Journal of Experimental Psychology:
Learning, Memory, and Cognition
2013, Vol. 39, No. 6, 1742–1756

© 2013 American Psychological Association
0278-7393/13/$12.00 DOI: 10.1037/a0033743

Focusing the Search: Proactive and Retroactive Interference and the Dynamics of Free Recall

Nash Unsworth
University of Oregon

Gene A. Brewer
Arizona State University

Gregory J. Spillers
University of Georgia

Targeting information in long-term memory is an important cognitive ability, but one that is not well understood. In this study, 4 experiments were conducted to examine the influence of proactive and retroactive interference on memory targeting. Participants were given either 1 or 2 lists and asked to recall List 1, List 2, or in some cases both lists. Multiple dependent measures were explored including the proportion of items recalled, number of intrusions output, and recall latency to arbitrate between 4 extant accounts of memory targeting. In general, recalling either List 1 or List 2 resulted in lower probability of recall, recall of more intrusions, and longer recall latencies compared to when recalling a list alone, suggesting both proactive and retroactive interference. These results suggest that long-term memory targeting is guided by noisy temporal-contextual cues (unless other salient cues are present) that activate both relevant and irrelevant memoranda that are then subjected to a postretrieval monitoring process.

Keywords: memory search, recall dynamics, interference

Our ability to selectively recall information from the recent past is an important feature of our long-term memory system. Research suggests that in many situations individuals are quite apt at selectively targeting items in memory, leading to their eventual recall. Despite evidence for this ability to selectively target information in memory, little is still known about how we actually accomplish this difficult task. Our goal in the present study was to better examine situations in which participants are asked to target information in memory in the presence of interfering information in the hopes of better elucidating how participants focus their search of memory to the desired information.

Focusing the Search Set in Memory

Free recall, wherein participants are presented with a list of items and are asked to recall the items in any order they wish, is one of the oldest and most heavily studied tasks in memory research (Crowder, 1976; Murdock, 1974; Tulving, 1968). An important result from studies of free recall is the finding that individuals are quite apt at selectively targeting items from the most recently presented list and rarely recall items from previous lists. This finding, among others, has resulted in many memory models assuming that context (in particular temporal context) plays a large role in allowing the memory system to selectively focus the search such that only a subset of relevant items (the search set) are activated (Anderson & Bower, 1972; Howard & Kahana, 2002; Mensink & Raaijmakers, 1988; Raaijmakers & Shiffrin, 1980). In such models it is assumed that the search set is determined, in part, by the match between context stored in the items and context present during retrieval such that the greater the overlap between the two, the more likely an item is to be included in the search set and subsequently recalled (e.g., Mensink & Raaijmakers, 1988). These models nicely account for the fact that most items that are recalled are from the most recently presented list with only a few intrusions from prior lists.

One issue with these models, however, occurs when the current context does not match context stored with the desired information. For instance, if asked “What did you have for dinner last Thursday?” how would you go about retrieving the answer? It is unlikely that your current context would provide much of a match to the context for the Thursday night in question, and thus it should be very difficult to retrieve the desired information if utilizing only the current context. Rather, it has been suggested that we must somehow attempt to reinstate the prior context to retrieve the desired information. That is, rather than use the current context as a probe, we must attempt to reconstruct the context for the time in question and use this information as a probe. Work by Shiffrin (1970a) has suggested that context reconstruction is possible when asked to recall information not from the current list, but from the list just prior to the current list. Specifically, Shiffrin devised the list-before-last recall task in which participants are asked not to

---

Nash Unsworth, Department of Psychology, University of Oregon; Gene A. Brewer, Department of Psychology, Arizona State University; Gregory J. Spillers, School of Law, University of Georgia.

We thank Henry Roediger and Geoff Ward for comments on a previous version of this article.

Correspondence concerning this article should be addressed to Nash Unsworth, Department of Psychology, University of Oregon, Eugene, OR 97403. E-mail: nashu@uoregon.edu

recall the current list of items, but rather to recall the prior list of items. That is, participants are given a list of items, followed by another list of items, and at recall participants are asked to recall the prior list, with the prior and current lists changing with each new list that is presented. Shiffrin also manipulated both the length of the target list and the length of the intervening list such that sometimes the lists were composed of five items and sometimes they were composed of 20 items. Shiffrin reasoned that if participants could not isolate the target list, then the size of the intervening list should matter, but if participants could fully isolate the target list, then the size of the intervening list would have no effect on the probability of recall. In three experiments Shiffrin found that the length of the target list mattered, but the length of the intervening list had no effect. Shiffrin concluded that participants could focus their search exclusively on the target list, and thus only the size of target list determined the likelihood of recall. Subsequent studies have largely corroborated Shiffrin's findings (Jang & Huber, 2008; Ward & Tan, 2004; but see Smith, 1979). Specifically, when participants recall between lists, the size of the intervening list has no effect on recall probabilities, suggesting that participants are able to isolate the target list. Recent work in our laboratory, however, has found the situation to be bit more nuanced. In fact, the mere presence of an intervening list can result in lowered recall probabilities, more intrusions, and longer recall latencies compared to control conditions with no intervening lists (Unsworth, Spillers, & Brewer, 2012). We suggested that participants could reconstruct the context of the prior list for the most part, but that items from the intervening list were also included, leading to retroactive interference effects.

Additional evidence for the notion that intervening items can influence recall comes from studies of the list-before task where there is no recall between lists. For example, Ward and Tan (2004) found an effect of intervening list length when participants were given two lists during encoding and then were cued after list presentation to recall either the current list or the prior list. Specifically, participants tended to recall more words when the intervening list was short compared to when the intervening list was long. Furthermore, in a control condition where participants were given only a single list and asked to recall it, Ward and Tan found that performance was much better than when an intervening list was presented, suggesting that retroactive interference played a role in recalling the prior list. Recently Jang and Huber (2008) replicated these effects, suggesting that when there is no recall between lists, the size of intervening list did have an effect. To account for these results, Jang and Huber suggested that on some retrieval attempts participants can correctly reconstruct the target list context, leading to an effect of target list length, but no effect of intervening list length. On other retrieval attempts, participants cannot effectively reconstruct the target list context and thus rely on context present during recall, which activates the intervening list to a greater extent than the target lists. This suggests that participants can effectively retrieve from only the target lists some of the time, and this is especially true when recall occurs between lists, suggesting that a recall test can lead to a release from interference (e.g., Szpunar, McDermott, & Roediger, 2008). However, when there is no recall between the lists, participants must rely on context present during the recall period as a cue, leading to more interference from the intervening list and the recall of intrusions from the intervening list (e.g., Jang & Huber, 2008; Sahakyan & Hendricks, 2012; Smith, 1979).

Work by Epstein and colleagues complicates this overall picture, however. Specifically, Epstein (1969, 1970) presented participants with two lists of items, and at test participants were cued to recall List 1, List 2, or both lists. Epstein found that recall was superior when participants had to recall only one list compared to both lists. Epstein referred to this as the "only effect" and suggested that it was due to selective search processes working at retrieval, which allowed participants to select only one list at a time while excluding the other list even with no recall between lists. Similar to the work reviewed previously, this suggests that participants can isolate one list, leading to little or no interference from the other list (see Epstein, 1972, for a review). However, Epstein and colleagues always compared either List 1 or List 2 recall to recall of both lists, but did not compare recall to a condition in which only a single list was presented. It is possible that when either List 1 or List 2 recall is compared to single list recall, there will be some interference from the other list as suggested by Unsworth et al. (2012).

Dynamics of Free Recall

The work reviewed thus far has focused almost exclusively on probability of recall. However, an examination of recall latency can also be informative in terms of better understanding how participants isolate target items in free-recall tasks. *Recall latency* refers to the time point during the recall period when any given item is recalled, and mean recall latency is simply the average time it takes to recall items. For instance, if items are recalled 5 s, 10 s, and 15 s into the recall period, mean recall latency would be 10 s. Prior work (Bousfield & Sedgewick, 1944; Indow & Togano, 1970; McGill, 1963; Roediger, Stellon, & Tulving, 1977; Rohrer & Wixted, 1994) has suggested that cumulative recall curves are well described by a cumulative exponential

$$F(t) = N(1 - e^{-\lambda t}), \tag{1}$$

where $F(t)$ represents the cumulative number of items recalled by time t, N represents asymptotic recall, and λ represents the rate of approach to asymptote. If given enough time to recall, N should equal the number of items recalled. However, these items can be recalled either quickly or slowly, and this information is captured by λ. When items are recalled quickly during the recall period, λ is relatively large, whereas when items are recalled slowly during the recall period, λ is relatively small.

Overall recall latency distributions are consistent with search models of free recall (Rohrer, 1996; Shiffrin, 1970b). In these models it is assumed that during recall a retrieval cue activates a subset of representations in memory that are related to the cue in some fashion. This delimited subset is known as the search set, and during recall, item representations are sampled (with replacement) from the search set (Raaijmakers & Shiffrin, 1980; Rohrer, 1996; Shiffrin, 1970b). After an item has been sampled, it must then be recovered. Items whose strength exceeds some critical threshold will be recovered and can be recalled, whereas weak items that do not exceed the threshold will not be recovered (Rohrer, 1996). Finally, after an item has been recovered, it is subjected to a monitoring and editing process that attempts to determine whether the item is correct and recalled, or incorrect and not recalled.

