Evidence for proactive interference in the focus of attention of working memory

Lauren M. Carroll
Annie Jalbert
Alexander M. Penney
Ian Neath
Aimée M. Surprenant
Gerald Tehan

1 Memorial University of Newfoundland, St. John's, NL, Canada

2 University of Southern Queensland, Springfield, Australia

Correspondence:
Ian Neath
Psychology Department
Memorial University of Newfoundland
St. John's, NL
A1B 3X9
Canada

Phone: +1 (709) 737-8495
Fax: +1 (709) 737-2430
Email: ineath@mun.ca
Proactive interference (PI) occurs when an earlier item interferes with memory for a newer item. Whereas some researchers (e.g., Surprenant & Neath, 2009a) argue that PI can be observed in all memory systems, some multiple systems theorists (e.g., Cowan, 1999) propose that items in the focus of attention of working memory are immune to PI. Two experiments tested whether PI occurs when the to-be-remembered items are assumed, by multiple-systems theorists, to be held in the focus of attention. In each experiment, subjects saw four trials in a row with the same type of to-be-remembered items, followed by four trials in a row with a different type of material. On each trial, only 3 stimuli were shown, which is below the capacity limit of the focus of attention, and subjects were asked if a probe item was one of those 3 items seen. In both experiments, response time increased from Trial 1 to Trial 4, suggesting that items from the earlier trials interfered with memory on the later trials. In addition, release from PI was shown in that response times decreased with a change of materials. The results replicate those first reported by Hanley and Scheirer (1975), and pose a problem for theorists who argue that parts of short-term memory are immune to PI.
Proactive interference (PI) occurs when an earlier item interferes with memory for a newer item. Although it is generally taken for granted that PI is a fundamental source of forgetting in long-term memory (LTM), there remains a debate about whether PI is observable in short-term memory (STM) or working memory (WM). There are some theorists who posit that at least part of WM is immune to PI, and there are many studies that fail to observe PI when the number of items to be remembered is below the presumed capacity of this store (see Surprenant & Neath, 2009a, for a review). The present work was designed to replicate one of the few studies that did observe significant PI in a short-term or working memory situation, Hanley and Scheirer (1975). Observing PI under such conditions will have implications for those theories that propose that a small number of items can be registered in a limited capacity store thereby making them immediately accessible and isolated from all preceding memories.

A consistent theme in PI research is that some forms of memory are immune to PI (Halford, Maybery & Bain, 1986; Wickens, Moody & Dow, 1981) with the most recent example being Cowan’s (1999) embedded processes model. This model distinguishes between (1) long-term memory; (2) working memory, which is considered to be the activated portion of long-term memory; and (3) the focus of attention. The focus of attention has many similarities to William James’ (1890, p. 609) description of primary memory: “An object of primary memory is not thus brought back; it never was lost; its date was never cut off in consciousness from that of the immediately present moment. In fact, it comes to us as belonging to the rearward portion of the present space of time, and not to the genuine past.”

Although Cowan’s (1999) embedded processes model shares some features with traditional short-term and working memory theories, it differs in a number of important ways. For example, it emphasizes the idea that working memory is a set of processes carried out in
long-term memory, and that working memory is simply the subset of information in long-term memory that is temporarily activated. In contrast, Baddeley (2003) views working memory as a fundamentally separate memory system that does not overlap with long-term memory, and working memory itself is further fractionated into subsystems. Cowan’s view also emphasizes the role of the focus of attention in memory performance whereas there is no comparable contribution from attentional components in Baddeley’s model.

A third difference, and the focus of this paper, is that one component of the embedded processes model is immune to PI, whereas Baddeley’s (2003) working memory, just like other forms of memory, is susceptible to PI. Indeed, immunity to PI has been instrumental in determining the capacity of the focus of attention, which is estimated to be approximately 4 unrelated items (Cowan, 2000). Because of this limit, “one can observe proactive interference (PI) in retrieval only if there are more than four items in a list to be retained… This presumably occurs because four or fewer items are, in a sense, already retrieved; they reside in a limited-capacity store, eliminating the retrieval step in which PI arises” (p. 103).

