

Experience with proactive interference diminishes its effects: mechanisms of change

Christopher N. Wahlheim · Larry L. Jacoby

Published online: 19 November 2010

© The Psychonomic Society 2010

Abstract In three experiments, we examined the mechanisms by which prior experience with proactive interference (PI) diminished its effects. Cued recall tasks conforming to an A–B, A–D paradigm were used to induce PI effects. Experiment 1 showed that reduced PI was not due to a reduction in attention to the source of PI. Experiment 2 revealed that participants' awareness of PI effects on memory performance increased with experience, resulting in a shift in encoding processes. Experiment 3 demonstrated that changes in encoding provided additional support for recollection that further enhanced participants' ability to constrain their retrieval processing to the appropriate source of information at the time of test. These results can be interpreted as showing that experience with PI enhances awareness of its effects and allows individuals to adjust their learning and retrieval strategies to compensate for such effects.

Keywords Proactive interference · Metacognition · Source-constrained retrieval · Memory training

Proactive interference (PI) is a potent source of forgetting. For serial learning, Underwood (1957) reviewed experiments in which the effects of PI were examined and found that the probability of recall decreased dramatically from .80 when there was no preceding list to approximately .20 following 20 preceding lists. Paired-associate learning also shows large effects of PI (for a review, see Anderson & Neely, 1996). The question we ask in this article is whether participants are unaware of the effects of PI during a first

encounter with PI but become aware of its effects and successfully implement procedures to diminish those effects during a second encounter with PI. Although the effects of PI are widely known by memory researchers, little has been done to investigate people's ability to diminish its effects. Findings of an ability to adapt learning in ways that diminish the effects of PI are important for gaining a better understanding of the effects of PI and, potentially, hold import for applied purposes.

In commonplace situations, people show knowledge of PI effects by taking actions to avoid such effects. For example, in card games, it is common for the dealing of cards to pass around the table, creating the problem of remembering who dealt last so as to decide whose turn it currently is to deal. This problem is typically solved by playing with two decks of cards, with one of the decks being placed in front of the person who is to deal next while the other deck is being dealt. However, rather than being generally aware of effects of PI, such awareness might be restricted to particular situations. For card playing, it is likely that problems produced by PI were discovered as a result of experience in the situation, and only then were steps taken to eliminate its effects. Similarly, combating PI in a list-learning situation might depend on prior experience in the situation of a sort that makes people aware of its effects.

To our knowledge, the only investigation of ability to diminish PI as a result of prior experience with its effects was done by Jacoby, Wahlheim, Rhodes, Daniels, and Rogers (2010). Participants in their experiments were given two rounds of experience with PI in a paired-associate learning task. In each round, two lists were presented, with the relation between the pairs within the lists being varied. To measure PI, performance in an interference condition that presented the same cue but a different response for the two lists (A–B, A–D) was compared with that in a control

C. N. Wahlheim (✉) · L. L. Jacoby
Department of Psychology, Washington University in St. Louis,
St. Louis, MO 63130, USA
e-mail: cnwahlheim@gmail.com

condition for which both the cue and the response (rest, C–D) occurred only in list 2. Other pairs in the second list retained the response from the first list (A–B, A–B). This facilitation condition was meant to encourage attention to list 1 items during both of the two rounds and also to encourage carryover of responses between the two lists, increasing PI. Test items consisted of a cue word and a word fragment that could be completed with either the target or its competing response that appeared in list 1 for interference items. This was intended to further increase PI and to allow effects in the form of intrusion errors to be examined better. The procedure for the second round was the same as that for the first, except that new materials were used (e.g., for interference pairs, E–F, E–G).

The results from the experiments by Jacoby et al. (2010) revealed that PI was diminished on the second, as compared with the first, encounter with PI for both young and older adults. Experiment 2 in that series provided results that suggested that participants became aware of PI only as a result of experience with its effects. A self-allocated study time procedure was employed to examine differences in study time for control and interference pairs. It was found that participants did not spend more time studying interference pairs than control pairs in round 1 but did so in round 2—presumably, as an attempt to diminish the effects of PI that they became aware of in round 1. The probability of intruding a list 1 response as an error when attempting to recall the list 2 response for interference pairs was lower on the second than on the first round. As important, confidence judgments were educated by prior experience, such that the difference in confidence between correct recalls and intrusion errors was larger on the second round than on the first round. The probability of an intrusion error being held at the highest level of confidence decreased across rounds for both young and older adults. This reduction in high-confidence intrusion errors is important because PI is likely to have a deleterious effect in everyday situations only when confidence in the accuracy of erroneous responses is high (e.g., Hay & Jacoby, 1996).

The goals of the experiments reported in the present article were to replicate the results reported by Jacoby et al. (2010) for young adults and to further explore the means by which prior encounters with PI diminish its effects. An anonymous reviewer of the article by Jacoby et al. (2010) suggested that the reduction in PI observed in their experiments might have resulted from reduced attention to list 1 during the second, as compared with the first, round. Against that possibility, performance on facilitation pairs in which both the cue and the response were the same for the two lists (A–B, A–B) did not worsen across rounds, whereas a reduction in attention to list 1 would be expected to produce an effect of that sort. Regardless, Experiment 1

in the present study more directly tested whether memory for list 1 items decreased across rounds. To anticipate, the results of Experiment 1 revealed that prior experience with PI diminished its effects, even though there was no decrease in memory for list 1, the source of PI, across rounds. Experiments 2 and 3 further investigated the means by which prior experience with PI reduced its effects.