According to search models of this type, *N* reflects the number of target items in the search set whose absolute strength exceeds some threshold (e.g., Rohrer, 1996). Recall latency, and λ, reflect the number of items within the search. Thus, the larger the search set, the longer on average it will take to recall any given item. Important evidence for this type of model, as well as for a distinction by *N* and λ, comes from a number of studies that have manipulated aspects of free recall and found that some variables affect *N* but have no effect on λ, whereas other variables seem to primarily affect λ. For example, manipulating presentation duration increases *N* but has no effect on λ, whereas increasing list length leads to increases in both *N* and λ (Rohrer & Wixted, 1994). Additionally, Wixted and Rohrer (1993) examined the build and release of proactive interference and found that as proactive interference increased and probability of recall subsequently decreased, overall recall latency increased. Similar to the list length effects, this is presumably because as proactive interference built up, more items were included in the search set. Thus, although *N* decreased, this was due to a change in size of the search set rather than strength of items, given that the search set was likely composed of both target items and intrusions from prior lists (see also Unsworth, 2009). This work suggests that recall latency provides an index of overall search set size (Shiffrin, 1970b).

The Present Study

The goal of the present study was to examine how individuals focus their search on a target list that is either preceded by (proactive interference) or followed by (retroactive interference) another list. As with Ward and Tan's (2004) Experiment 3, we presented participants with one or two lists of words. Following a brief distractor task, participants were instructed to recall either List 1 or List 2. Thus, when presented with two lists, participants were required to target certain items in the presence of an irrelevant list that came either before or after the target list. When presented with only a single list, participants should be able to focus their search primarily on the target list (although there will be some interference from prior lists).

The main question in the current study was, how do we focus our search on target information? Specifically, if presented with two lists of items, can we focus on only one, or will we focus on both? Prior work suggests that it may be possible to focus on one list (Epstein, 1972; Shiffrin, 1970a), whereas other work suggests it is difficult to focus only on one list (Roediger & Tulving, 1979; Unsworth et al., 2012; Ward & Tan, 2004). In the present study we were interested in examining four possible explanations of how individuals search for information in the presence of either proactive or retroactive interference. In each case, predictions for overall probability of recall, recall latency, and intrusions will be given in order to examine which possibility provides the best account of the data. Importantly, as will be seen, each possibility predicts a different pattern of results in terms of the different recall measures. Thus, it is the overall pattern of results across measures, rather than any one measure (i.e., probability of recall), that distinguishes the different possibilities. Four experiments were conducted to examine these possibilities.

The first possibility, the Isolated Context hypothesis, suggests that at recall participants reconstruct or reinstate the target list context sufficiently such that the target list is isolated and probability of recall is driven by the target list with no interference from the other list (Shiffrin, 1970a; see also Klein, Shiffrin, & Criss, 2007). For example, Klein et al. (2007), in discussing Shiffrin's (1970a) results, suggested that "participants could reconstruct a context cue enabling them not only to access the older list, but also to focus sufficiently on that prior list to prevent interference by items on the intervening list" (p. 178). This possibility predicts that probability of recall and recall latency should be the same when one is required to recall either from two lists or from one list alone (i.e., control lists). That is, if there is no interference from the other list in the two list conditions and participants can isolate their search only to the target list, then probability of recall and recall latency should be the same when asked to recall from a target list in the presence of an irrelevant list as well as when asked to recall control lists. As noted previously, some of the prior work examining list-before-last recall has provided support for this possibility in that the size of the intervening list does not matter (at least when testing is required between lists), suggesting that participants can isolate the target list. Additionally, work by Epstein (1972) suggests that even with no recall between lists, it is possible to selectively search one list and exclude the other. Problematic for this possibility, however, is the finding that participants do emit intrusions from the intervening list, suggesting that the target list is not perfectly isolated (e.g., Jang & Huber, 2008; Smith, 1979; Unsworth et al., 2012).

The second possibility, the Recent Context hypothesis, suggests that participants rely on context present at recall to search for items (e.g., Ward & Tan, 2004). Items whose context matches context present at recall will receive the strongest activation and will be the most likely to be sampled and recalled. Thus, recently presented items will likely share the most context with the retrieval cue and should be the most likely to be recalled. When asked to recall List 1 in the presence of List 2, this suggests that in order to search for items from the List 1, participants would search back in time using the recall context as a cue and thus would activate items from the target list (List 1) as well as all items from the most recently presented intervening list (List 2). That is, the search set includes all of the target items as well as all of the intervening items. This predicts that when asked to recall List 1 items in the presence of List 2, there should be substantial retroactive interference, leading to a lower probability of recall, more intrusions (from List 2), and a longer recall latency, compared to control lists. Importantly, this possibility predicts that there should be more retroactive interference than proactive interference, given that List 2 shares more contextual features with the recall context than List 1. Thus, output from List 1 should be associated with lower probability of recall, recall of more intrusions, and longer recall latency than List 2 in this situation. Additionally, recall of List 2 should resemble recall of control lists. As noted previously, support for this view largely comes from studies in which there was no testing between lists (e.g., Jang & Huber, 2008; Ward & Tan, 2004; but see Smith, 1979).

An alternative, the Both Context hypothesis, is that there is not sufficient differentiation between List 1 and List 2, and thus the context at recall activates List 1 and List 2 items to the same extent. This possibility suggests that participants would search both lists at once, as one big list, and simply edit out intrusions. Thus, this view predicts that there should be both proactive and retroactive interference to the same extent, leading to lower probability of recall, recall of more intrusions, and longer recall latency, compared to control lists. This view is consistent with early search models of recall that suggested that in such situations proactive and retroactive interference effects would be the same. For example, Shiffrin (1970b) noted that

> in interference terms, the model proposed for free recall predicts both retroactive and proactive effects (nonspecific), and in fact assumes these to be equal. That is, any new item added to a list, whether prior to or subsequent to the item of interest simply has the effect of increasing the size of the search-set by some given amount; the increased size of the search-set causes reduced retrieval probability. (p. 437; see also Shiffrin & Atkinson, 1969)

The final possibility, the Noisy Context hypothesis, suggests that participants are generally able to reconstruct the target list context, but this reconstruction is noisy, leading to the inclusion of some intrusions. That is, this possibility suggests that participants can generally reconstruct the context for the target list, but given that there is some uncertainty about which items were actually presented on that list relative to other list items (intrusions), participants cast a wider net to ensure that the target information will be included in the search set. Thus, the reconstruction of context is noisy, leading to a slightly larger than normal (i.e., relative to control lists) search set that encompasses target items and some intrusions (Unsworth et al., 2012). This hypothesis is the same as the Both Context hypothesis in suggesting that both proactive and retroactive interference should occur and be of equal magnitude in terms of resulting in lower recall probabilities, recall of more intrusions, and longer recall latencies, compared to control lists, but these possibilities differ in the amount of interference that is predicted. Specifically, the Both Context hypothesis suggests that both List 1 and List 2 are included in the search set in their entirety, leading to one big list. The Noisy Context hypothesis, however, suggests that not all of the irrelevant list is included in the search set but only some irrelevant items are included (perhaps half). Thus, the Noisy Context hypothesis predicts that there should be both proactive and retroactive interference, but not as much as is predicted by the Both Context hypothesis.

Experiment 1

The purpose of Experiment 1 was to examine the four hypotheses of interest. Participants were presented with one or two lists of words and, following a brief distractor task, were instructed to recall either List 1 or List 2. Probability of recall, intrusions, and recall latency were examined.

Method

Participants and design. Participants were 26 undergraduate students recruited from the subject pool at the University of Georgia. Participants received course credit for their participation. Each participant was tested individually in a laboratory session lasting approximately 30 min. Words were nouns selected from the Toronto word pool (Friendly, Franklin, Hoffman, & Rubin, 1982). Words were initially randomized and placed into the lists, and all participants received the same lists of words. Twenty lists of 10 words each were created.

Procedure. Participants were tested individually. Participants received a total of 12 experimental trials. On four trials participants were presented with two lists of words and were instructed to recall List 1. On four trials participants were presented with two lists of words and were instructed to recall List 2. On the remaining four trials participants were presented with one list of words and were instructed to recall only that list. The different trial types were randomly mixed.

For each trial participants were told that they would be presented with either one or two lists of words and that following a brief distractor task they would be prompted to recall from one of the lists. They were instructed to read the words silently as they were presented and to recall the words in any order they wished during the recall period. Each trial began with a Ready signal onscreen for 3 s, followed by the statement *List 1* onscreen for 3 s, followed by a series of words presented one at a time in the center of the screen for 1 s each, with a 1-s blank screen in between the presentation of each word. After presentation of the first list, participants either began the distractor task (in the control lists) or were presented with the statement *List 2* onscreen for 3 s, followed by a series of words presented one at a time in the center of the screen for 1 s each, with a 1-s blank screen in between the presentation of each word. Following the second list (or the first list in the control conditions), participants engaged in a 16-s distractor task before recall: Participants saw eight three-digit numbers appear for 2 s each and were required to write the digits in descending order (e.g., Rohrer & Wixted, 1994; Unsworth, 2007). At recall, participants saw three question marks appear in the middle of the screen along with a message instructing them to recall either List 1 or List 2. Participants had 60 s to recall as many of the words as possible in any order they wished. Participants typed their responses and pressed Enter after each response, clearing the screen.

Results

Proportion recalled. As shown in Table 1, the proportion of items recalled was greater in the control lists compared to Lists 1 and 2, which had similar levels of proportion correct. These observations were supported by a 3 (list) within-subjects analysis of variance (ANOVA) on proportion correct, demonstrating a significant main effect of list, $F(2, 50) = 18.78$, $MSE = .01$, $p < .01$, $\eta_p^2 = .43$. Follow-up comparisons suggested that proportion correct was higher for the control lists than either List 1 or List 2 (both $ps < .01$), but there was no difference between List 1 and List 2 ($p > .31$). These results demonstrate that the presence of a preceding or subsequent list resulted in worse performance compared to control lists indicating both proactive and retroactive interference. Importantly, there was no difference in the magnitude of the interference effects.