Although Cowan’s (1999, 2000) model is the most well-developed, other theorists have also posited immunity to role to PI in what is variously termed primary memory, short-term memory, or immediate memory. For example, Schweickert’s (1993) multinomial model holds that some items need not be retrieved. Within the model, “subjects carry out immediate recall by first attempting direct readout and then guessing if that fails” (p. 70; see also Estes, 1991). "Direct readout" is not further explained, but the implication is that the information needed is still resident in consciousness and therefore does not need retrieving. At the time of the memory test, there is a probability, I, that an item is intact and can be directly read-out, and a probability, 1-I, that the item is not intact. If the item is intact, it is correctly recalled. This suggests that PI
cannot affect such items. Similarly, Wickens, Moody, and Dow (1981, p. 17) state that “the retrieval act is not required in the PM [primary memory] situation.”

Of course, there are numerous studies that demonstrate PI in short-term and working memory tasks when the number of to-be-remembered items is below span but greater than four items (for a review, see Humphreys & Tehan, 1999). Thus, from Cowan’s perspective, it might be argued that in those tasks at least some of the items are no longer in the focus of attention and thus must be retrieved. The key issue is whether PI can be demonstrated when the number of to-be-remembered items does not exceed the capacity of Cowan’s focus of attention.

There are a number of studies that support the prediction of Cowan’s (2000) embedded process model that items in the focus of attention are immune to PI. For example, Halford, Maybery, and Bain (1988, Experiment 1) used a variant of the Sternberg (1966) memory scanning procedure. The subjects saw a list of words presented all at once; there could be 4 or 10 words in the list, presented for 5.8 or 13 seconds, respectively. One second after the list disappeared, a probe appeared. The subject's task was to indicate if the probe was or was not in the list. PI was manipulated by using categorized lists. There were 3 consecutive test trials using words in one taxonomic category (i.e., trees), then 3 consecutive trials using another category (e.g., sports), then 3 consecutive trials using another category (e.g., spices), and so on. Trial 1 for any one category was designated as a low PI trial, as there were no trials immediately preceding it on which the same type of items had occurred. Trial 3 for any one category was designated as a high PI trial, because it followed two trials that contained the same types of items. With 10 items, subjects were faster and more accurate for the low PI condition than the high PI condition, suggesting PI was observed. With only 4 items, however, there was no difference between the low PI and high PI conditions. Halford et al. reported a second experiment that used rhyme rather
than semantic categories, but also included lists of 6 and 8 items. Again, the response times for the 4-item lists did not differ between the low PI and high PI conditions, but did for the 6-, 8-, and 10-item lists.

Other studies have also failed to find evidence of PI when lists are shorter than 4 items. For example, Cowan, Johnson, and Saults (2005) used a procedure similar to that of Halford et al. (1988) that relied on category membership to produce interference. They found effects of PI with lists lengths of 4 and more, but not with a list length of less than four. Similarly, Wickens et al. (1981) also used taxonomic category to induce interference, and observed PI for longer list lengths but not for shorter list lengths.

