Intrusions and the Decision to Terminate Memory Search

J. Isaiah Harbison (isaiah.harbison@gmail.com)
Department of Psychology, University of Maryland at College Park
College Park, MD 20742

Eddy J. Davelaar (eddy.davelaar@gmail.com)
Depatiment of Psychological Sciences, Birkbeck University of London
Malet Street, WC13 7HX, London, UK

Erica C. Yu (erica.c.yu@gmail.com)
Erika K. Hussey (erikahussey@gmail.com)
Michael R. Dougherty (mdougher@umd.edu)
Department of Psychology, University of Maryland at College Park
College Park, MD 20742

Abstract

Little is known about the how the decision is made to terminate memory search, though there have been several recent attempts to uncover this process. In one recent study, Miller et al. (2012), re-analyzed data from a large number of free-recall experiments and identified intrusions as a factor that influenced search termination decisions. One potential problem with this re-analysis is that all the data were drawn from experiments in which it was impossible to determine if or when search was terminated. Using data from experiments in which search termination decisions were directly measured, we confirmed Miller at al.'s (2012) original findings but also demonstrated that intrusions influence the time taken to generate the final retrieval and the time between the final retrieval and search termination. The pattern of data is consistent with a simple, sample-with-replacement model in which intrusions are less active than items from the target list.

Keywords: recall; memory search termination; stopping rules

Every search of memory is eventually terminated. When an individual decides to terminate their own memory search (e.g., when they are not interrupted or given a fixed time limit for search), what factors influence this decision? The long history of memory research is relatively silent regarding this question as most memory recall experiments give participants a pre-determined amount of time to search memory (a closed-interval) and have no method of determining when or, even if, participants terminate their search before the retrieval interval expires. When participants are allowed to terminate their own search, as is the case in the open-interval design discussed below, a number of dependent variables emerge that allow the measurement of memory search termination decisions. These variables include the total time spent in search (total time or T_T), which is controlled by the participant in this design. The total time can be divided into the time from the beginning of search to the time of the final retrieval (time-to-last retrieval or T_L) and the time between the final retrieval and search termination (exit latency or E_L; Dougherty & Harbison, 2007). These variables allow for the testing of different memory search stopping rules previously proposed in the literature, with much of the available data uniquely supporting the cumulative-failures stopping rule (Harbison, Dougherty, Davelaar, & Fayyad, 2009) proposed within the search of associative memory model (SAM; Raaijmakers & Shiffrin, 1981). According to this rule, every retrieval attempt that does not produce a new retrieval is counted as a retrieval failure and search is terminated when the number of these failures reaches a threshold.

Recent research, however, has suggests that search termination might also be influenced by the presence of memory intrusions (Miller, Weidmann, & Kahana, 2012; Unsworth, Brewer, & Spiller, 2011). Miller et al. (2012) showed that memory search was more likely to be terminated after an intrusion from a previous list (prior list intrusion or PLI), an extra list intrusion (ELI), or after outputting a list word that had previously been retrieved (repetition). They suggested that the increase in the probability of stopping may be due to the effect such retrieval errors have on subsequent recall. Each retrieved word is thought to influence subsequent retrievals either by the use of the retrieved word as a cue for subsequent retrieval attempts (Kimball, Smith, & Kahana, 2007; Sirotin, Kimball, & Kahana, 2005) or by the retrieved item shifting the contextual retrieval cues closer to the retrieved items own context (Howard & Kahana, 2002). Intrusions and repetitions then would decrease the probability of retrieving a new target list word. A PLI would increase the relative probability of another word from the prior list being retrieved; an ELI would increase the probability of sampling related extra-list words, and repetitions would increase the probability of retrieving other words that have already been retrieved.

One potential problem with the Miller et al. analysis is that they had to infer when participants terminated search since the experiments they used in their analysis used a closed-interval design. To determine when participants might have terminated search, they used data from an open-interval experiment (Dougherty & Harbison, 2007) to set an inter-retrieval time longer than participants were found to search memory before terminating search. However, the factors they identified as increasing the probability of search termination would also slow down retrieval. That is, if intrusions increase the probability of sampling non-target words, this decreases the probability of sampling target list words. When the probability of sampling a word decreases, the expected time to sample that word increases. Therefore, it might be that intrusions did not increase the probability of search termination but simply slowed retrieval sufficiently for no additional words to be output in the retrieval interval. Participants might have continued searching but to no avail.