Experiment 1

In Experiment 1, we examined the possibility that the reduction in PI as a result of prior experience found by Jacoby et al. (2010) was due to reduced attention to list 1 items in the second round. To more directly test that account, cued recall of a small set of filler items from list 1 was tested immediately following the presentation of list 1 on each of the two rounds. The filler items that were tested did not overlap across rounds, nor did they overlap with the critical items used to manipulate the relationship between items in lists 1 and 2. As in the experiments by Jacoby et al. (2010), the critical items maintained the same response across lists (A–B, A–B), had the same cue but a different response in the two lists (A–B, A–D), or were control items and appeared only in the second list (rest, C–D).

The goal of testing filler items was not to examine the effects of prior testing on PI but, rather, was to examine memory for list 1 items so as to detect any reduction in attention to list 1 items across rounds. Reduction in PI across rounds in the condition that included tests of filler items from list 1 was compared with that in a condition that did not include such tests—a condition that was the same as that used by Jacoby et al. (2010). Given that Jacoby et al. (2010) showed no effects of round on recall of facilitation items (i.e., A–B, A–B pairs), we expected testing of list 1 items to reveal that memory for list 1 would not change across rounds and also expected the reduction in PI across rounds to be the same for tested and nontested conditions. This pattern of results would provide evidence that the reduction in PI across rounds was not due to reduced attention to list 1.

However, testing filler items from list 1 might increase list differentiation for the nontested critical items and, so, produce differences in PI between the tested and nontested conditions. Tulving and Watkins (1974) found that testing memory for paired associates influenced the accessibility of list 1, as compared with list 2, responses in an A–B, A–D paradigm when participants were asked on a final test to recall responses from both lists. When the first list was tested, participants were better able to remember the most recent responses (i.e., $D > B$), but the opposite was true (i.e., $D < B$) when the first list was not tested. Recent investigations of PI (Brewer, Marsh, Meeks, Clark-Foos, &

Hicks, 2010; Szpunar, McDermott, & Roediger, 2008) have shown that testing decreases PI in free recall learning. In contrast to experiments showing testing effects on PI, we only tested filler items, not the critical items that were used to examine the effects of PI. Any reduction in PI produced by testing filler items should be revealed by differences in PI between the tested and nontested conditions.

Method

Participants

Seventy-two Washington University undergraduates participated in exchange for course credit or \$10/hr. Thirty-six participants were randomly assigned to each between-subjects group (list 1 test: present vs. absent). All the participants were tested individually.

Design and materials

A 3 (item type: facilitation vs. control vs. interference) \times 2 (round: first vs. second) \times 2 (list 1 test: present vs. absent) mixed design was used. Item type and round were manipulated within subjects, and presence vs. absence of list 1 test was manipulated between subjects. The materials consisted of 152 three-word sets that included one cue word (e.g., *knee*) and two associatively related responses (e.g., *bone*, *bend*). The bulk of these sets were selected from the norms reported by Jacoby (1996), and the remaining sets were created from the norms reported by D. L. Nelson, McEvoy, and Schreiber (1998). The responses in each set had the same number of letters and could complete the same word fragment (e.g., *b_n_*). Six groups of 20 sets served as critical items, and two groups of 10 sets served as items to be tested in the list 1 test condition (these items were included but not tested in the test-absent condition). The remaining 12 sets served as buffers and practice items. Groups of items were matched on word frequency and length of cue and response words. The groups were rotated through each of the within-subjects conditions, resulting in six formats. Three primacy and three recency buffers served as practice items on each round of testing and remained constant across formats.

In the test-present group, each round consisted of four phases: list 1, list 1 test, list 2, and list 2 test. The test-absent group included the same phases, with the exception that filler items from list 1 were not tested in either round. List 1 consisted of 54 word pairs (e.g., *knee bone*), including 4 pairs used as buffers in list 2, 10 pairs to be tested after list 1, and 40 critical pairs. Each pair appeared three times each, for a total of 162 presentations. List 2 consisted of 66 word pairs (six buffers, 60 critical items).

Three buffers appeared at the beginning and end of the list to prevent primacy and recency effects. Twenty critical pairs represented the facilitation, the control, and the interference conditions. The facilitation pairs were the same as those in list 1 (e.g., *knee bone*, *knee bone*), the control pairs appeared exclusively in list 2 (e.g., *lamb wool*), and the interference pairs consisted of cues from list 1 paired with new responses that appeared only in list 2 (e.g., *apple core*, *apple worm*). Each list 1 test consisted of 10 cues paired with word fragments, and each list 2 test consisted of 66 pairs arranged in the same manner. The fragments (e.g., *apple_or_*) could be completed with either the target word or its competitor from an interference pair (e.g., *core*, *worm*). Cue–fragment pairs were selected such that only the target and its competitor would complete the fragment with a word that was associatively related to the cue. The six buffer pairs from study were used for practice, and the remaining 60 pairs served as critical items.

Procedure

All the stimuli were presented in white lowercase letters on a black background in the center of a computer screen. In list 1, word pairs were presented three times each in a fixed random order, with the restriction that no more than three pairs from the same condition were presented consecutively. Each pair was presented for 2 s, followed by a 500-ms interstimulus interval (ISI). Participants were instructed to read each word aloud quickly and accurately.

Participants in the test-present group were then presented with cue–fragment pairs in random order, and were told to complete the fragment with the target that was paired with the cue in list 1. Items remained on the screen until a response was made, followed by a 500-ms ISI. No effort was made to equate the spacing between lists 1 and 2 in the list 1 test-present versus test-absent groups, because the list 1 test was quite short (i.e., 10 items).