Intrusions. Intrusions were classified as items that had been presented in the current task (either before or after the target list) but were not on the target list (i.e., intraexperimental intrusions). As shown in Table 1, there were more intrusions emitted per list during recall of List 1 and List 2 compared to the control lists, but the number of intrusions emitted per list was similar for List 1 and List 2. These observations were supported by a 3 (list) within-subjects ANOVA on the number of intrusions per list, demonstrating a significant main effect of list, $F(2, 50) = 28.41$, $MSE = .33$, $p < .01$, $\eta_p^2 = .53$. Follow-up comparisons suggested that fewer intrusions were emitted during recall of List C than either List 1 or

Table 1
Proportion Correct, Number of Intrusions per List, and Recall Latency for Each List Type in All Four Experiments

Measure	List 1	List 2	List C	List Both
		Experiment 1		
Proportion correct	0.36 (0.02)	0.33 (0.03)	0.48 (0.02)	
Number of intrusions	1.31 (0.17)	1.11 (0.14)	0.18 (0.05)	
Recall latency	18.51 (1.05)	17.69 (1.01)	14.82 (1.03)	
		Experiment 2		
Proportion correct	0.30 (0.03)	0.27 (0.02)	0.44 (0.02)	
Number of intrusions	2.76 (0.32)	2.66 (0.30)	1.20 (0.24)	
Recall latency	24.72 (0.99)	24.40 (0.87)	22.45 (0.99)	
		Experiment 3		
Proportion correct	0.33 (0.02)	0.31 (0.02)	0.43 (0.03)	0.26 (0.02)
Number of intrusions	1.59 (0.31)	1.59 (0.30)	0.36 (0.10)	0.55 (0.10)
Recall latency	17.51 (0.91)	17.56 (0.97)	14.88 (0.82)	20.12 (0.82)
		Experiment 4		
Proportion correct	0.61 (0.02)	0.61 (0.03)	0.62 (0.02)	
Number of intrusions	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	
Recall latency	16.47 (0.63)	16.04 (0.67)	16.37 (0.86)	

Note. Standard errors are shown in parentheses.

List 2 (both $ps < .01$), but there was no difference between List 1 and List 2 ($p > .27$). For List 1, 89% of intrusions came from the intervening list (List 2), whereas 2% came from the list from the immediately prior trial and the remaining 9% from two or more trials back. For List 2, 75% of intrusions came from the immediately preceding list (List 1), whereas the remaining 25% came from two or more trials back. Finally, for control lists, 47% of intrusions came from the immediately preceding trial, whereas the remaining 53% came from two or more trials back. Participants emitted more intrusions while trying to recall a list in the presence of an interfering list (List 1 or List 2) compared to control lists. Further, the majority of these intrusions came from the interfering list, suggesting the presence of both proactive and retroactive interference, which were similar in magnitude.

Recall latency. First, we examined cumulative recall functions for each of the conditions. Shown in Figure 1 are the cumulative recall functions for List 1, List 2, and the control lists. Consistent with previous research (Rohrer & Wixted, 1994; Wixted & Rohrer, 1993), the cumulative recall curve was well described by a cumulative exponential. As shown in Figure 1, the symbols represent the data, and the lines represent the best fitting cumulative exponential. As can be seen, fewer items were recalled from List 1 and List 2 compared to the control lists. Furthermore, rate of approach to asymptotic performance (λ) was slower for List 1 and List 2 compared to the control lists, but λ did not seem to differ between List 1 and List 2. Shown in Table 2 are the parameter estimates for each condition after fitting a cumulative exponential to the cumulative recall curves.[1]

In addition to examining the cumulative latency distributions and the parameter estimates from fitting the cumulative exponential, recall latency was directly computed. As shown in Table 1, recall latency was shorter for the control lists compared to List 1 and List 2, with List 1 and List 2 demonstrating similar recall latency values. These observations were supported by a 3 (list) within-subjects ANOVA on recall latency, demonstrating a significant main effect of list, $F(2, 50) = 6.26$, $MSE = 15470959$, $p < .01$, $\eta_p^2 = .20$. Follow-up comparisons suggested that recall latency was shorter for the control lists than either List 1 or List 2 (both $ps < .01$), but there was no difference between List 1 and List 2 ($p > .46$). Thus, recall latency was longer when trying to recall a target list associated with an irrelevant list compared to control lists. Increased recall latency associated with List 2 recall in the presence of List 1 is consistent with prior research, suggesting proactive interference effects on recall latency (Wixted & Rohrer, 1993; see also Unsworth, 2009). Novel to the current study is the finding of increased recall latency associated with List 1 recall in the presence of List 2, suggesting retroactive interference effects on recall latency (which were of similar magnitude to the proactive interference effects).

Discussion

The results from Experiment 1 were relatively straightforward, demonstrating clear interference effects in terms of both retroactive and proactive interference. Specifically, in comparison to control lists, List 1 and List 2 were associated with lower recall levels, higher rates of intrusions, and longer recall latencies, suggesting that the presence of a nontarget list presented before or after the target list interfered with the recall of target items. Furthermore, as suggested by Shiffrin (1970b), proactive and retroactive interference effects were of a similar magnitude across

---

[1] Note that although the cumulative exponential fit the data well, there are clear systematic deviations of fit. Specifically, the cumulative exponential tends to miss the early part of the curve, then slightly overestimates the data. As shown by Vorberg and Ulrich (1987), this pattern is expected when item strengths vary and recall is not entirely random. Furthermore, given that there is generally a few seconds' pause before the first item is output (Rohrer & Wixted, 1994), the fit will tend to be slightly overestimated early on. Despite these variations, the simple search model still provides a useful interpretation of the data.

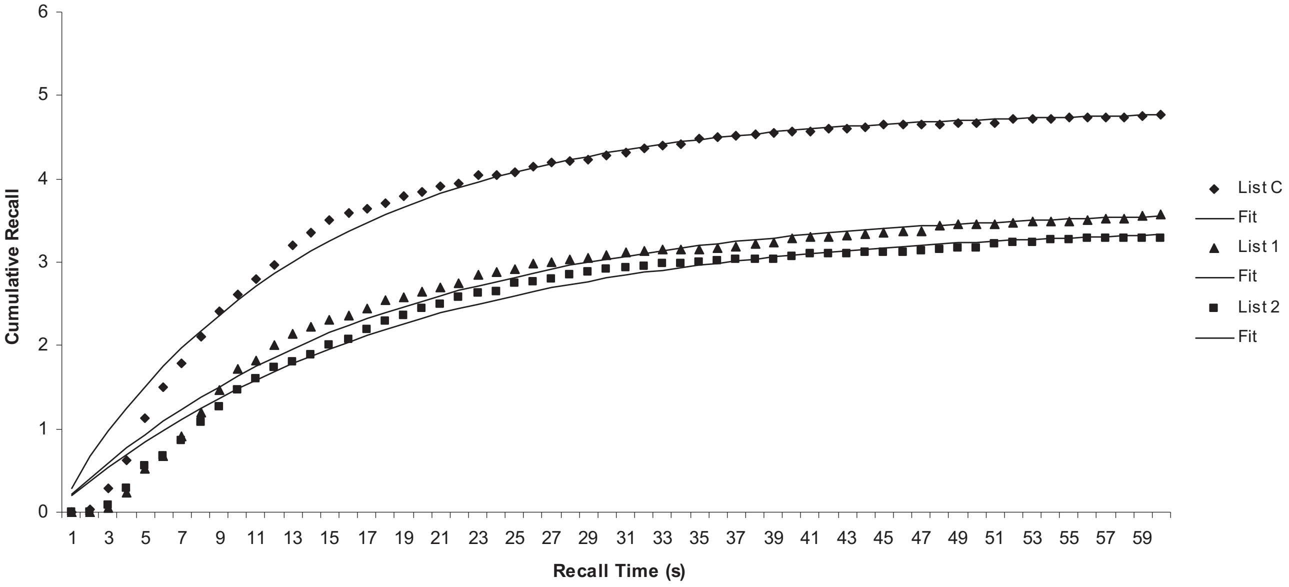

Figure 1. Cumulative recall curves as a function of recall time and list condition. List C = control lists for Experiment 1. Symbols represent the observed data, and the solid line represents the best fitting exponential.

all three measures of interest. These results are consistent with the Both Context hypothesis and Noisy Context hypothesis outlined in the introduction. Recall that these two suggest that there is insufficient differentiation between List 1 and List 2, leading the context cue used at recall to activate both List 1 and List 2 items to the same extent. This means that participants then search both lists at once, as one big list. However, if participants were searching both lists at the same time, one would expect more intrusions to be emitted than actually were. That is, participants emitted slightly more than one intrusion per list, but if both lists were active to the same extent, one would expect that there would be roughly equal numbers of correct items and intrusions emitted. It seems possible that participants are generating intrusions but recognize them as incorrect and are editing them out before recalling them. This possibility is tested in Experiment 2.

Table 2
Parameter Estimates Obtained From Fitting the Cumulative Recall Curves to a Cumulative Exponential for Each Experiment

List	λ	*N*	VAF
	Experiment 1		
List C	.075	4.82	.98
List 1	.059	3.66	.97
List 2	.056	3.45	.98
	Experiment 3		
List C	.077	4.36	.97
List 1	.055	3.71	.98
List 2	.057	3.12	.97
List Both	.046	5.62	.98
	Experiment 4		
List C	.058	6.43	.98
List 1	.059	6.44	.98
List 2	.062	6.37	.98

Note. λ = rate of approach to asymptotic performance; *N* = asymptotic performance; VAF = variance accounted for.