In the present study, we used data from experiments that used an open-interval paradigm to examine whether intrusions increase the probability of search termination when participants were required to indicate when they terminated search. Second, we tested the hypothesis that the retrieval of an intrusion changes the probability of subsequent retrieval types. Third, we tested if terminating search after an intrusion changes the temporal variables of search termination (T_T, T_L, and E_L). Fourth, we evaluated whether the results could be modeled using a cumulative-failures stopping rule, and if so what assumptions were needed.

Open- vs Closed-Interval Retrieval

The difference between the standard free recall paradigm, or the closed-interval design, and the open-interval design is depicted in Figure 1. In the closed-interval design participants are given a predetermined length of time for retrieval (e.g., 60 seconds). In this design, the decision to terminate search is obscured. During the retrieval interval participants might continue to search memory throughout the entire interval, terminate search immediately after the final retrieval, or even retrieve items after terminating search (e.g., a participant might terminate search then stumble upon another list word while letting their mind wander). In contrast, the open-interval design allows participants to continue search until they decide to terminate their own search (depicted in Figure 1 with an "X"). As mentioned above, three temporal variables emerge as measures of memory search termination and it has been found that the T_T and T_L increases while the E_L decreases with the number of items retrieved (Dougherty & Harbison, 2007; Harbison et al., 2009; Unsworth, Brewer, & Spillers, 2011).

The open-interval design is therefore particularly useful for examining search termination decisions and provides not only a method of replicating the Miller et al. results, but also a method of extending them. Using this design it is possible to determine if search is terminated

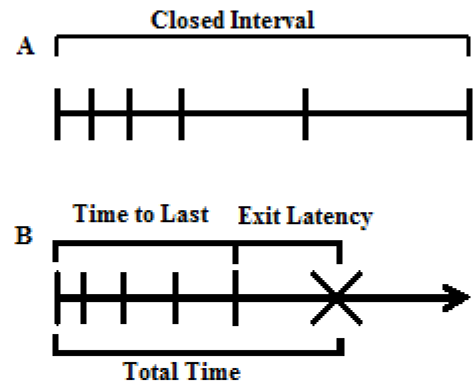

Figure 1: Comparison of A) Closed-Interval and B) Open-Interval Experimental Designs.

more quickly if an intrusion is the final outputted word relative to when a word from the target list is the final retrieval.

We re-analyzed data from two previous experiments using the open-interval design to test for these results. Importantly, the data from these two experiments utilized what we call a multi-target cued recall paradigm. Within this paradigm, participants studied multiple separate lists of words successively, with all words from each list paired with a single cue word. After studying all the lists, participants were presented with a cue and asked to recall as many words as possible that had appeared with the cue. This aspect of the experiments is particularly useful for examining the role of intrusions on search termination because it should increase the probability of PLIs. The multi-target cue task can be contrasted with the standard procedure for list recall experiments, which is to have participants study a single list at a time and then, possibly after a filler task to clear short-term memory (e.g., solving simple math problems), prompt participants to retrieve words from the studied list.

Analysis of Previous Experiments

As mentioned above, both experiments used a multi-target cued recall paradigm. For the first experiment (Dougherty & Harbison, 2007), three lists were shown in each block for four total blocks of lists providing data from twelve lists per participant. For the second (Experiment 1, Harbison et al., 2009), four lists were shown in each of the four blocks providing sixteen lists per participant. The two experiments also differed in list length. There were ten words per list in the first experiment and eight words per list in the second. Both experiments used lists of high (KF ¿ 50) and low (KF ¡ 10) frequency words drawn from the MRC Psycholinguistic Database (Wilson, 1988; available from http://www.psy.uwa.edu.au/mrcdatabase.uwa_mrc.htm). For the second experiment, in addition to high and low frequency lists, two of the lists per block were a mix of high and low frequency words (four of each). One mixed-frequency list had a high frequency cue word per block and one had a low frequency cue word. In both experiments, word frequency did not have a significant effect on the memory search termination variables. Therefore, the analyses collapsed across lists of high and low frequency words.

