conditions of high interference. In general, training under conditions of high interference holds promise as a means of improving ability to recollect, thereby reducing false memory.

Given the results of our experiments, it may be useful to revisit procedures that have been shown to produce dramatic false remembering (for reviews, see Loftus, 2004; Roediger, 1996). Many of the studies on false memory have drawn conclusions from a single session of performance. The dramatic levels of false memory found in those studies might not have persisted across repeated applications of the procedures used to produce them. As in the case of PI, the effectiveness of the procedure might diminish as a function of prior experience. For example, it might be more difficult to mislead people to falsely remember having been lost in a mall if they have had prior experience being misled in a similar way (cf. Loftus, 1997).

The results of our present experiments suggest that being forewarned of a manipulation that could produce false memory may not be sufficient to avoid its effects. Being told that the study list included interference pairs did not diminish the effects of PI, but prior experience with PI did. More generally, prior experience in situations that produce false memory might diminish the likelihood of false memory in similar situations. Being forewarned and experienced might allow one to be forearmed against such effects.

AUTHOR NOTE

REFERENCES

ANDERSON, M. C., & NEELY, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Handbook of perception and cognition: Memory* (2nd ed., pp. 237-313). San Diego: Academic Press.

BALOTA, D. A., DUCHEK, J. M., SERGENT-MARSHALL, S. D., & ROEDIGER, H. L., III (2006). Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer's disease. *Psychology & Aging*, **21**, 19-31.

BENJAMIN, A. S. (2003). Predicting and postdicting the effects of word frequency on memory. *Memory & Cognition*, **31**, 297-305.

CRAIK, F. I. M., & BYRD, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 191-211). New York: Plenum.

DODSON, C. S., & KRUEGER, L. E. (2006). I misremember it well: Why older adults are unreliable eyewitnesses. *Psychonomic Bulletin & Review*, **13**, 770-775.

HASHER, L., ZACKS, R. T., & MAY, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance. Interaction of theory and application* (pp. 653-675). Cambridge, MA: MIT Press.

HAY, J. F., & JACOBY, L. L. (1996). Separating habit and recollection: Memory slips, process dissociations, and probability matching. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 1323-1335.

HAY, J. F., & JACOBY, L. L. (1999). Separating habit and recollection in young and older adults: Effects of elaborative processing and distinctiveness. *Psychology & Aging*, **14**, 122-134.

JACOBY, L. L., BISHARA, A. J., HESSELS, S., & TOTH, J. P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General*, **134**, 131-148.

JACOBY, L. L., & DALLAS, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **110**, 306-340.

JACOBY, L. L., DEBNER, J. A., & HAY, J. F. (2001). Proactive interference, accessibility bias, and process dissociations: Valid subjective reports of memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 686-700.

JACOBY, L. L., & RHODES, M. G. (2006). False remembering in the aged. *Current Directions in Psychological Science*, **15**, 49-53.

JENNINGS, J. M., & JACOBY, L. L. (2003). Improving memory in older adults: Training recollection. *Neuropsychological Rehabilitation*, **13**, 417-440.

JENNINGS, J. M., WEBSTER, L. M., KLEYKAMP, B. A., & DAGENBACH, D. (2005). Recollection training and transfer effects in older adults: Successful use of a repetition-lag procedure. *Aging, Neuropsychology, & Cognition*, **12**, 278-298.

KELLEY, C. M., & SAHAKYAN, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory & Language*, **48**, 704-721.

KORIAT, A., & GOLDSMITH, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, **103**, 490-517.

KORIAT, A., MA'AYAN, H., & NUSSINSON, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General*, **135**, 36-69.

LOFTUS, E. F. (1997). Creating false memories. *Scientific American*, **277**, 70-75.

LOFTUS, E. F. (2004). Memories of things unseen. *Current Directions in Psychological Science*, **13**, 145-147.

MAY, C. P., HASHER, L., & KANE, M. J. (1999). The role of interference in memory span. *Memory & Cognition*, **27**, 759-767.

MCCABE, D. P., & SMITH, A. D. (2002). The effect of warnings on false memories in young and older adults. *Memory & Cognition*, **30**, 1065-1077.

NELSON, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, **95**, 109-133.

NELSON, T. O., & LEONESIO, R. J. (1988). Allocation of self-paced study time and the "labor-in-vain effect." *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 676-686.

REBOK, G. W., & BALCERAK, L. J. (1989). Memory self-efficacy and performance differences in young and old adults: The effect of mnemonic training. *Developmental Psychology*, **25**, 714-721.

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

SON, L. K., & METCALFE, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 204-221.