Experiment 2

The purpose of Experiment 2 was to examine the notion that participants are generating intrusions during the recall of List 1 or List 2, but are editing those intrusions out before they can be recalled. To examine this possibility, we had participants perform the same recall task as Experiment 1, but with a variant of externalized free recall (Kahana, Dolan, Sauder, & Wingfield, 2005; Unsworth, Brewer, & Spillers, 2010). In this task participants were instructed to recall all of the target items as in the Experiment 1, and participants were further instructed to recall any words that came to mind during the recall phase even if they knew that the word was not from the current list. Allowing participants to recall all items that come to mind serves to minimize the editing process by making recall uninhibited (Bousfield & Rosner, 1970; Kahana et al., 2005; Roediger & Payne, 1985; Unsworth et al., 2010). Furthermore, in order to examine monitoring processes, Kahana et al. (2005) instructed participants to press a key immediately after any response that the participant knew was incorrect. Thus, in this version of externalized free recall, participants are free to generate all items that come to mind (both correct and intrusions) and can indicate whether they identify the item as a correct or an intrusion. This should allow for a more fine-grained examination of the editing of intrusions. Additionally, given that the numbers of intrusions are likely to increase with this method, it should be possible to examine output dynamics for correct items and intrusions as a function of list to determine whether correct items and intrusions are equally accessible throughout the recall period or whether there are differences in accessibility. If participants are searching both lists as if it were one big list and search is relatively random, then one would expect equal numbers of correct responses and intrusions throughout the recall period. If, however, some items are more active than others (due to an overlap in contextual features associated with the items and contextual features associated with the context cue at retrieval) then one would expect some items to be more accessible early on and other items to be more accessible later on.

Method

Participants and design. Participants were 26 new undergraduate students recruited from the subject pool at the University of Georgia. Participants received course credit for their participation. Each participant was tested individually in a laboratory session lasting approximately 30 min.

Procedure. Participants performed the same recall task as in Experiment 1 with the exception that they were instructed not only to recall all of the items from the target list, but also to recall any other words that came to mind during the recall phase, even if they knew that the word was not presented on the target list. That is, participants were instructed to try to recall only items from the target list, but if other items came to mind, then they should recall those as well. Furthermore, the participants were instructed that if they recalled a word that they knew was incorrect, they should press the space bar to indicate that the response was incorrect.

Results

Proportion recalled. As shown in Table 1 and consistent with Experiment 1, the proportion of items recalled was greater in the control lists compared to Lists 1 and 2, which had similar levels of proportion correct. These observations were supported by a 3 (list) within-subjects ANOVA on proportion correct, demonstrating a significant main effect of list, $F(2, 50) = 21.23$, $MSE = .01$, $p < .01$, $\eta_p^2 = .46$. Follow-up comparisons suggested that that proportion correct was higher for the control lists than either List 1 or List 2 (both *p*s $< .01$), but there was no difference between List 1 and List 2 ($p > .38$).

Intrusions. As shown in Table 1, there were more intrusions from other lists (intraexperimental intrusions) emitted per list during recall of List 1 and List 2 compared to the control lists, but the number of intrusions emitted per list was similar for List 1 and List 2. Importantly, more intrusions were emitted in the current experiment than in Experiment 1. These observations were supported by a 3 (list) within-subjects ANOVA on the number of intrusions per list, demonstrating a significant main effect of list, $F(2, 50) = 26.59$, $MSE = .74$, $p < .01$, $\eta_p^2 = .52$. Follow-up comparisons suggested that fewer intrusions were emitted during recall of control list than either List 1 or List 2 (both *p*s $< .01$), but there was no difference between List 1 and List 2 ($p > .71$). For List 1, 68% of intrusions came from the intervening list (List 2), whereas 13% came from the list from the immediately prior trial, and the remaining 19% came from two or more trials back. For List 2, 65% of intrusions came from the immediately preceding list (List 1), whereas the remaining 35% came from two or more trials back. Finally, for control lists, 22% of intrusions came from the immediately preceding trial, whereas the remaining 78% came from two or more trials back.

Next we computed the percentage of times that participants correctly identified intrusions as errors for each list. For List 1 participants correctly rejected 63% ($SE = 0.06$) of intrusions, for List 2 participants correctly rejected 61% ($SE = 0.05$) of intrusions, and for control lists participants correctly rejected 85% ($SE = 0.05$) of intrusions. A within-subjects ANOVA on lists suggested a main effect of list, $F(2, 36) = 7.17$, $MSE = .05$, $p < .01$, $\eta_p^2 = .29$. Follow-up comparisons suggested that participants were better at rejecting intrusions on control lists compared to either List 1 or List 2 (both *p*s $< .01$), but there was no difference between List 1 and List 2 ($p > .77$). Additionally, note that correct rejection rates for both List 1 and List 2 were greater than chance (both *p*s $< .05$).

Output dynamics for corrects and intrusions. In our next set of analyses we examined output dynamics for correct responses and intrusions for List 1 and List 2 in order to gain a better understanding of when correct responses and intrusions are likely to be output during the recall period and whether this differs as a function of list. To examine this, we computed the proportion of responses that were either correct or intrusions for each output position and for each list. Shown in Figure 2 are the resulting output functions for List 1 (Figure 2A) and List 2 (Figure 2B). As can be seen, although equal numbers of correct responses and intrusions were recalled in List 1 and List 2, the two lists differed in when responses were output during the recall period. Specifically, when recalling List 1 items (in the presence of List 2), participants tended to start out by recalling correct items, but by the third output position there was no difference in recalling correct items or intrusions. When recalling List 2 items (in the presence of List 1), participants were more likely to begin recall with correct items and then, starting around Output Position 5, predominantly recalled intrusions. These observations were supported by a 2 (response type: correct vs. intrusion) × 2 (list) × 16 (output position) within-subjects ANOVA. The ANOVA demonstrated a main effect of output position, $F(15, 375) = 70.83$, $MSE = .06$, $p < .01$, $\eta_p^2 = .74$, suggesting that there were fewer responses at later output positions as well as a Response Type × Output Position interaction, $F(15, 375) = 5.84$, $MSE = .14$, $p < .01$, $\eta_p^2 = .19$, suggesting that the relative proportions of corrects and intrusions changed as a function of output position. Importantly, there was a significant Response Type × List × Output Position interaction, $F(15, 375) = 1.71$, $MSE = .12$, $p < .05$, $\eta_p^2 = .06$, suggesting that the Type × Output Position interactions were different for List 1 and List 2. That is, correct items and intrusions were recalled roughly equally across output positions (except for the first two output positions) when recalling from List 1, but when recalling from List 2, correct items were more likely to be recalled early on, whereas intrusions were more likely to be recalled at later output positions, suggesting potential differences between List 1 and List 2 in the accessibility of items during the recall period. An additional result worth mentioning is that the main effect of response type was not significant, suggesting that equal numbers of correct responses and intrusions were emitted. Specifically, for List 1, 2.99 correct responses were recalled along with 2.76 intrusions. For List 2, 2.74 correct responses were recalled along with 2.66 intrusions. Thus, there were equal numbers of correct items and intrusions recalled for both lists when using the externalized free-recall procedure.

Transition probabilities. Next, we examined transition probabilities between all of the item types to better determine the relations among the different item types in order to examine potential clustering of responses based on the previous recall of