In list 2, pairs were presented in random order for 2 s, followed by a 500-ms ISI. Participants were instructed to read the pairs aloud and to study them for an upcoming memory test. Finally, on the list 2 test, cue–fragment pairs were presented in a fixed random order, with the restriction that no more than three pairs from the same condition were presented consecutively. Pairs remained on the screen until a response was made and were followed by a 500-ms ISI. Participants were instructed to complete the fragments with targets from list 2. Following each of their responses, participants made confidence judgments regarding the likelihood that their response was correct on a scale from 0 (*wild guess*) to 100 (*certain correct*). All responses were made aloud and were recorded by the experimenter. Participants were then given corrective feedback following each response. For target responses, “Correct” appeared in

green ink, whereas “Incorrect” appeared in red ink for nontarget responses. The entire procedure was then repeated in a second round with a new set of materials.

Results and discussion

Preliminary inspection of the results revealed that the probability of correctly responding to items in the facilitation condition was near ceiling, and, so, we analyzed results from that condition separately from performance in other conditions. Our primary interest was in whether prior experience with PI would diminish its effects. Diminishing effects of PI would be evidenced by a significant interaction between round (first vs. second encounter with PI) and item type (control vs. interference) in the probability of intrusion errors. 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 for purposes of comparison with intrusion errors to interference items. The probability of producing either the target or its alternate was extremely high for facilitation (.98), control (.94) and interference (.98) items, and, so, errors other than producing an alternate response as an intrusion error did not enter into the analyses.

The probability of a correct response for items tested from list 1 was near ceiling and did not differ between round 1 and round 2 (.97 vs. .97), $t < 1$. On the test of memory for list 2, performance for facilitation items was near ceiling and did not differ depending on whether or not items that appeared only on list 1 had been tested (.92 vs. .91) and did not differ across rounds 1 and 2 (.92 vs. .92), $F_s < 1$. These results allow one to dismiss the possibility that any reduction in PI across rounds was caused by reduced memory for list 1, the source of PI, across rounds.

The analysis of intrusion errors revealed that PI was reduced across rounds. A 2 (test) \times 2 (item type) \times 2 (round) ANOVA revealed a significant interaction between round and item type (control vs. interference), $F(2, 140) = 9.82$, $\eta_p^2 = .12$. Further analysis revealed significantly more intrusion errors in the interference condition in the first (.38) than in the second (.33) round, $t(71) = 2.93$, $d = 0.36$. The main effect of whether or not list 1 items were tested did not approach significance, $F(1, 70) = 1.19$, nor did the interaction of prior testing with item type or rounds, $F_s < 1$. Consequently, the results were collapsed across those groups for display in the top section of Table 1. As is shown in that table, the probability of an intrusion error decreased across rounds for interference pairs but did not do so for control pairs.

We examined the extent to which confidence judgments were diagnostic of correct responding by measuring the

Table 1 Probability of list 1 intrusion as a function of item type and round: experiments 1–3

	Item Type		
	Facilitation	Control	Interference
Experiment 1			
Round 1	.06 (.01)	.13 (.01)	.38 (.02)
Round 2	.07 (.01)	.14 (.01)	.33 (.02)
Experiment 2			
Round 1	.04 (.01)	.10 (.02)	.27 (.03)
Round 2	.06 (.01)	.12 (.02)	.19 (.03)
Experiment 3			
Round 1	.07 (.01)	.19 (.03)	.35 (.04)
Round 2	.08 (.02)	.18 (.02)	.30 (.03)

Note: standard errors of the means are presented in parentheses

resolution of confidence judgments. *Resolution* refers to the item-level correlations between confidence and accuracy. These correlations provide information regarding how well confidence judgments discriminate between items that were correctly, as compared with incorrectly, recalled. That is, they provide an estimate of the extent to which high-confidence judgments are associated with correct responses and low-confidence ratings are associated with incorrect responses. Positive gamma correlations indicate effective discrimination between correct and incorrect responses. Note that in subsequent analyses, gamma correlations could not always be computed for all the participants, as a result of a constant value on one of the variables. That is, gamma correlations could not be computed in cases in which recall performance was perfect or when the same confidence judgment was given for all the items. Consequently, the degrees of freedom are lower in some analyses than would be expected, given the sample size.

To assess resolution, we computed the gamma correlations for each combination of conditions and then entered those correlations into an ANOVA (see T. O. Nelson, 1984, for a detailed rationale for using gamma). The results from that analysis revealed a significant interaction between item type and prior testing of list 1, $F(1, 67) = 4.82$, $\eta_p^2 = .07$. As is shown in Table 2, resolution was higher for control pairs than for interference pairs. For the group without prior tests of list 1, this difference in resolution was decreased by resolution's increasing for interference pairs from round 1 to round 2, particularly for interference items. We replicated the finding of an increase in resolution for the interference condition across rounds in Experiment 2. In contrast, the results for the group that had a prior list 1 test showed a slight decrease in resolution between rounds 1 and 2 for both control items and interference items. We have no

Table 2 Resolution for confidence judgments as a function of item type, list 1 test, and round: experiment 1

List 1 Test	Item Type	
	Control	Interference
Absent		
Round 1	.70 (.06)	.46 (.05)
Round 2	.72 (.06)	.56 (.06)
Present		
Round 1	.77 (.06)	.38 (.05)
Round 2	.72 (.06)	.37 (.06)

Note: standard errors of the means are presented in parentheses

explanation for this finding of a slight decrease in resolution for interference items, and it was not replicated in later experiments. As will be seen, however, we do find a significant decrease in later experiments in the resolution of confidence judgments for control pairs across rounds.

In sum, the results of Experiment 1 provide evidence that prior experience with PI reduces its effects and that the reduction is not due to reduced memory for the source of PI. The probability that items presented in list 1 would be recalled did not differ across rounds, whereas it would be expected to do so if attention were reduced in the second round.