Cowan et al. (2005) note two studies – Monsell (1978) and McElree and Dosher (1989) – that did find PI with list lengths shorter than 4. Unlike the Cowan et al. and Halford et al. (1988) studies, those of Monsell and McElree and Dosher did not rely on semantic relations to generate PI. Rather, both of these studies found evidence of PI when examining performance on negative trials. RTs were slower when the negative probe was an item from an earlier trial compared to when it was a novel item. These results are similar to those reported earlier by Atkinson, Herrmann, and Wescourt (1974). They showed subjects 2, 3, 4, or 5 words, followed by a probe. The probe could be one of the items in the list (a P or positive trial), or it could be a word that was not on the list (an N or negative trial). Atkinson et al. manipulated the number of times the probe word had been experienced. N1 meant that the word was being seen for the first time. N2 meant that the word was being seen for the second time, having also been seen as a distractor on the immediately preceding trial. N3 meant that the word was being seen for the third time, having been seen on the preceding list as both a member of the list and as the probe. Mean RT systematically increased for correct responses on N1, N2, and N3 trials. Even when there were
only 2 items in the memory set, the mean RT for an N1 trial was 634 ms compared to 658 ms and 668 on N2 and N3 trials respectively. The same was true for trials on which there were 3 items in the memory set: RTs increased from 661 to 679 to 714 ms.

Thus, the available evidence is that PI is not observable on short lists (i.e., lists with about 4 or fewer items) except in the case when a prior item is used as a probe and even then, only on negative trials. Given this, Cowan et al. (2005, p. 297) concluded that the results from the available studies on PI in short lists is “broadly consistent” with the claim that items in the focus of attention are immune to PI.

There is one other study that demonstrated PI in a situation in which the number of to-be-remembered items was less than the presumed capacity of the focus of attention, but curiously, this study, Hanley and Scheirer (1975), has not been directly replicated, and indeed, has been cited only once to our knowledge (by Sanders & Willemsen, 1978). Hanley and Scheirer also used a variant of the Sternberg task, in which subjects saw 3 items presented sequentially at a rate of 800 ms each, with a 200 ms interval between items. Two seconds after presentation of the final item, a probe item was presented, and subjects were to indicate if the item was one of the three just seen. Every four trials, the to-be-remembered items changed from two-digit numbers to two letters (or vice versa), just as is done in the build-up and release of PI experiments (e.g., Wickens, 1970). Hanley and Scheirer found that subjects were slower on Trial 4 of a given stimulus type than on Trial 1, exactly what would be expected if PI built up over trials. What is most important is that this effect was found for both positive and negative trials, was not restricted to just recent negatives, and the number of to-be-remembered items was below the capacity of the focus of attention.
Why did Hanley and Scheirer (1975) find evidence of the build-up of PI when the other studies cited above did not? One possibility is that the stimuli were not sufficiently confusable to allow for the observation of PI in such short lists. Halford et al. (1988, Experiment 1) and Cowan et al. (2005) used unique words and categories on every trial. Based on the results from a very different paradigm, Tehan and Humphreys (1996, 1998) have argued that both the presence of PI and immunity to PI are a function of the relation between the cues available to the participant and the particular test. We suggest that in the Sternberg task, semantic cues are sufficient to discriminate the 4 items in the current memory set from that of previous trials so no PI is demonstrated. With less meaningful information, such as nonwords or two-digit numbers, the cues available at test should be less discriminable and this will allow PI to appear. Another contributing factor may have been that Hanley and Scheirer had 4 trials for PI to build up whereas, for example, Halford et al. had only three.

Although the results of Hanley and Scheirer (1975) are consistent with the claim that PI can be observed in the focus of attention, there are two possible problems with their study. First, as noted above, there are no published replications. It could be the case that this pattern is not replicable. Second, it is possible that a multiple systems theorist might argue that letter pairs are not unitized, and therefore, subjects were really being asked to remember 6 items rather than 3. The purpose of the two experiments reported below is to replicate the original findings using two new types of stimuli to ensure that only 3 items need be retained on each trial. If PI is observed, it will be evidence against the multiple systems claim, and will support the more general view that PI is observable in all memory systems.

Experiment 1
Experiment 1 was designed to replicate Hanley and Scheirer (1975). As in the original study, a Sternberg memory-scanning paradigm was used in which subjects saw three items and were then asked to indicate if a probe item was one of the three items just seen. Every four trials, the type of stimuli changed. One type of stimulus was the two-digit numbers used by Hanley and Scheirer, allowing for easier comparison between our results and theirs. Rather than using letter pairs, however, we used four-letter, one-syllable nonwords. The rationale was that pronounceable four-letter nonwords (e.g., hice, lant, ronk) should be processed verbally and thus be unitized, and three of them should be below the capacity limit of the focus of attention. Moreover, unlike the words used by Halford et al. (1988) and Cowan et al. (2005), the nonwords are less meaningful and therefore more likely to be confused.