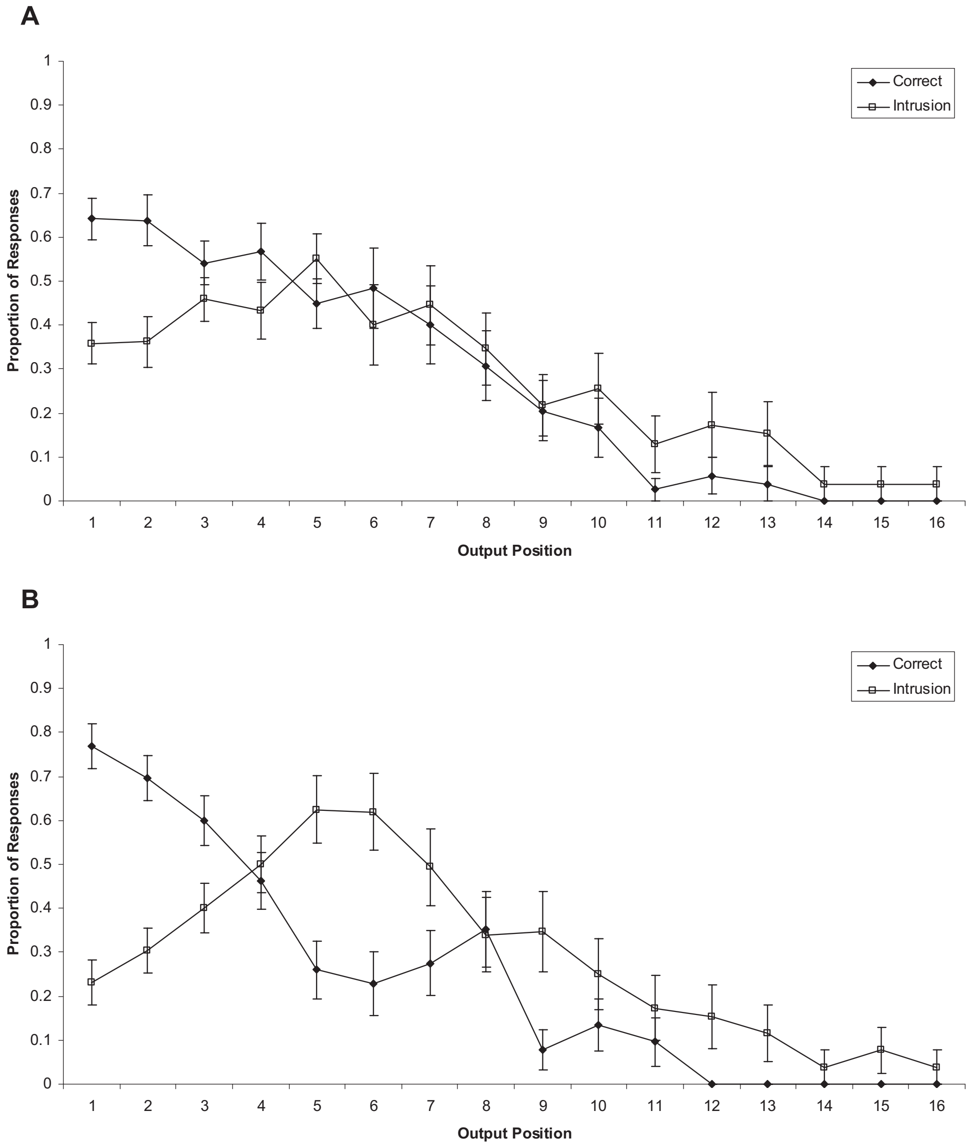

Figure 2. (A) Proportion of responses as a function of output position and response type (correct and intrusions) for List 1. (B) Proportion of responses as a function of output position and response type (correct and intrusions) for List 2. Error bars represent 1 standard error of the mean.

similar item types (see also Unsworth et al., 2010; Zaromb et al., 2006). Specifically, we computed transition probabilities for each possible transition for correct recalls, intraexperimental intrusions, and extraexperimental intrusions individually. We included extraexperimental intrusions (recall of an item that was not presented during the experiment) given that participants typically recall a large number of extraexperimental intrusions with this procedure (Kahana et al., 2005; Unsworth et al., 2010). Indeed, in the current study participants recalled on average 8.40 (SE = 1.59) extraexperimental intrusions per list. We computed the probability of recalling a correct item followed by another correct item, as well as the probabilities of recalling a correct item followed by an intraexperimental or followed by an extraexperimental intrusion. These transition probabilities were calculated separately for each response type. Shown in Table 3 are the resulting transition probabilities. As can be seen, when a participant recalled a correct item, the probability that the next item the participant recalled would be another correct item was very high. The probability of recalling another item type was much smaller. Note that the comparisons in Table 1 are only meaningful within a row and not across rows, given that the transitions within a row were divided by the same baseline but those in different rows were divided by different baselines. The transition probabilities within a row sum to 1.0. When recalling an intraexperimental intrusion, participants tended to recall another intraexperimental intrusion next. Likewise, when recalling extraexperimental intrusions, the item recalled tended to be an extraexperimental intrusion. Examining transition probabilities for each list type separately suggested overall similar results.

Table 3

Recall Transition Probabilities Between Correct Items, Intraexperimental Intrusions, and Extraexperimental Intrusions for All Items, for List 1 Items Only, and for List 2 Items Only in Experiment 2

Item	Correct	Intra	Extra
	All items		
Correct	.64	.19	.17
Intra	.15	.60	.25
Extra	.11	.21	.68
	List 1 only		
Correct	.69	.15	.16
Intra	.14	.61	.25
Extra	.11	.21	.68
	List 2 only		
Correct	.59	.22	.19
Intra	.17	.58	.25
Extra	.11	.21	.68

Note. Correct = correct responses; intra = intraexperimental intrusions; extra = extraexperimental intrusions.

The only difference seemed to be that participants were slightly more likely to transition between correct items and intraexperimental intrusions when recalling from List 2 compared to when recalling from List 1. This likely reflects the shift seen around Output Position 5 when participants shift from recalling predominantly correct items to recalling intraexperimental intrusions. Overall there was a clear pattern of clustering in the data, with similar item types being recalled in succession.

Recall latency. As shown in Table 2, recall latency was shorter for the control lists compared to List 1 and List 2, with List 1 and List 2 demonstrating similar recall latency values consistent with Experiment 1. These observations were supported by a 3 (list) within-subjects ANOVA on recall latency, demonstrating a significant main effect of list, $F(2, 50) = 3.32$, $MSE = 1073431$, $p < .05$, $\eta_p^2 = .12$. Follow-up comparisons suggested that that recall latency was shorter for the control lists than either List 1 or List (both $ps < .05$), but there was no difference between List 1 and List 2 ($p > .48$).

Discussion

Overall the results from Experiment 2 were consistent with the results from Experiment 1. Importantly, additional findings were gained by using the externalized free-recall technique in Experiment 2. Unlike in Experiment 1, which demonstrated that participants recalled more correct items than intrusions, using externalized free recall in Experiment 2 suggested that roughly equal numbers of correct items and intrusions were emitted on List 1 and List 2. Additionally, the results suggested that participants were better at rejecting intrusions on the control lists compared to either List 1 or List 2. Thus, using externalized free recall revealed that participants not only include many intrusions in their search sets for List 1 and List 2, but also have more difficulty monitoring these intrusions relative to lists that are not accompanied by additional material. There was also a clear pattern of clustering within the data in which participants recalled similar items in succession, suggesting that the prior item's retrieved context likely cued recall of successive items (Howard & Kahana, 2002). Finally, although the results for List 1 and List 2 were comparable in terms of overall levels of recall, number of intrusions, and recall latency, an examination of output dynamics suggested that there were some differences between List 1 and List 2. Specifically, when asked to recall List 1 in the presence of List 2, participants recalled roughly equal numbers of correct items and intrusions across output positions (except for the first two output positions), suggesting that intrusions were as accessible as correct items under conditions of retroactive interference. When asked to recall List 2 in the presence of List 1, however, correct items were more likely to be recalled early on, and intrusions were more likely to be emitted later in the recall period. This suggests that under conditions of proactive interference, correct items were more accessible early in the recall period than intrusions, likely due to a greater overlap in contextual features between correct items and context during retrieval. This is consistent with prior work on List 2 dominance effects in proactive and retroactive free-recall studies (e.g., Albert & Schulz, 1975). Thus, although proactive and retroactive interference effects were comparable, there were also important differences in terms of the accessibility of items.

Experiment 3

The results thus far suggest that when asked to recall a target list in the presence of an interfering list, participants demonstrate lower levels of recall, higher numbers of intrusions, and longer recall latencies compared to control lists. In addition, the results indicate that the magnitude of proactive and retroactive interference effects was roughly equal. As noted previously, these results are in line with the Both Context and Noisy Context hypotheses outlined in the introduction. Recall that these two hypotheses suggest that there is not sufficient differentiation between List 1 and List 2, leading to the context cue at recall activating both List 1 and List 2 items to the same extent. However, there is one important difference between these two possibilities. As outlined in the introduction, the Both Context hypothesis suggests that participants activate both List 1 and List 2 in their entirety, leading to one big list. The Noisy Context hypothesis suggests that not all of the irrelevant list is included in the search set but only some irrelevant items are included. Thus, these two hypotheses differ in the overall size of the search set. The Both Context hypothesis suggests that the search set includes 20 items (all 10 items from List 1 and all 10 items from List 2), whereas the Noisy Context hypothesis suggests that all of the target items are included in the search set but only some of the irrelevant items are included. To determine which of these two possibilities provides a more likely account of the results, we had participants perform the same recall task as in Experiment 1, in which they were required to recall List 1, List 2, or control lists. An additional condition was included in this experiment in which, after being presented with both List 1 and List 2 (and after completing the distractor task), participants were instructed to recall both lists in any order they wanted (e.g., Epstein, 1972). Thus, on some trials participants had to recall only List 1 items, List 2 items, or both List 1 and List 2 items. If the Both Context hypothesis is correct and participants are searching through one large list of 20 items when recalling from either List 1 or List 2, then recall latency should be the same as when they are asked to recall both List 1 and List 2. If the Noisy Context hypothesis is correct and participants include some of the irrelevant list in the search set, but not all irrelevant items, then recall latency for either List 1 or List 2 should be shorter than when asked to recall both lists, but longer than when asked to recall control lists.

Method

Participants and design. Participants were 26 undergraduate students recruited from the subject pool at the University of Oregon. Participants received course credit for their participation. Each participant was tested individually in a laboratory session lasting approximately 30 min.

Procedure. Participants performed a similar recall task as in Experiment 1. Participants received a total of 12 experimental trials. On three trials participants were presented with two lists of words and were instructed to recall List 1. On three trials participants were presented with two lists of words and were instructed to recall List 2. On three trials participants were presented with two lists of words and were instructed to recall both lists in any order they wanted. On the remaining three trials participants were presented with one list of words and were instructed to recall only that list. The different trial types were randomly mixed.

Results

Proportion recalled. As shown in Table 1, the proportion of items recalled was greater in the control lists compared to Lists 1 and 2 (which had similar levels of proportion correct) and compared to when participants had to recall both lists (which seemed to have lower levels of recall compared to all other lists). These observations were supported by a 4 (list) within-subjects ANOVA on proportion correct, demonstrating a significant main effect of list, $F(3, 75) = 18.45$, $MSE = .01$, $p < .01$, $\eta_p^2 = .43$. Follow-up comparisons suggested that proportion correct was highest for the control lists compared to all other lists (all $ps < .01$) and lowest for both lists (all $ps < .05$), but there was no difference between List 1 and List 2 ($p > .37$). Examining List 1 and List 2 items separately in the both-lists condition suggested a slightly higher proportion of items were recalled from List 1 ($M = .28$, $SE = .02$) than List 2 ($M = .23$, $SE = .02$) items ($p < .08$). However, a fewer proportion of items were recalled from List 1 when required to recall both lists than when to recall List 1 only ($p < .05$). The same was true for List 2 items ($p < .05$).