The lack of an effect of testing on PI contrasts with recent findings that testing reduces PI in free recall learning (e.g., Brewer et al., 2010; Szpunar et al., 2008). There are many differences between our experiment and their experiments. Among those differences is the fact that we examined PI in the learning of paired associates, whereas they examined free recall learning. Also, we included facilitation pairs (i.e., A–B, A–B pairs), and the presence of those pairs might have reduced the list differentiation that would otherwise have resulted from testing. Most important, perhaps, we tested only filler items, not the critical items used to assess PI. In contrast, those finding testing effects on PI tested items that were used to assess PI.

Experiment 2

In Experiment 2, we returned to the procedure of not testing memory for list 1, so as to replicate the results reported by Jacoby et al. (2010, Experiment 2). As did they, we used a self-allocated study time procedure to show that prior experience with PI was required for participants to become aware of and reduce its effects. The major difference between our experiment and theirs was that they used a free

versus forced responding technique (Koriat & Goldsmith, 1996) for tests of PI, whereas we did not do so. We did, though, use a modified form of that technique in Experiment 3, which is described when that experiment is introduced.

Jacoby et al. (2010) found that study time did not differ for interference and control pairs in the first round, showing that participants were unaware of memory difficulties produced by PI. However, in the second round, as a result of prior experience with PI, participants devoted more time to studying interference pairs than to studying control pairs, so as to diminish the effects of PI. The attempt was successful in that PI was reduced in the second round, as compared with the first. As important, the ability of confidence judgments to discriminate between correct responses and intrusion errors for interference items improved across rounds. Jacoby et al. (2010) interpreted these results as evidence for a qualitative change in the basis for responding across rounds. They argued that participants relied more heavily on source-constrained retrieval in the form of recollection in the second round, whereas the less constrained fluency with which a response came to mind was the primary basis for responding in the first round.

Analyses of the results from the present experiment examined differences in the resolution of confidence judgments in order to gain evidence of a qualitative change in the basis for confidence across rounds. As was described earlier, the resolution of confidence judgments measures the extent to which confidence in a response predicts its accuracy at the level of individual items. We expected an interaction between item type (control vs. interference) and rounds to be produced by a qualitative change in the basis for confidence across rounds. Participants' confidence judgments in the first round were expected to rely heavily on a fluency heuristic (Jacoby, Bishara, Hessels, & Toth, 2005; Jacoby & Dallas, 1981), with the highest confidence held for responses that most easily came to mind. Reliance on fluency would provide a valid basis for confidence judgments for control items, because a competitor was not presented for those items earlier. However, for interference items, reliance on fluency would result in erroneous confidence in intrusion errors, because of encounters with the intruding response during list 1. In the second round, having gained experience with PI, participants were expected to shift from heavy reliance on fluency to the ability to recollect the list 2 presentation of a tested item, a more valid basis for confidence in responses to interference items. A qualitative change in bases for confidence of this sort would result in improved resolution of confidence judgments across rounds for interference items. In contrast, for control items, shifting away from reliance on fluency might produce the opposite effect of reducing resolution, because fluency serves as a valid basis for confidence in

responses to control items. Abandoning this valid basis for confidence might reduce resolution for control items, because fluency may serve as a basis for high confidence in responses whose prior study encounter could not be recollected.

Method

Participants

Twenty-four Washington University undergraduates participated in exchange for course credit or \$10/hr. All the participants were tested individually.

Design, materials, and procedure

The design, materials, and procedure were identical to those used in Experiment 1, with the following exceptions. No items were employed for tests of memory for list 1 responses, resulting in a total of 132 three-word sets (e.g., *knee- bone, bend*). In both rounds, list 1 included 44 items presented 3 times each, for a total of 132 presentations. In addition, study time for list 2 items was allocated by participants. Participants were informed about the relationship between list 1 and list 2 pairings and were told to study the pairs presented in list 2 until they had been completely learned. Participants pressed the space bar upon completion of studying each pair.

Results and discussion

As was found in Experiment 1, performance on facilitation items was near ceiling (.95) and did not differ across rounds. The probability of producing either the target or its alternate was extremely high for facilitation (.99), control (.99), and interference (.99) items, and, so, errors other than producing an alternate response as an intrusion error did not enter into the analyses.

The middle section of Table 1 shows that the pattern of intrusion errors was similar to that found in Experiment 1. PI effects were diminished by prior experience, as is shown by a significant interaction indicating that the difference between intrusion errors for control and interference conditions was smaller in the second round than in the first, $F(1, 23) = 13.62$, $\eta_p^2 = .37$. Also, self-paced study had a tendency to further diminish PI, in comparison with the results obtained in Experiment 1, which employed a fixed study time procedure.

As was predicted, the resolution of confidence judgments, as indexed by the gamma correlation between confidence and accuracy, increased across rounds for the

interference condition (.39 vs. .57) but decreased slightly across rounds for the control condition (.86 vs. .82), $F(1, 15) = 8.14$, $\eta_p^2 = .35$. This finding provides evidence that participants profited from prior experience with PI by shifting the bases for confidence judgments from reliance on a fluency heuristic in the first round to reliance on ability to recollect in the second round. As was described earlier, a shift of this sort would increase the resolution of confidence judgments for interference items but would reduce resolution for control items. To anticipate, this pattern of results was replicated in Experiment 3.