Table 1: Mean Number Retrieved

Source	List	PLI	ELI
Dougherty & Harbison, 2007	3.40	.25	.18
Exp 1. Harbison et al., 2009	2.12	.53	.42
New Experiment	1.84	.32	.06

Table 1 shows the mean number of list words, PLIs, and ELIs per participant per list. Note that the number of PLIs though still small made up 6% and 17% of the total number of items retrieved, respectively. Following the procedure of Miller et al. (2012), we examined the probability of terminating search as a function of both the previous retrieval type (list word, PLI, ELI) and output position. Note that for all results reported in this study statistical significance was determined by biased-corrected and accelerated (BCa) bootstrap estimates of the 95% confidence intervals. Here, we test the differences between list words and the two types of intrusions. The results from both previous experiments were consistent with the PLIs and ELIs increasing the probability of search termination, as shown in Figure 2A, but none of these differences were significant for the Dougherty and Harbison (2007) experiment. For the Harbison et al. (2009) experiment, shown in Figure 2B, output positions two and three were significant for both ELIs and PLIs, also ELIs were significant for the first position. Therefore, these results replicated those of Miller et al. (2012).

The hypothesis for why intrusions increase the probability of search termination was also tested. As mentioned above, the explanation is that intrusions increase the probability of non-list items being sampled after an initial intrusion. These intrusions could be words from a previous list, as should be the case for PLIs, or words from outside the experiment, as should be the case for ELIs. Figure 3A and B shows the probability of list words, PLIs, ELIs, and search termination immediately after each type of retrieval. For both experiments, after a PLI participants were more likely to terminate search, less likely to output a list word, and, importantly, more likely to output another PLI relative to after they had generated a list word. In contrast, after ELIs there were only significant differences in the probability of generating list words and terminating search and this was only found to be significant in the Harbison et al. Experiment

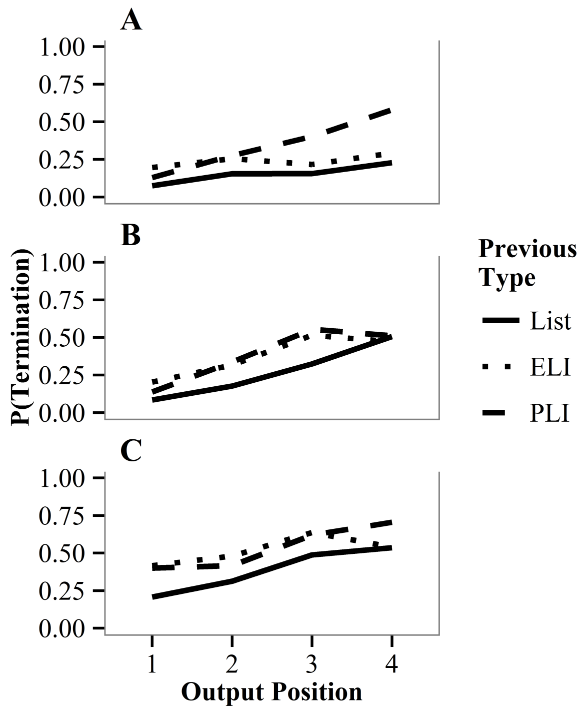

Figure 2: Probability of terminating search as a function of the previous retrieval and output position for A) Dougherty & Harbison, 2007; B) Harbison et al., 2009, Exp 1; and C) the new experiment.

1 data.

These results are consistent with the explanation proposed by Miller et al.(2012). Not only are PLIs associated with an increase in the probability of termination, but they are also associated with an increase in the probability of generating other words from previous lists. The support for ELIs is mixed, but still consistent with this explanation.

Overall the pattern of results using the open-interval design were consistent with those reported using the closed-interval design. Search termination was more likely after an intrusion, especially a PLI. Furthermore, the retrieval of a PLI does appear to be correlated with an increase in the probability of retrieving words from previous lists and a decrease in the probability of retrieving subsequent list words. However, the replication of difference in the probability of terminating search by type and output position was only significant for one of the two re-analyzed experiments. Therefore, to further test Miller et al.'s results a new experiment using the open-interval and multi-target cued recall was conducted.