Intrusions. More intrusions emitted per list during recall of List 1 and List 2 compared to the control lists and both lists, but the number of intrusions emitted per list was similar for List 1 and List 2. These observations were supported by a 4 (list) within-subjects ANOVA on the number of intrusions per list, demonstrating a significant main effect of list, $F(3, 75) = 13.21$, $MSE = .86$, $p < .01$, $\eta_p^2 = .35$. Follow-up comparisons suggested that fewer intrusions were emitted during recall of the control lists than any of the other lists (all $ps < .01$), and more intrusions were emitted during recall of List 1 and List 2 compared to the other lists (all $ps < .01$), but there was no difference between List 1 and List 2 ($p > .99$). For List 1, 85% of intrusions came from the intervening list (List 2), whereas 2% came from the list from the immediately prior trial, and the remaining 13% came from two or more trials back. For List 2, 72% of intrusions came from the immediately preceding list (List 1) whereas the remaining 28% came from two or more trials back. For the control lists, 20% of intrusions came from the immediately preceding trial, whereas the remaining 80% came from two or more trials back. Finally, for both lists, 13% of intrusions came from the immediately preceding trial, whereas the remaining 87% came from two or more trials back.

Recall latency. Shown in Figure 3 are the cumulative recall functions for List 1, List 2, the control lists, and both lists. As can be seen, more items were recalled in both lists and the control lists than for either List 1 or List 2. Although of course a lower proportion of items was recalled in the both-lists condition. Furthermore, rate of approach to asymptotic performance (λ) was slowest for both lists, fastest for the control lists, and intermediate for List 1 and List 2, which did not seem to differ in λ. Shown in Table 2 are the parameter estimates for each condition after fitting a cumulative exponential to the cumulative recall curves.

As shown in Table 1, recall latency was shortest for the control lists, longest for both lists, with List 1 and List 2 demonstrating similar recall latency values between the control and both lists. These observations were supported by a 4 (list) within-subjects ANOVA on recall latency, demonstrating a significant main effect of list, $F(3, 75) = 9.18$, $MSE = 12943821$, $p < .01$, $\eta_p^2 = .27$. Follow-up comparisons suggested that recall latency was shortest for the control lists compared to all other lists (all $ps < .05$), longest for both lists (all $ps < .01$), with Lists 1 and 2 falling between the control and both lists and no difference between List 1 and List 2 ($p > .96$). Examining recall latencies for List 1 and List 2 items from the both-lists condition suggested no differences in recall latency for List 1 and List 2 items ($p > .15$), with both having longer recalling latencies than the other conditions (all $ps < .01$). These results demonstrate that recall latency was longer when trying to recall a target list associated with an irrelevant list compared to control lists, but recall latency was not as long as when trying to recall both lists at the same time.

Discussion

The results for Experiment 3 were consistent with those for Experiments 1 and 2, suggesting that List 1 and List 2 were associated with lower recall levels, higher rates of intrusions, and longer recall latencies than control lists. When asked to recall both lists, recall levels were lower and recall latency was longer than when asked to recall only List 1 or List 2. This replicates the “only effect” demonstrated by Epstein and colleagues (e.g., Epstein, 1972). Consistent with Epstein’s interpretation, this suggests that the overall size of the search set was larger when asked to recall both lists than when asked to recall either List 1 or List 2. Thus, these results suggest that when searching for either List 1 or List 2, participants do not simply include all items from both lists, creating one big list. Rather, it seems that some items from the irrelevant list are included in the search set, but not all of the irrelevant items are included. The fact that recall latency for List 1 and List 2 fell almost exactly in between the control lists and both lists suggests that the search sets for List 1 and List 2 are likely composed of roughly half of the irrelevant list, leading to an overall search set of roughly 15 items. Thus, although participants demonstrated some selectivity in their search of only List 1 or List

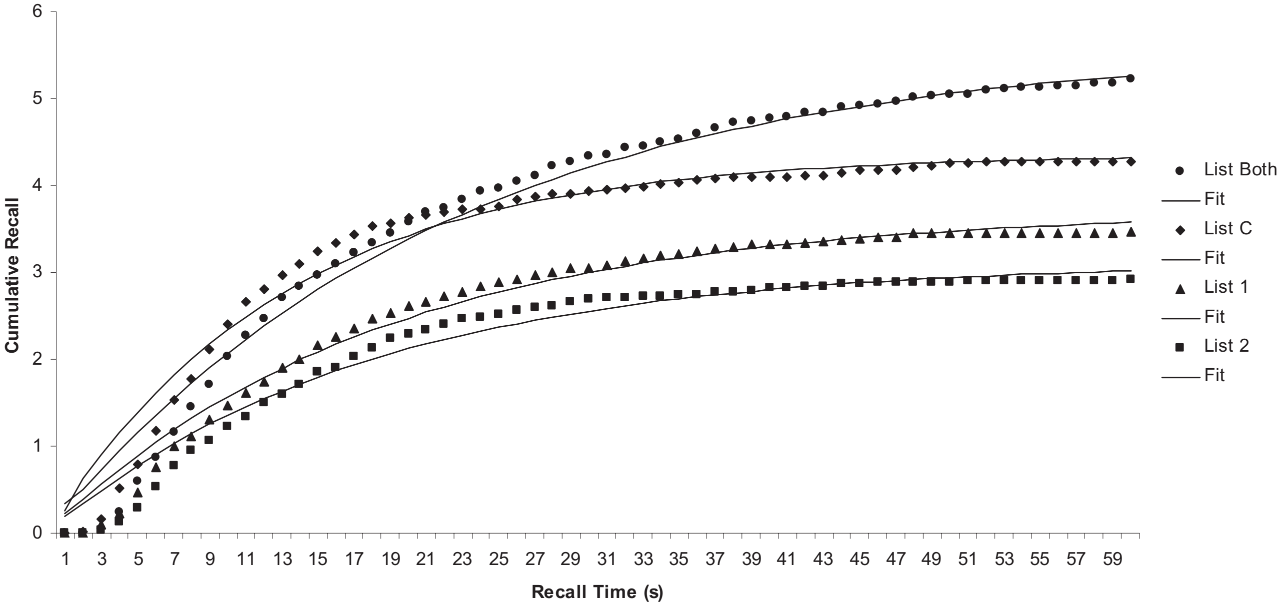

Figure 3. Cumulative recall curves as a function of recall time and list condition. List C = control lists for Experiment 3. Symbols represent the observed data, and the solid line represents the best fitting exponential.

2 (as suggested by Epstein, 1972), they could not fully isolate the target list, leading to some interference from the excluded list.

Experiment 4

The prior experiments suggested that participants have difficulties in focusing their search on only the target information and include items from other lists. The purpose of Experiment 4 was to see whether participants can adequately focus their search on target information when that information is made distinct relative to information from other lists. That is, if temporal context by itself cannot isolate one list over another (especially when the lists are presented in close temporal proximity), then perhaps other means can be used to differentiate the two lists, leading to the isolation of one list. Prior research in the context of proactive interference has found that changes in semantic categories between lists result in a release from proactive interference such that performance on the new category of items is similar to performance on the first list of the experiment (e.g., Wickens, 1972). Theoretically, the release from proactive interference is due to the fact that the new semantic category distinguishes target items on the current list from items on prior trials, leading to a focusing of the search set on only the target information (e.g., Gardiner, Craik, & Birtwistle, 1972). In line with this interpretation, Wixted and Rohrer (1993) found that changes in semantic category between lists reduced recall latency, suggesting a narrowing of the search set. In a similar vein, Marsh, Landau, and Hicks (1996) found that providing participants with a postinformation cue indicating differences between List 1 and List 2 resulted in enhanced recall of List 1 and in shorter recall latencies for List 1 compared to a group that was not given a postinformation cue. Thus, in the context of retroactive interference, this suggests that making the lists distinct should reduce retroactive interference and allow participants to focus on their search on the target list. To examine this, we had participants perform the same recall task as in Experiment 1, but each list was composed of a new semantic category, and participants were instructed to recall the target list and were provided with a category cue for that list.

Method

Participants and design. Participants were 27 new undergraduate students recruited from the subject pool at the University of Georgia. Participants received course credit for their participation. Each participant was tested individually in a laboratory session lasting approximately 30 min.

Procedure. Participants performed the same recall task as in Experiment 1 with the exception that each list was now composed of items from a single category unique to that list. Words and category labels were from Murdock (1976), excluding the six most frequent words for each category. Categories included body parts, male names, sports, types of boats, spices, kitchen utensils, vegetables, birds, fields of study, articles of clothing, chemicals, weather, flowers, insects, fabrics, tools, animals, colors, musical instruments, and trees. At recall, participants saw three question marks appear in the middle of the screen along with a message instructing them to recall either List 1 or List 2 along with the appropriate category label for the target list.

Results

Proportion recalled. As shown in Table 1, there were virtually no differences between the lists in the proportion of correct items recalled. These observations were supported by a 3 (list) within-subjects ANOVA on proportion correct, demonstrating the lack of a significant main effect of list, $F(2, 52) = 0.007$, $MSE = .01$, $p < .99$, $\eta_p^2 = .00$.

Intrusions. As shown in Table 1, there were no intrusions emitted in any of the lists by any participant.