The opposite effects of prior experience with PI on resolution for control and interference items provides evidence of a qualitative shift in the bases for confidence that was aimed toward reducing high-confidence erroneous responses. Reliance on ability to recollect, rather than on fluency, as a basis for confidence would serve to reduce high-confidence errors to interference items and, so, would increase the resolution of confidence judgments in that condition. In contrast, for control items, fluency serves as a valid basis for confidence, and, so, its abandonment would decrease the resolution of confidence judgments for that condition. These opposite effects on the resolution of confidence judgments cannot be accounted for by positing a quantitative shift involving familiarity or any other single basis for confidence.

Prior experience with PI resulted in participants' altering their study of interference pairs. The pattern of study time allocation (Table 3) revealed that study time reflected item difficulty across item types better in the second round than in the first round. Study time did not differ between the control condition and the interference condition in the first round, whereas less time was spent on items from the control condition than on those from the interference condition in the second round, $F(1, 23) = 8.93$, $\eta_p^2 = .28$. The least amount of study time was allocated to items in the facilitation condition in both rounds, $F(1, 23) = 24.80$, $\eta_p^2 = .52$. Finally, the overall amount of study time decreased from the first round to the second round, $F(1, 23) = 11.27$, $\eta_p^2 = .33$.

Overall, these results replicate the results reported by Jacoby et al. (2010, Experiment 2). Allowing self-allocated study time further diminished the effects of PI in round 2, as compared with the results in Experiment 1, in which study time was experimenter controlled and, typically, less than was self-allocated by participants. Furthermore, the resolution of confidence judgments for interference items increased across rounds but slightly decreased across rounds for control items. Whereas we examined resolution, Jacoby et al. (2010) examined the effects of prior experience on differences in confidence judgments for correct responses and intrusion errors. Corresponding analyses of differences in confidence judgments for correct

Table 3 Study time allocation as a function of item type and round: experiment 2

	Item Type		
	Facilitation	Control	Interference
Round 1	4417 (933)	6418 (981)	6357 (959)
Round 2	3105 (456)	4354 (526)	5211 (660)

Note: standard errors of the mean are presented in parentheses. Observations exceeding 2.5 *SDs* above or below the mean in each within-participant condition were trimmed prior to analysis. Fewer than 3% of all observations were excluded

responses and errors in the present study revealed a pattern of results consistent with those in Jacoby et al. (2010) but are not reported in detail because they provide no new information. The increase in study time for interference, relative to control, pairs shows that participants became aware of the greater difficulty of interference items as a result of prior experience with PI and increased their study time as an attempt to overcome the effects of PI.

The overall reduction in study time across rounds suggests that prior experience with PI, as well as with the form of test, produced qualitative, as well as quantitative, changes in encoding processes. It is noteworthy that the probability of an error did not increase across rounds for facilitation and control items and substantially decreased for interference items, although the amount of time devoted to study decreased in the second round for all conditions. These shifts in study time provide evidence that participants became aware of the effects of PI and, consequently, modified their encoding strategies to better support recollection at the time of test.

What was the basis for the allocation of study time across types of pairs (facilitation, control, and interference pairs)? One possibility is that after gaining experience with PI, study time was allocated on the basis of the ease with which an alternative response came to mind. Doing so would result in additional study time being allocated to interference items, which had a readily accessible competitor. Alternatively, participants may have relied on their ability to detect a change in responses between list 1 and list 2 as a basis for allocating study time. With this strategy, little study time would be allocated to a facilitation pair that was recognized as having the same response as in list 1, whereas additional time would be allocated to an interference pair that was recognized as having a changed response.

Experiment 3

The results of Experiment 2 provided evidence that prior experience with PI reduced its effects by producing a shift

from reliance on fluency to greater reliance on recollection as a basis for responding and confidence in responses. Experiment 3 was designed to further examine the mechanisms underlying the reduction of PI by exploring changes in retrieval processes. Experiment 3 was similar to Experiment 2, with one change being that study time was held constant, as in Experiment 1. Also, participants were given the option of withholding responses. 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 (e.g., Koriat & Goldsmith, 1996). This option made it possible for participants to withhold incorrect responses to interference items so as to further diminish PI effects (cf. Jacoby et al., 2010, Experiment 2).

In addition to being forced to respond to each test item, participants were told that they were to retrieve the list 2 response but should also report any other candidate responses that came to mind prior to the response that they output. They were told to stop producing candidates when they thought that the list 2 response had been reached. This was done to measure how often participants reported that a competing list 1 response came to mind prior to or simultaneously with the target response from list 2 (cf. Halamish, Goldsmith, & Jacoby, 2008). Doing so is useful for understanding how experience with PI reduces its effects. One possibility is that experience results in participants being more likely to use source memory to reject a competing response after it comes to mind (cf. Johnson, Hashtroudi, & Lindsay, 1993; Winograd, 1968). On this source memory account, the reduction in PI that results from prior experience is produced by participants' improved ability to reject a strong competitor (i.e., a list 1 response) after it comes to mind. Alternatively, the results of Experiment 2 suggest that the reduction of PI by prior experience was produced by participants' learning to constrain retrieval processes so as to recollect the list 2 response. Heavier reliance on recollection would be expected to result in the competing response being less likely to come to mind (cf. Jacoby, Debner, & Hay, 2001; Jacoby, Shimizu, Velanova, & Rhodes, 2005).

Both the source memory and recollection accounts predict that the number of intrusion errors to interference pairs will decrease across rounds but will do so for different reasons. Examining changes in PI across rounds separately for cases in which an alternative response did or did not come to mind provides a means of choosing between the two accounts. The source memory account focuses on response competition and holds that prior experience improves the ability to choose the target over other responses that come to mind. The prediction is that when alternatives to a produced response are reported as having come to mind, the probability of the produced response being correct will increase across rounds. In contrast, the

recollection account holds that prior experience increases ability to constrain retrieval processes such that only the target response comes to mind. The prediction is that when no other responses are reported as having come to mind, the probability of the produced response being the correct one will increase across rounds.