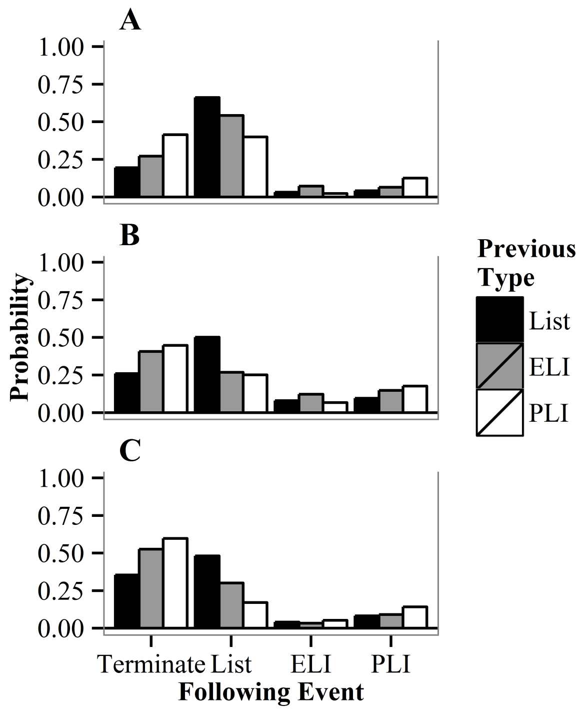

Figure 3: Probability of retrieval type or termination as a function of the previous retrieval type.

Experiment

103 Participants were randomly assigned to one of two conditions. Participants in one condition were given two blocks of four high frequency word lists followed by two blocks of low frequency word lists while participants in the other condition were given four blocks of low frequency word lists. This manipulation was designed to test a hypothesis about the role of individual differences in motivation on the influence of previous retrieval experience on stopping decisions (Dougherty & Harbison, 2007) which is outside of the scope of the present study. As was found in the previous experiments, word frequency did not influence stopping decisions outside of the impact on number of words retrieved. Therefore, consistent with the previous experiments, all lists were combined for the purpose of examining stopping decisions.

Stimuli

Both high and low frequency words were drawn from MRC linguistics database (Wilson, 1988), the same source for lists in the previous experiments, and the same criteria for high and low frequency words were used. Lists of eight to be recalled words and one cue word were random generated for each participant.

Procedure

The same open-interval, multi-target cued recall procedure was used as in the previous experiments. During learning, participants were presented with the cue word and each list word from the first, second, third, and fourth list. The learning phase was then repeated. Thus, participants saw each list word with that list's cue word twice. Participants were then asked to retrieve the words from the first, second, third, and fourth lists, in that order. There were a total of four blocks of lists. As before, participants were provided an open-interval for retrieval. They indicated when they were finished retrieving from each list by saying 'Stop" and pressing the space bar.

Results and Discussion

The exit latency and total time results replicated the findings from previous open-interval experiments (Dougherty & Harbison, 2007; Harbison et al., 2009; Unsworth et al., 2012). E_Ls were negatively correlated with number retrieved and T_L and T_T were positively correlated with number retrieved. The mean within participant gammas of -0.293, 0.224, and 0.472, respectively, were each significant.

Intrusions and Search Termination As shown in Figure 2C, the results were again consistent with Miller et al. (2012). PLIs and ELIs were consistently more likely to be the final word retrieved before termination across output positions relative to target list words. However, only the output positions with the most participants contributing to them, the first three positions, were significantly different between the intrusion types (both PLI and ELI) and list words.

The shift in the probability of retrieving subsequent list items and intrusions after retrieving an intrusion was found in the present experiment, matching the results of the re-analyzed experiments. Participants were more likely to have a PLI after a PLI and more likely to terminate search while also being less likely to retrieve a target item. Likewise, after an ELI, participants were more likely to terminate and less likely to generate a target word. One difference from the results reported above is that participants were also more likely to generate an ELI after an ELI in the present experiment. This pattern is shown in Figure 3C.

Intrusions and the Time Course of Termination The open-interval design also allows for the testing of temporal effects of intrusions. Specifically, does the generation of an intrusion have a different profile in terms of exit latency (E_L), time-to-last (T_L), and total time (T_T)? To test this, we compared temporal variables when search was terminated after both types of intrusions with instances where retrieval was terminated after a target list word was generated. Note that rate transformations were used for the purposes of the analyses and that we again used BCa to estimate the 95% confidence intervals.