Recall latency. Shown in Figure 4 are the cumulative recall functions for List 1, List 2, and the control lists. As can be seen, there were no differences in the number of items recalled for each list, and the rate of approach to asymptotic performance (λ) was roughly equal across the different lists. Shown in Table 2 are the parameter estimates for each condition after fitting a cumulative exponential to the cumulative recall curves.

As shown in Table 1, recall latency was roughly equal for the different lists. These observations were supported by a 3 (list) within-subjects ANOVA on recall latency, demonstrating a lack of a significant main effect of list, $F(2, 52) = 0.53$, $MSE = 6328714$, $p > .59$, $\eta_p^2 = .01$.

Discussion

The results for Experiment 4 were very straightforward. Presenting participants with distinct lists of items from unique semantic categories resulted in similar levels of performance for all three lists in terms of proportion correct, intrusions, and recall latency. Thus, making each list distinct allowed participants to focus their search on the target information, effectively reducing both proactive and retroactive interference, leading to similar levels of recall, similar recall latencies, and the recall of zero intrusions across all lists and all participants.

General Discussion

In four experiments, using a two-list paradigm, we examined participants' ability to focus their search on target list items in the presence of irrelevant items that came either before (proactive interference) or after (retroactive interference) the target list. In Experiment 1, it was shown that recalling either List 1 or List 2 resulted in lower levels of recall, higher rates of intrusions, and longer recall latencies compared to when recalling a single control list. Importantly, these effects were equal for both List 1 and List 2. Using an externalized free-recall procedure, in Experiment 2 we replicated those results from Experiment 1 and also demonstrated that participants recalled the same number of correct items and intrusions when recalling List 1 or List 2, but not when recalling control lists. Interestingly, participants' monitoring abilities were worse when recalling from List 1 or List 2 compared to control lists, and participants were close to chance in recognizing intrusions as errors. Thus, not only were intrusions likely to be recalled, these intrusions were also very likely to be confused as correct items by the participants. Additionally, Experiment 2 demonstrated that although performance was similar when recalling from List 1 and List 2, differences did emerge in output dynamics such that correct items and intrusions were equally likely to be emitted when recalling from List 1, but that correct items were more likely to be recalled early on and intrusions later when recalling from List 2. This suggests that there were differences in recalling from List 1 and List 2 in terms of the accessibility of items. Finally, examining transition probabilities in Experiment 2 demonstrated that participants clustered responses based on item type, suggesting that the prior item's retrieved context was used as a cue for successive items (Howard & Kahana, 2002). In Experiment 3, where asked to recall both List 1 and List 2 on some trials, participants demonstrated that when recalling both lists, recall levels were lower and recall latency was longer than when recalling either List 1 or List 2 (the "only effect"), but performance was superior when given only one list, suggesting that participants could not fully isolate the target list. Finally, in Experiment 4, when presented with lists of category items unique to each list and provided with a postinformation cue (e.g., Marsh et al., 1996), participants eliminated differences between List 1 and List 2 with control lists, suggesting that both proactive and retroactive interference were eliminated.

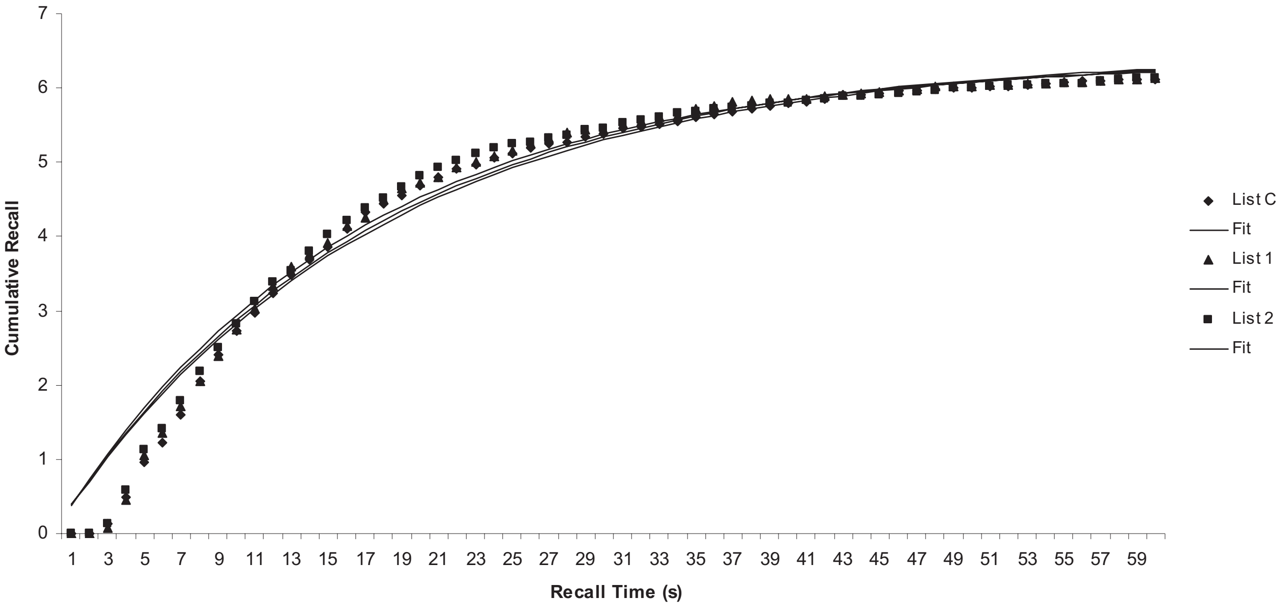

Figure 4. Cumulative recall curves as a function of recall time and list condition. List C = control lists for Experiment 4. Symbols represent the observed data, and the solid line represents the best fitting exponential.

Collectively, these results are most consistent with the Noisy Context hypothesis discussed in the introduction. Specifically, the Noisy Context hypothesis suggests that participants can generally reconstruct the context for the target list, but due to some uncertainty about which items were actually presented on target list relative to the irrelevant list, participants cast a wider net to ensure that the target information will be included in the search set. Thus, probability of correct recall is reduced when recalling from List 1 or List 2 because prior or intervening list items are included in the search set. Likewise, this view predicts that the search set will be centered on the target list, but items presented in close temporal proximity to the target list (either before or after the target list) will also be included in the search set. Importantly, not all intruding items will be in the search set, but rather only those items that share enough contextual features will be included. Thus, intrusions should be recalled from either the prior or intervening list. Finally, given that intrusions are included in the search set, this view predicts that recall latency should be longer when recalling from either List 1 or List 2 compared to the control lists. Overall, the current pattern is very much in line with predictions from this possibility.

When searching memory, individuals attempt to focus their search on target information by reconstructing the context associated with the target information. However, given this is a difficult process and that there is some uncertainty about what information is correct versus incorrect, individuals will rely on a somewhat noisy contextual reconstruction process in which the search set is centered on the target information but irrelevant information is also included in the search set. By casting a wider net and including irrelevant information in the search set, there is a greater likelihood of actually sampling the target information as well as a greater likelihood of sampling irrelevant information. Thus, there are both costs and benefits to casting a wider net. The overall effect of including irrelevant information into the search set is an increase in interference (both proactive and retroactive) that lowers recall probabilities, increases intrusion rates, and increases recall latency.[2] The current results suggest that it is possible to reconstruct the context of target information in order to focus the search on this information. At the same time, the current results suggest that this context reconstruction process is far from perfect in that irrelevant information will be included along with target information in the search set. Importantly, not all irrelevant items are included in the search set, thus allowing for a greater probability of recalling target items than would occur if individuals attempted to search both lists at the same time. Furthermore, the current results rule out additional possibilities about how individuals may focus their search by suggesting that participants do not exclusively rely on the context at recall as a cue, but rather attempt to reconstruct the context of the target list.

In general, the current results suggest that when presented with two lists of items in close temporal proximity, participants cannot focus exclusively on the target list. Prior results suggesting that participants can selectively search only one list (e.g., Epstein, 1972; Shiffrin, 1970a) are challenged by the fact that when recalling either List 1 or List 2, performance is worse compared to recalling only one list alone. That is, prior research has suggested that when participants recall between lists, as in the list-before-last paradigm, the act of recall leads to list isolation (e.g., Jang & Huber, 2008). However, recent work with the list-before-last paradigm suggests that the lists are not completely isolated, as there is some interference from the intervening list (e.g., Unsworth et al., 2012). Additionally, although there is an “only effect” such that performance is better when recalling either List 1 or List 2 compared to recalling both lists (Epstein, 1972), recalling only one list alone leads to the best performance, given that there is some interference in the either List 1 or List 2 condition. Again this suggests that participants cannot fully select only the target list, but rather some items from the inferring list are included in the search set. However, that is not to say that list isolation is impossible. Rather, when the two lists are made sufficiently distinct (e.g., by using different types of materials, spacing out the lists, or presenting the lists in different modalities), it should be possible to isolate the target list, eliminating interference effects. Indeed, the results from Experiment 4 suggested that when the two lists are composed of items from different categories and participants are provided with category cues at recall, it is possible to focus exclusively on the target list with no interference from the other list. This suggests that context alone does a poor job of isolating a list unless other means are used to make the lists distinct from one another.

The current results also suggest the importance of monitoring and editing processes inherent in recall. Specifically, the results suggest that although participants tended to emit more intrusions when recalling List 1 or List 2 compared to the control lists, the overall number of intrusions emitted was quite low, suggesting that participants likely generated intrusions but recognized these intrusions as such and edited them out before they were recalled. The results from Experiment 2 were consistent with this notion by demonstrating that with an externalized free-recall procedure, individuals emitted more intrusions than with a standard recall procedure. Furthermore, participants were generally good at recognizing the intrusions as errors. Thus, although participants tended to include intrusions in their search sets, they were quite good at recognizing those intrusions as incorrect and editing them out so that they were not emitted during the recall period.