Method

Participants

Twenty-four Washington University undergraduates participated in exchange for course credit or \$10/hr. All the participants were tested individually.

Design, materials, and procedure

The design, materials, and procedure were identical to those used in Experiment 2, with the following exceptions. The study duration for items in list 2 was fixed at 2 s in both rounds. At the time of test, participants were told that their primary task was to retrieve the responses presented in list 2. In addition, participants were told to report any other responses related to cues that fit the fragments as they came to mind during retrieval of list 2 responses. Participants were allowed to produce a maximum of two candidate responses and were told to stop once they thought that the list 2 response had been given. This was done in an attempt to restrict participants' candidate production to intraexperiment responses. Participants reported the candidates aloud, and an experimenter typed them onto the computer screen.

After giving their final responses, participants rated their confidence aloud, using the same scale as that in Experiments 1 and 2. Next, participants were given a report option that allowed them to either volunteer or withhold responses. A screen appeared that read "Report?" with the options to respond yes or no aloud. Participants were told that they should volunteer responses only when they felt sufficiently confident that the responses were correct. As an incentive to use the report option to maximize their memory accuracy, five points were added to a running total for correct volunteered responses, whereas 15 points were deducted for incorrect volunteered responses. For withheld responses, no points were gained or lost, and no feedback was provided. The running point total was displayed in the upper right-hand corner of the screen.

Results and discussion

As was found in Experiments 1 and 2, performance on facilitation items was near ceiling (.91) and did not differ

across rounds. The probability of producing either the target or its alternate was extremely high for facilitation (.98), control (.95), and interference (.99) items, and, so, errors other than producing a list 1 intrusion did not enter into the analyses.

In the following analyses, forced-report tests include all the responses given at the time of test, whereas free-report tests include only the volunteered responses. The probabilities of an intrusion error for the forced- and free-report tests are shown in Table 4. Performance for facilitation pairs was near ceiling, and, so, there was little room for free report to improve accuracy for that condition. An analysis that included only the control and interference conditions revealed that allowing free report did more to increase accuracy of responding for control items than for interference items, $F(1, 23) = 8.07$, $\eta_p^2 = .26$. The triple interaction of report option, round, and item type (control vs. interference) only approached significance, $F(1, 23) = 2.88$, $p = .10$, $\eta_p^2 = .11$, suggesting that there was a tendency for the accuracy advantage gained by allowing free report for control items, as compared with interference items, to be larger in round 1 than in round 2. However, the results reported by Jacoby et al. (2010) gave a priori grounds for analyzing results separately for free versus forced responding, although the relevant interaction only approached significance. These further analyses revealed that although intrusions on interference items decreased across rounds for both forced and free report, the difference was significant only for free report, $t(23) = 2.17$, $d = 0.35$. The results are similar to those reported by Jacoby et al. (2010) in suggesting that the option to not respond allowed participants to further diminish the effects of PI.

An analysis of gamma correlations between confidence judgments and accuracy on the forced-report test revealed a significant interaction between item type (control vs. interference) and round, $F(1, 19) = 7.00$, $\eta_p^2 = .27$. The resolution of confidence judgments increased across rounds for interference items (.31 vs. .43) but decreased across rounds for control items (.75 vs. .60). The form of the

Table 4 Probability of list 1 intrusion as a function of item type, report, and round: experiment 3

	Item Type		
	Facilitation	Control	Interference
Forced			
Round 1	.07 (.01)	.19 (.03)	.35 (.04)
Round 2	.08 (.02)	.18 (.02)	.30 (.03)
Free			
Round 1	.04 (.01)	.09 (.02)	.32 (.04)
Round 2	.05 (.01)	.12 (.03)	.26 (.04)

Note: standard errors of the means are presented in parentheses

interaction is the same as that found in Experiment 2. Again, the interaction provides strong evidence of a qualitative change in the basis for confidence across rounds. A quantitative change cannot account for the opposite effects of prior experience with PI on resolution for control and interference items.

Decreased reliance on a fluency heuristic as a basis for confidence would decrease the resolution of confidence judgments for control items, because fluency serves as a valid basis of confidence for those items. Although recollection can be considered a more certain basis for confidence than is fluency, the failure to recollect does not always lead to incorrect responses. That is, fluency can still drive accurate responding in the absence of recollection for control items. Thus, confidence judgments that take fluency and recollection into account are likely to be more accurate assessments of memory performance than are those based on recollection success. In contrast, a decrease in reliance on fluency, along with increased reliance on recollection, as a basis for confidence for responses to interference items would increase the resolution of confidence judgments for interference items, because fluency is often a misleading basis for confidence in interference situations.

The reduction in resolution across rounds for control items was larger in Experiment 3 than in Experiment 2. This likely resulted from the requirement at test to report any responses that came to mind prior to the response that was output. That requirement, along with prior experience with PI, likely encouraged the abandonment of reliance on a fluency heuristic for confidence, to the disadvantage of the resolution of confidence judgments for responses to control items. In turn, the increase in resolution across rounds was smaller for interference items in Experiment 3 than in Experiment 2. This difference can be explained as arising because the self-allocated study time procedure employed in Experiment 2 allowed greater opportunity to overcome the effects of PI by studying in ways that enhanced later recollection.

Analysis of whether another response came to mind prior to the response that was output is informative with regard to means by which experience with PI diminished the probability of an intrusion error for interference items. Increased reliance on recollection during the second round was expected to increase the probability that a produced response was correct when no alternative response was reported as having come to mind. The probability of participants' saying that another response came to mind prior to the response that they output on tests for interference items was low on both round 1 and round 2 (.35 vs. .31).