**Method**

*Subjects.* Thirty-two members of the Memorial University community, all between 19 and 30 years old, volunteered to participate and received a small honourarium. All were native speakers of English.

*Materials.* The number stimuli were fifty-six two-digit numbers, assembled using the digits 2-9 but omitting all numbers such as 22, 33, 44, etc. Fifty-six one-syllable four-letter nonwords were selected at random from an online database (Medler & Binder, 2005; see the appendix for the specific non-words used).

*Design.* The experiment used a two stimulus type (numbers or nonwords) × two test type (probe present or absent) × four (trial number) completely within-subjects design.

*Procedure.* Subjects were informed that they would see three items presented one at a time followed by a probe item and were asked to decide whether the probe item had been present on the previous list. They were asked to respond as quickly and as accurately as they could by
pressing one of two computer keys. The right index finger pressed the /-key to indicate a positive response (i.e., the probe was on the list) and the left index finger pressed the z-key to indicate a negative response (i.e., the probe was not on the list). All items were shown in a 36 point monospaced font.

As in Hanley and Scheirer’s (1975) study, the type of stimulus alternated every four trials. A positive test probe was presented in half the test trials and a negative test probe was presented in the other half of the trials. A trial consisted of the presentation of three nonwords one at a time followed by a nonword probe or three two-digit numbers followed by a two-digit number probe. No number or nonword appeared more than once every 28 trials, and negative probes had not been seen at all during the previous 28 trials. The order of the trials was randomly determined for each subject.

Each trial began with the presentation of a fixation point in the middle of the computer screen. One second later, three items (nonwords or two digit numbers) were presented one at a time for 800 ms followed by a 200 ms blank interval. Following the third item, a warning sign (“????????????”) appeared to signal the end of the trial. One second later, the probe item was presented and remained on the screen until a response was made. The next trial began when the subject pressed the space bar. Subjects received 8 unscored practice trials prior to the 104 experimental trials. The order of trials was randomized for each subject.

Results and Discussion

For each subject, the overall mean RT for correct responses was calculated. Then, any RT that was more than three standard deviations above or below the mean was discarded and performance on that trial was counted as incorrect. A total of 84 trials were thus discarded, a
mean of 2.63 discarded trials per subject. All discarded trials had RTs greater than 3 standard deviations above the mean, and were evenly distributed over types of trials.

Figure 1 shows the mean RTs for correct responses for trials 1 through 4 as a function of stimulus type (numbers or nonwords) and test type (probe present or probe absent). The pattern of results for numbers replicated Hanley and Scheirer (1975): responses were slower on Trial 4 than on Trial 1, consistent with the idea that PI builds up over trials. The pattern for nonwords shows the same pattern as the numbers.

The RT data were analyzed using a two stimulus types (numbers vs. nonwords) × two test types (probe present vs. probe absent) × 4 trial numbers repeated measures analysis of variance. For all analyses, an alpha level of 0.05 was adopted.

There was no significant main effect of stimulus type ($F(1,31) < 1$). The mean RT for numbers was 895 ms compared to 888 ms for nonwords. The main effect of test type just failed to reach the adopted significance level ($F(1,31) = 3.45, MSE = 22160$, partial $\eta^2 = 0.100, p > 0.07$), although the pattern was for faster RTs on probe present trials compared to probe absent trials (879 vs. 903 ms). The main effect of trial number was significant ($F(3,93) = 6.69, MSE = 7288$, partial $\eta^2 = 0.178$). In general, RTs were faster on earlier trial numbers and slower on later trial numbers. The mean RTs for trials 1 to 4 were 873, 876, 903, and 913 ms, respectively. A contrast revealed a significant linear trend for trial number ($F(1,31) = 13.15, MSE = 10290$, partial $\eta^2 = 0.298$), but no quadratic component ($F(1,31) < 1$).