These results have important implications for current theories of recall. In particular, most current theories of recall typically rely on the notion that when searching memory, individuals recall items based on the match between context stored in the items and the context present during retrieval, such that the greater the overlap between the two, the more likely an item has in being included in the search set and subsequently recalled (e.g., Howard & Kahana, 2002; Mensink & Raaijmakers, 1988; Raaijmakers & Shiffrin, 1980). The notion that context at retrieval acts as the primary cue and dictates the extent to which relevant items will be recalled works quite well in standard recall paradigms where participants are asked to recall the most recently presented list. However, these models would seem to be less capable of accounting for patterns of recall in paradigms where one not only has to recall the most recently presented list, but also has to recall items presented on

---

[2] In addition to proactive and retroactive interference, there seemed to be evidence for output interference in the data. Within each experiment there were roughly the same total number of words recalled (correct plus intrusions) across the different conditions. This suggests the possibility of output interference influencing the rate of retrieval, with any word retrieved in response to the same cue effectively slowing down the retrieval of as-yet-unrecalled words. We thank Geoff Ward for noticing and suggesting this.

prior lists, as in the current experiments and in the list-before-last recall task (Shiffrin, 1970a). In particular, the current results along with prior work (e.g., Jang & Huber, 2008; Sahakyan & Hendricks, 2012; Shiffrin, 1970a; Smith, 1979; Unsworth et al., 2012) suggest that these models need to be augmented with some form of contextual reconstruction process that allows participants to attempt to reconstruct the context of the target list rather than merely rely on context at retrieval as a cue. Furthermore, the current work suggests that this process will need to be somewhat noisy in that perfect reconstruction is unlikely (unless the lists are sufficiently distinct and there are sufficient retrieval cues present), leading to the inclusion of some intrusions in the search set. Finally, the current results suggest the need for some sort of monitoring and editing process (see Sirotin, Kimball, & Kahana, 2005, for a recent example) in these types of models in order to account for the fact that participants can be quite good at recognizing intrusions as such and that they actively edit these intrusions, thus preventing them from being emitted. Future work is needed to better understand the situations in which the search set can and cannot be focused on only target information, the processes that are utilized in order to reconstruct context for events without relying exclusively on the present context, and how participants use context to monitor the products of retrieval and how this influences editing decisions during recall. For now, the current results suggest that people target information in memory using noisy temporal-contextual cues that activates a range of accurate memoranda but also activates some irrelevant memoranda unless the target information is distinct.

References

Albert, D., & Schulz, U. (1975). A model for free recall of tow lists: Individual recall sequences and the effect of List-2-dominance. *Acta Psychologica, 39,* 161–174. doi:10.1016/0001-6918(75)90031-1

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review, 79,* 97–123. doi:10.1037/h0033773

Bousfield, W. A., & Rosner, S. R. (1970). Free vs. uninhibited recall. *Psychonomic Science, 20,* 75–76.

Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of restricted associative responses. *Journal of General Psychology, 30,* 149–165. doi:10.1080/00221309.1944.10544467

Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.

Epstein, W. (1969). Poststimulus output specification and differential retrieval from short-term memory. *Journal of Experimental Psychology, 82,* 168–174. doi:10.1037/h0028045

Epstein, W. (1970). Facilitation of retrieval resulting from postinput exclusion of part of the input. *Journal of Experimental Psychology, 86,* 190–195. doi:10.1037/h0029982

Epstein, W. (1972). Mechanisms of directed forgetting. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 6, pp. 147–191). New York, NY: Academic Press. doi:10.1016/S0079-7421(08)60386-2

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instruments, 14,* 375–399. doi:10.3758/BF03203275

Gardiner, J. M., Craik, F. I. M., & Birtwistle, J. (1972). Retrieval cues and release from proactive inhibition. *Journal of Verbal Learning and Verbal Behavior, 11,* 778–783. doi:10.1016/S0022-5371(72)80012-4

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46,* 269–299. doi:10.1006/jmps.2001.1388

Indow, T., & Togano, K. (1970). On retrieving sequences from long-term memory. *Psychological Review, 77,* 317–331. doi:10.1037/h0029395

Jang, Y., & Huber, D. E. (2008). Context retrieval and context change in free recall: Recalling from long-term memory drives list isolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 112–127. doi:10.1037/0278-7393.34.1.112

Kahana, M. J., Dolan, E. D., Sauder, C. L., & Wingfield, A. (2005). Intrusions in episodic recall: Age differences in editing of overt responses. *Journals of Gerontology: Series B. Psychological Sciences and Social Sciences*, *60,* P92–P97. doi:10.1093/geronb/60.2.P92

Klein, K. A., Shiffrin, R. M., & Criss, A. H. (2007). Putting context into context. In J. S. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger, III* (pp. 171–189). New York, NY: Psychology Press.

Marsh, R. L., Landau, J. D., & Hicks, J. L. (1996). The post-information effect and reductions in retroactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1296–1303. doi:10.1037/0278-7393.22.5.1296

McGill, W. J. (1963). Stochastic latency mechanism. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 309–360). New York, NY: Wiley.

Mensink, G., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review, 95,* 434–455. doi:10.1037/0033-295X.95.4.434

Murdock, B. B. (1974). *Human memory: Theory and data*. Potomac, MD: Erlbaum.

Murdock, B. B. (1976). Item and order information in short-term serial memory. *Journal of Experimental Psychology: General, 105,* 191–216. doi:10.1037/0096-3445.105.2.191

Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 14, pp. 207–262). New York, NY: Academic Press. doi:10.1016/S0079-7421(08)60162-0

Roediger, H. L., & Payne, D. G. (1985). Recall criterion does not affect recall level or hypermnesia: A puzzle for generate/recognize theories. *Memory & Cognition, 13,* 1–7. doi:10.3758/BF03198437

Roediger, H. L., Stellon, C. C., & Tulving, E. (1977). Inhibition from part-list cues and rate of recall. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 174–188. doi:10.1037/0278-7393.3.2.174

Roediger, H. L., & Tulving, E. (1979). Exclusion of learned material from recall as a postretrieval operation. *Journal of Verbal Learning and Verbal Behavior, 18,* 601–615. doi:10.1016/S0022-5371(79)90334-7

Rohrer, D. (1996). On the relative and absolute strength of a memory trace. *Memory & Cognition, 24,* 188–201. doi:10.3758/BF03200880

Rohrer, D., & Wixted, J. T. (1994). An analysis of latency and interresponse time in free recall. *Memory & Cognition, 22,* 511–524. doi:10.3758/BF03198390

Sahakyan, L., & Hendricks, H. E. (2012). Context change and retrieval difficulty in the list-before-the-last paradigm. *Memory & Cognition, 40,* 844–860. doi:10.3758/s13421-012-0198-0

Shiffrin, R. M. (1970a). Forgetting: Trace erosion or retrieval failure? *Science, 168,* 1601–1603. doi:10.1126/science.168.3939.1601

Shiffrin, R. M. (1970b). Memory search. In D. A. Norman (Ed.), *Models of human memory* (pp. 375–447). New York, NY: Academic Press.

Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review, 79,* 179–193. doi:10.1037/h0027277

Sirotin, Y. B., Kimball, D. R., & Kahana, M. J. (2005). Going beyond a single list: Modeling the effects of prior experience on episodic free recall. *Psychonomic Bulletin & Review, 12,* 787–805. doi:10.3758/BF03196773

Smith, A. D. (1979). The interaction between age and list length in free recall. *Journal of Gerontology, 34,* 381–387. doi:10.1093/geronj/34.3.381

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1392–1399. doi:10.1037/a0013082

Tulving, E. (1968). Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 2–36). Englewood Cliffs, NJ: Prentice Hall.

Unsworth, N. (2007). Individual differences in working memory capacity and episodic retrieval: Examining the dynamics of delayed and continuous distractor free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 1020–1034. doi:10.1037/0278-7393.33.6.1020

Unsworth, N. (2009). Variation in working memory capacity, fluid intelligence, and episodic recall: A latent variable examination of differences in the dynamics of free recall. *Memory & Cognition, 37,* 837–849. doi:10.3758/MC.37.6.837

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2010). Understanding the dynamics of correct and error responses in free recall: Evidence from externalized free recall. *Memory & Cognition, 38,* 419–430. doi:10.3758/MC.38.4.419

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2012). Evidence for noisy contextual search: Examining the dynamics of list-before-last recall. *Memory, 20,* 1–13. doi:10.1080/09658211.2011.626430

Vorberg, D., & Ulrich, R. (1987). Random search with unequal search rates: Serial and parallel generalizations of McGill's model. *Journal of Mathematical Psychology, 31,* 1–23. doi:10.1016/0022-2496(87)90033-2

Ward, G., & Tan, L. (2004). The effect of the length of to-be-remembered lists and intervening lists on free recall: A re-examination using overt rehearsal. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1196–1210. doi:10.1037/0278-7393.30.6.1196

Wickens, D. D. (1972). Characteristics of word encoding. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 191–215). Washington, DC: Winston.

Wixted, J. T., & Rohrer, D. (1993). Proactive interference and the dynamics of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1024–1039. doi:10.1037/0278-7393.19.5.1024

Zaromb, F. M., Howard, M. W., Dolan, E. D., Sirotin, Y. B., Tully, M., Wingfield, A., & Kahana, M. J. (2006). Temporal associations and prior-list intrusions in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 792–804. doi:10.1037/0278-7393.32.4.792

Received June 20, 2012
Revision received May 29, 2013
Accepted May 30, 2013 ■