When the conditional probability of an intrusion error given that no other response came to mind prior to the response that was output was examined, the probability of an intrusion for interference items decreased across rounds (.43 vs. .34), $t(23) = 2.21$, $d = 0.39$. This result is in accord with the expected

heavier reliance on recollection during the second round than during the first (i.e., the source-constrained retrieval account). In contrast, the source-monitoring account predicts that the reduction in PI across rounds results from an increased ability to successfully resolve response competition by choosing the correct response when alternative responses have also come to mind. Against that account, when alternatives to the correct response were reported as having come to mind, there was little decrease in the probability of an intrusion error across rounds (.30 vs. .28), $t(21) < 1$.

An unexpected finding was that the probability of an intrusion error for interference items was lower when another response came to mind prior to the response that was output than when another response did not come to mind, $F(1, 21) = 5.93$, $\eta_p^2 = .22$. We return to discussion of this unexpected finding after further considering conclusions that can be drawn regarding the recollection and source memory accounts of reduced PI.

Although reliance on source memory in rejecting a competing response after it comes to mind might play some role, our results provide evidence that increased reliance on recollection is largely responsible for prior experience with PI diminishing its effects. When only one response comes to mind for a test item in the interference condition, that response is likely to be the competitor if one relies upon overall strength, because of the multiple presentations of the competitor in list 1. However, as a result of prior experience with PI, participants became more likely to engage in source-constrained retrieval, so as to recollect the list 2 response (i.e., the target). Doing so increased the probability of their producing the target without other alternatives coming to mind. That is, participants shifted from reliance on a less-constrained fluency basis for responding in round 1 to greater reliance on recollection as a basis for responding in round 2. The reduction in intrusion errors across rounds when alternative responses did not come to mind converges with the effects of prior experience on the resolution of confidence judgments in providing evidence for a shift toward heavier reliance on recollection.

Why was the probability of an intrusion error lower when other responses were reported as having come to mind? One possibility is that source memory might sometimes serve to reject a competitor after it comes to mind. An incorrect response's coming to mind as a competitor for the target would be more advantageous than would having an intrusion error be sufficiently strong as to be produced without the target's coming to mind. A high percentage of the responses that were produced without other responses coming to mind were intrusion errors. Obviously, the target's not coming to mind made it less likely to be produced as a response, as compared with the case in which it came to mind along with alternative responses. Although the probability of recollection in-

creased across rounds, a less constrained, strengthlike basis for responding still played a role in the second round, as evidenced by the high probability of an intrusion error being presented without an alternative response coming to mind. The recall advantage of other items' coming to mind might reflect the strengthlike basis for responding.

An alternative possibility is that the advantage of other items' coming to mind might result from participants' reliance on mediation processes for correct responding to interference items. The materials in our experiments were selected to produce an overlap in letters between a target and its alternative response, so as to allow both to be a possible completion for the same word fragment. As well as this orthographic similarity, for some pairs, there was an associative relation between a target word and its competitor. In this vein, Barnes and Underwood (1959) investigated retroactive interference with the materials being such that list 1 and list 2 responses were associatively related (A–B, A–B) and found evidence that access to list 1 response was sometimes mediated by the list 2 response. Although extraexperimental associations between list 1 and list 2 responses could potentially account for the benefit of another response's coming to mind prior to a final response, the average forward-associative strength between list 1 and list 2 responses (see D. L. Nelson et al., 1998) for items in our experiments was quite low (.03). Thus, the advantage of other responses' coming to mind may have reflected mediation processes. However, such mediation played little, if any, role in the increased resistance to PI across rounds. The probability of other responses being reported as having come to mind did not increase across rounds, nor was there a substantial decrease across rounds in the probability of an intrusion error when other responses were reported as having come to mind. Either one or both effects would be expected had mediation played a role in the reduction of PI across rounds.

General discussion

The results from our experiments replicate the results reported by Jacoby et al. (2010) by showing that prior experience with PI diminished its effects. In addition, Experiment 1 provided evidence that the reduction in PI across rounds did not occur because of reduced attention in the second round to list 1, the source of PI. Ability to recall list 1 responses was shown to be near ceiling and to not change across rounds. The results from Experiments 2 and 3 converge to produce strong evidence that prior experience reduces its effects by means of a shift toward heavier reliance on recollection as a basis for responding. That shift reflects both a change in encoding and a change in retrieval processes.

Changes in the allocation of study time revealed that participants became more sensitive to the difficulty of

interference items after experience with PI. Participants did not devote more time to interference than to control items in the first round, but did so in the second round (Experiment 2). Participants were more aware of PI effects in the second than in the first round and compensated for those effects by devoting more study time to interference than to control items. In addition to a quantitative shift, it is likely that there was also a *qualitative* shift in encoding processes. Less study time was devoted to all the items in the second round, yet accuracy for interference items increased across rounds. One possibility is that participants studied items in ways that allowed them to better encode the list membership of items during the second round.

The resolution of confidence judgments in the interference condition improved as a result of prior experience with PI. In contrast, the resolution of confidence judgments for control items decreased or was unchanged across rounds. The finding that the resolution of confidence judgments in the interference condition improved across rounds provides evidence that prior experience with PI resulted in a shift away from a fluency heuristic (e.g., Benjamin, Bjork, & Schwartz, 1998; Kelley & Lindsay, 1993) to reliance on the success of recollection as a more valid basis for confidence. Further evidence of a shift in bases for confidence is provided by the finding that monitoring the resolution of confidence judgments decreased across rounds for control items. This reduction would be expected if participants switched to reliance on recollection success, reducing reliance on less constrained accessibility of a response. The fluency with which a response comes to mind is a valid basis for confidence for control pairs but is a misleading basis of confidence for interference pairs. For interference pairs, if one is not constraining retrieval in ways that support recollection, the competing response is likely to fluently come to mind.