No interactions approached significance; indeed, for all, $F<1$. 
The accuracy data were also analyzed using a two stimulus types (numbers vs. nonwords) × two test types (probe present vs. probe absent) × 4 trial numbers repeated measures analysis of variance. There was no significant main effect of stimulus type ($F(1,31) < 1$), with the overall proportion of correct responses 0.891 for numbers compared to 0.938 for the nonwords. There is no indication of a speed-accuracy tradeoff. Neither the main effect of test type ($F(1,31) = 2.04$, $MSE = 0.020$, partial $\eta^2 = 0.062$) nor the main effect of trial number ($F(3,93) = 2.11$, $MSE = 0.011$, partial $\eta^2 = 0.069$) were significant.

The interaction between stimulus type and test type was significant ($F(1,31) = 5.52$, $MSE = 0.012$, partial $\eta^2 = 0.151$), due primarily to larger difference between numbers present and numbers absent trials relative to the difference between nonwords present and nonwords absent trials. The interaction between stimulus type and trial number approached significance ($F(3,93) = 2.34$, $MSE = 0.098$, partial $\eta^2 = 0.070$, $p > 0.07$), reflecting the tendency for accuracy to numbers to decrease slightly more over trials than accuracy to nonwords. Finally, neither the interaction between test type and trial nor the three-way interaction were significant (both $F$s < 1).

The RT data were also analyzed to examine release from PI. For each subject, the RT on Trial 4 of a block was compared to the RT on the immediately following Trial 1 (given correct responses to both). Response times were significantly slower on Trial 4 than on Trial 1 (912 ms vs 864 ms, respectively; $F(1,31) = 22.67$, $MSE = 1633$, partial $\eta^2 = 0.422$). In addition, a difference score was created (RT on Trial 4 minus RT on the immediately following Trial 1 in which both trials had a correct response). Then, the mean of these difference scores was computed for each subject. Twenty three subjects had a positive score, indicating slower responding on Trial 4 than Trial 1, and nine had a negative score, indicated faster responding on Trial 4 than Trial 1. This is significant by a sign test.
Experiment 1 replicated the finding of Hanley and Scheirer (1975) that response times on Trial 4 were slower than response times on Trial 1 when the stimuli were numbers. In addition, this finding was extended to nonwords. It is less plausible that short, pronounceable nonwords are retained or processed in terms of their components rather than as unitary items, because the easiest processing strategy is to just read them as unitary items. Nonetheless, one might have some objection to the claim that three nonwords (e.g., frim, tomp, blek) do not exceed the capacity of the focus of attention. The purpose of Experiment 2 was to replicate Experiment 1 using a different type of stimuli.

Experiment 2

For Experiment 2, we wanted a second kind of stimulus type that is unlikely to be processed in terms of its components. As two kinds of stimuli are required, we retained the two-digit numbers to facilitate comparison with Experiment 1. For the second type, we used common symbols (e.g., !@#$%).

Method

Subjects. Thirty-two different members of the Memorial University community volunteered to participate. As in Experiment 1, the participants were all between 19-30 years old, were all native speakers of English, and received a small honourarium for participating.

Materials. Sixteen symbols were selected, all of which appear on common computer keyboards (see the appendix). To match these, sixteen of the 56 two-digit numbers used in Experiment 1 were also used (see the appendix).

Design and Procedure. The design and procedure were identical to Experiment 1, with the exception that no stimuli appeared more often than once every 10 trials. As in Experiment, no probe was a recent negative.
Results and Discussion.

As in Experiment 1, for each subject, the overall mean RT for correct responses was calculated. Then, any RT that was more than three standard deviations above or below the mean were discarded and performance on that trial was counted as incorrect. A total of 77 trials were thus discarded, a mean of 2.41 trials discarded per subject. As in Experiment 1, all such trials had RTs greater than 3 standard deviations above the mean, and were evenly distributed over types of trials.

INSERT FIGURE 2 ABOUT HERE

Figure 2 shows the mean response time for trials 1 through 4 as a function of stimulus type (numbers or symbols) and test type (probe present or probe absent). The pattern of results seen with both types of stimuli was similar to that observed in Experiment 1: There was slower responding on Trial 4 than on Trial 1, consistent with the idea that PI builds up over trials.

The RT data were analyzed using a two stimulus types (numbers vs. symbols) × two test types (probe present vs. probe absent) × 4 trial numbers repeated measures analysis of variance. There was a significant main effect of stimulus type \( (F(1,31) = 10.00, MSE = 11556, \text{partial } \eta^2 = 0.244) \), with faster overall responding to symbols than numbers (950 vs. 980 ms, respectively). The main effect of test type was also significant \( (F(1,31) = 7.37, MSE = 18638, \text{partial } \eta^2 = 0.192) \), with faster responding on probe present than on probe absent trials (949 vs. 982 ms, respectively).

The main effect of trial number was significant \( (F(3,93) = 4.97, MSE = 12631, \text{partial } \eta^2 = 0.138) \). In general, RTs were faster on earlier trial numbers and slower on later trial numbers.
The mean RTs for trials 1 to 4 were 938, 962, 970, and 992 ms, respectively. A contrast revealed a significant linear trend for trial number \((F(1,31) = 10.67, \text{MSE} = 17117, \text{partial } \eta^2 = 0.256)\), but no quadratic component \((F(1,31) < 1)\).

No interactions approached significance; indeed, for all, \(F<1\).

The accuracy data were also analyzed using a two stimulus types (numbers vs. symbols) \(\times\) two test types (probe present vs. probe absent) \(\times\) 4 trial numbers repeated measures analysis of variance. There was a significant main effect of stimulus type \((F(1,31) = 12.75, \text{MSE} = 0.008, \text{partial } \eta^2 = 0.291)\), with the overall proportion of correct responses 0.913 for numbers compared to 0.942 for the symbols. As in Experiment 1, there was no indication of a speed-accuracy tradeoff, as the symbols were responded to both more quickly and more accurately than the numbers. Neither the main effect of test type \((F(1,31) = 1.02, \text{MSE} = 0.018, \text{partial } \eta^2 = 0.032)\) nor the main effect of trial number \((F(3,93) = 1.70, \text{MSE} = 0.008, \text{partial } \eta^2 = 0.052)\) were significant.

Only one interaction was significant: The interaction between stimulus type and trial number \((F(3,93) = 4.51, \text{MSE} = 0.009, \text{partial } \eta^2 = 0.127)\). There was a smaller range in response accuracy over trials for symbols (max. = 0.948; min. = 0.928) than for numbers (max. = 0.941; min. = 0.878). All remaining interactions had \(F < 1\), except for the interaction between stimulus and test type \((F(1,31) = 1.93, \text{MSE} = 0.012, \text{partial } \eta^2 = 0.059)\).

As in Experiment 1, the RT data were also analyzed to examine release from PI. For each subject, the RT on Trial 4 of a block was compared to the RT on the immediately following Trial 1 (given correct responses to both). Response times were significantly slower on Trial 4 than on Trial 1 (990 ms vs. 935 ms, respectively; \(F(1,31) = 11.29, \text{MSE} = 4255, \text{partial } \eta^2 = 0.267\)). In addition, a difference score was created for each Trial 4 - Trial 1 sequence in which both trials
had a correct response. Then, the mean of these difference scores was computed. Twenty five subjects had a positive score, indicating faster responding on Trial 1 than Trial 4, and seven had a negative score, indicated slower responding on Trial 1 than Trial 4. This is significant by a sign test.