Finally, in Experiment 3, reductions in PI did not result from increased use of source memory to reject a competitor after it came to mind. Instead, as a result of prior experience with PI, participants more often produced the list 2 response without the competitor even coming to mind. The increase in the probability of a target response's coming to mind as the only response would be expected if experience with PI had the effect of encouraging participants to rely on recollection as a basis for responding. Recollection relies on preretrieval processes that employ the source of the target to elaborate the cues provided by retrieval so as to restrict access to the target item (e.g., Jacoby, Shimizu et al., 2005; Jacoby, Kelley, & McElree, 1999). This source-constrained retrieval, which operates prior to memory access, contrasts with postaccess source monitoring. Source monitoring serves to edit potential responses after they come to mind so as to withhold candidate responses that did not originate from the target source (e.g., Koriat & Goldsmith, 1996).

The results from the present experiments converge with those from earlier experiments in showing that success in avoiding the negative effects of PI depends on engaging in successful recollection. Hay and Jacoby (1999) examined the effects of PI by using procedures similar to those used in the present experiments. However, conditions were changed so as to implement a process dissociation procedure for estimating the probability of recollection. They gained evidence suggesting that PI results from a form of bias that has an effect only when recollection fails. Jacoby et al. (2001) reported similar results and related recollection to the subjective experience of remembering.

By emphasizing encoding processes in combination with retrieval processes, we join Postman and Underwood (1973) in their evaluation of the importance of list differentiation for identification of the list membership of a response after it comes to mind. They wrote that ". . . the critical factor is not the subject's ability to identify the list membership of whatever responses do occur but rather the mechanism governing the availability of alternative response repertoires for recall" (p. 24). We agree, as do our data. Learning to diminish the effects of PI relies more on enhancing recollection so as to prevent a competing response from coming to mind than on source memory, the ability to identify the list membership of a response after it comes to mind.

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 (2nd ed.)*, Memory (pp. 237–313). San Diego: Academic Press.
- Barnes, J. M., & Underwood, B. J. (1959). "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology*, *58*, 97–105.
- Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, *127*, 55–68.
- Brewer, G. A., Marsh, R. L., Meeks, J. T., Clark-Foos, A., & Hicks, J. L. (2010). The effects of free recall testing on subsequent source memory. *Memory*, *18*, 385–393.
- Halamish, V., Goldsmith, M., & Jacoby, L. L. (2008, November). *Source constrained recall: Strategic control of retrieval quality*. Poster presented at the annual meeting of the Psychonomic Society, Chicago.
- Hay, J. F., & Jacoby, L. L. (1996). Separating habit and recollection: Memory slips, process dissociations, and probability matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1323–1335.
- Hay, J. F., & Jacoby, L. L. (1999). Separating habit and recollection in young and older adults: Effects of elaborate processing and distinctiveness. *Psychology and Aging*, *14*, 122–134.
- Jacoby, L. L. (1996). Dissociating automatic and consciously controlled effects of study/test compatibility. *Journal of Memory and Language*, *35*, 32–52.
- 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., 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., Debnar, J. A., & Hay, J. F. (2001). Proactive interference, accessibility bias, and process dissociations: Valid subjective reports of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 686–700.
- Jacoby, L. L., Kelley, C. M., & McElree, B. D. (1999). The role of cognitive control: Early selection versus late correction. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 383–400). New York: Guilford Press.
- Jacoby, L. L., Shimizu, Y., Velanova, K., & Rhodes, M. G. (2005). Age differences in depth of retrieval: Memory for foils. *Journal of Memory and Language*, *52*, 493–504.
- Jacoby, L. L., Wahlheim, C. N., Rhodes, M. G., Daniels, K. A., & Rogers, C. S. (2010). Learning to diminish the effects of proactive interference: Reducing false memory for young and older adults. *Memory & Cognition*, *38*, 819–828.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.
- Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, *32*, 1–24.
- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490–517.
- Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*, 109–133.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. Retrieved from <http://www.usf.edu/FreeAssociation/>.
- Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. *Memory & Cognition*, *1*, 19–40.
- Szpunar, K. K., McDermott, K. B., & Roediger, H. L., III. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1392–1399.
- Tulving, E., & Watkins, M. J. (1974). On negative transfer: Effects of testing one list on the recall of another. *Journal of Verbal Learning and Verbal Behavior*, *13*, 181–193.
- Underwood, B. J. (1957). Interference and forgetting. *Psychological Review*, *64*, 49–60.
- Winograd, E. (1968). List differentiation as a function of frequency and retention interval. *Journal of Experimental Psychology*, *76*, 1–18.
- Christopher N. Wahlheim and Larry L. Jacoby, Department of Psychology, Washington University in St. Louis, MO. This research was supported by Binational Science Foundation Grant 2005356 and by a James S. McDonnell Foundation 21st Century Science Initiative in Bridging Brain, Mind, and Behavior Collaborative Award. We thank Sarah Arnspiger, Rissa Ivens, Miranda Lindberg, Danielle Hirschfeld, Carlee Beth Hawkins, Carole Jacoby, and Rachel Teune for their assistance with data collection. Correspondence concerning this article should be addressed to Christopher N. Wahlheim, Department of Psychology, Washington University, St. Louis, MO 63130. E-mail: cnwahlheim@gmail.com.