The change from nonwords to symbols did not affect the pattern of results: for two-digit numbers, nonwords, and symbols, performance was faster on Trial 1 than on Trial 4, a pattern consistent with the idea that PI builds up over trials. The results are consistent with the idea that PI builds up over trials as the same type of material is experienced again and again, and that there is release from PI when the stimulus type changes.

General Discussion

Experiments 1 and 2 replicate the finding of Hanley and Scheirer (1975) by showing that PI can be observed when the subjects are asked to remember only three items, one less than the estimated capacity of the focus of attention. As predicted, there was a significant increase in RT from Trial 1 to Trial 4 for numbers, nonwords, and symbols, indicative of a build-up of PI over trials, and then a significant decrease from Trial 4 to the next Trial 1, indicative of a release from PI. Most importantly, PI was observed on both positive and negative trials, and did not require presentation of a recent negative.

The results are inconsistent with Cowan’s (2000, p. 103) prediction that one “can observe proactive interference (PI) in retrieval only if there are more than four items in a list to be retained.” On each list, there were only three items that needed to be remembered. Nonetheless, the more preceding lists there were with the same type of stimulus item, the slower and less

---

1 We note that others have suggested that the capacity of the focus of attention may be less than 4, and even as small as 1 item (e.g., McElree, 2001; Verhaeghen & Basak, 2005).
accurate the responding. These results are important, as this basic paradigm has long been used as evidence that items in the focus of attention are immune to PI.

The major difference between our design and others is that the processing requirements were such that conditions for observing PI were more favorable. In the Halford et al. (1988) study, the key manipulation to induce PI involved semantic processing. For four-item lists, such as pine, elm, oak, palm, it is possible that simple maintenance rehearsal sufficed for the task. In contrast, the current studies did not rely on subjects using semantic processing. Rather, the to-be-remembered items were more similar to each other than the examples from taxonomic categories used earlier.

Although the particular paradigm used in Experiments 1 and 2 has been the focus of the debate, converging evidence for the claim that PI can be observed when items are presumed to be in the focus of attention has been found in other paradigms. For example, Brannelly, Tehan and Humphreys (1989) used set sizes of two or four items in a probe recognition task where two probes were presented on each trial, the first after a filled retention interval and the second immediately after the participant responded to the first probe. The assumption was that the first probe would show the effects of retrieval and of PI, but the second, because the list items were still in the focus of attention, would not show retrieval or PI effects. For current purposes the critical findings were that RTs to the second probe were very much faster than to the first probe, consistent with the notion that the list items were in the focus of attention, but PI effects were present for both positive and negative probes for both set sizes of two and four words.

As a second example, consider some results from a very different paradigm, immediate serial recall. Baddeley, Lewis, and Vallar (1984, Experiment 3) had subjects listen to lists of 2 to 7 items that were either similar-sounding or dissimilar sounding. Although no errors were made
with list lengths of 2, the data when the list length is 3 is suggestive of an effect of PI: there were more errors for 3-item lists of similar-sounding letters than dissimilar-sound letters in the slow condition (1 item every 1.5 s) but not in the fast condition. Because Baddeley et al. did not report their results as a function of serial position, attribution of errors to PI is not certain. However, more recent data is consistent with this account. Roodenrys and Miller (2008) found better recall of a 3-item list of dissimilar sounding items than similar sounding items. Most importantly, this difference was observable at the final position, where only PI could be playing a role.

The current results, as well as those of Brannelly et al. (1989) and others, are more in line with a view of memory that argues that memory principles are general and apply regardless of the hypothetical underlying memory system (Surprenant & Neath, 2009a). According to this latter view, all memory is cue driven, and one consequence of this is that all memory is potentially vulnerable to PI. PI need not always be observed; it is easy to set up an experiment in which subjects are induced to process items in such a way that, at test, interference is minimized. However, PI should be observed in all types of tasks when the experiment is set up appropriately.

According to this view, effects observed in other memory systems or at longer time scales should be seen in short-term and working memory, even when the to-be-remembered items are assumed to be in the focus of attention. We mention just three such examples. First, effects of redintegration can be observed with short lists (Roodenrys & Hinton, 2002). Because redintegration can occur only when retrieval occurs, this again is inconsistent with the idea of the focus of attention. Second, Tehan and Humphreys (1996) demonstrated cueing effects with similarly short lists. Had the items been in the focus of attention, the cues should not have had an effect. Finally, Coane, McBride, Raulerson and Jordon (2007) found evidence of false memory in the focus of attention. They used the Deese-Roediger-McDermott procedure, but with only
three items in the list, and showed that response times to process critical lures were significantly slower than response times to process unrelated items (see also Atkins & Reuter-Lorenz, 2008).

The present studies add to this very small literature that investigates whether certain effects found in what is generally accepted to be long-term memory can be observed in short-term memory paradigms and with very short lists. One possible interpretation of these studies is that there is no fundamental difference in the rules that operate in the short and long term and that there is no reason to suppose they are separate systems (Surprenant & Neath, 2009b).
References


Proactive Interference


Author Note

This work was part of an undergraduate honours thesis by the first author. AJ was supported by an NSERC Alexander Graham Bell Canada Graduate Scholarship; AMP was supported by an NSERC Undergraduate Student Research Award; IN and AMS by NSERC Discovery Grants, and GT was partially supported by a Newfoundland and Labrador Centre for Applied Health Research visiting lecturer award. AMP is now at Lakehead University. The last three authors are listed alphabetically. Correspondence can be sent to Ian Neath (ineth@mun.ca), Aimée M. Surprenant (asurpren@mun.ca), or Gerry Tehan (Gerry.Tehan@usq.edu.au).
Endnote

1. We presented all stimuli and collected responses on an Apple iMac computer. As there is no published documentation of the accuracy of this platform when measuring RTs, we built a dedicated hardware timer to assess the accuracy and variability in RT measurements. A forthcoming paper (Neath, Earle, Hallett, & Surprenant, 2010) will detail the results, but according to our results, this system adds approximately 50 ms to the true RT, and the variability is less than ± 4 ms. As this is distributed evenly over conditions, it should not affect any conclusions drawn about differences between conditions.
Appendix

Experiment 1:

Numbers

23 24 25 26 27 28 29 32 34 35 36 37 38 39 42 43
45 46 47 48 49 52 53 54 56 57 58 59 62 63 64 65
67 68 69 72 73 74 75 76 78 79 82 83 84 85 86 87
89 92 93 94 95 96 97 98

Non-words

bame bave birm blek dant dess dirm dulb fant fibe fove frim gant glep glis gret
habe hane hice huse jace jant jirm jonk kark kass kerm kout lant laub lawp loat
mafe mout mulb mume nabe nant nesk nume pafe pelk pime plue rabe rell risp ronk
sant slen snen swip tafe tirm tomp trup

Experiment 2:

Numbers

23 24 25 26
32 34 35 36
42 43 45 46
52 53 54 56
62 63 64 65

Symbols

! @ # $ % ^ & * { + = / ] ? | –
Figure Captions

Figure 1: Mean response time to correctly judge whether a probe was in the most recent list in Experiment 1. Proactive interference builds up across trials for both numbers and nonwords. Note: P = probe present, A = probe absent.

Figure 2: Mean response time to correctly judge whether a probe was in the immediately preceding list for Experiment 2. Proactive interference builds up over trials for both numbers and symbols. Note: P = probe present, A = probe absent.
Figure 2

The graph shows the relationship between Mean RT (ms) and Trial Number for different conditions labeled as Num P, Num A, Sym P, and Sym A. The data points indicate a trend of increasing Mean RT with increasing Trial Number for all conditions.