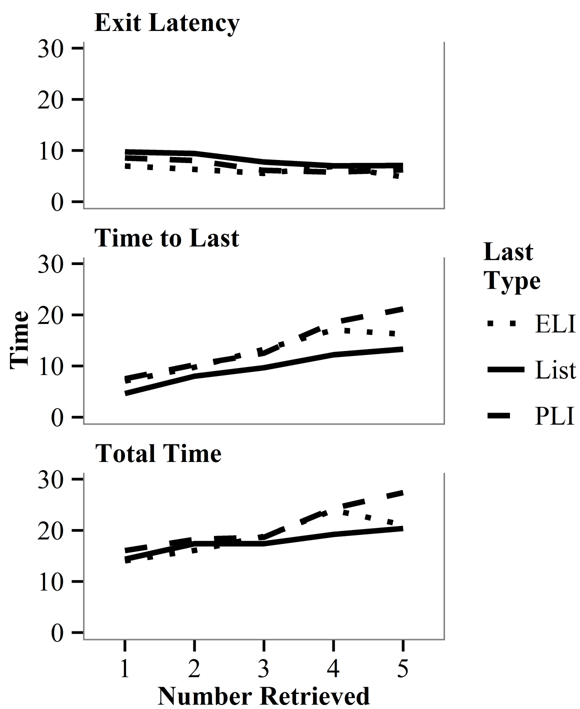

Figure 4: A) Exit Latency, B) Time to Last, and C) Total Time as a function of final retrieval type and total number retrieved.

We combined the data from the present experiment and the experiments included in the above reanalysis. The combined E_L, T_L, and T_T results are shown in Figure 4 as a function of final retrieval type and total number retrieved. The E_L after a PLI was significantly shorter than when the final word was from the target list for four of the five cases. Furthermore, the T_L was longer after a final PLI for all five total number retrieved. The pattern of results was less clear for the differences between target list words and ELIs. The E_L was shorter for three of the five cases and the T_L was longer for two of them. The results were even less consistent for T_T. The difference between list words and PLIs were significant for two output positions and the difference between list words and ELIs was only significant for one of the positions.

The temporal data add another portion of the picture of the role of intrusions in search termination. Combined, the present results suggest that participants are more likely to terminate search following an intrusion, more likely to generate a subsequent intrusion of the same type (particularly for PLIs), that the time to generate the final item is longer if it is an intrusion (particularly for PLIs), and that the time between the final retrieval and termination is shorter after an intrusion (particularly for PLIs). This pattern of data provides a new challenge for models of recall and particularly stopping rules to account for. The next section examines how well as simple sample-with-replacement model is able to account for these results.

Stopping Rules and Intrusions

For the present simulation, we tested the explanation that intrusions have lower activations relative to list words. Like intrusions, items with lower activation are retrieved later in recall. Here we tested if items with lower activation shared the additional characteristic of intrusions. The present focus is on the sampling process itself. Therefore, for the sake of simplicity, we assumed a set of activations instead of modeling encoding and the activation process (Harbison, Hussey, Dougherty, & Davelaar, 2012; Rohrer, 1996).

A sample-with-replacement model equipped with the cumulative failures stopping rule was used to predict search. The pattern of activations used to test the model was (.5, .5, .5, .5, 1, 1, 1.5, 1.5, 2, 2). We ran the sample-with-replacement procedure where the probability of sampling item i was determined by the act_i divided by the sum over all item activations. A retrieval attempt was successful if the sampled item had not previously been retrieved and if the activation of the sampled item was greater than the recovery threshold of 0.5. Four unrecoverable items were included in the activation pattern. The total number of retrieval failures was tracked and once this number exceeded the stopping threshold of 30 search retrieval was terminated as prescribed by the cumulative failure stopping rule (Harbison et al., 2009). The stopping threshold was within the range tested in previous applications of the model (Harbison et al., 2012) and the activation pattern used, specifically the inclusion of four items that were not recoverable, was chosen to increase the variability in the number of items retrieved. Also to this end, for each simulation run a subset of six items was chosen at random (and with equal probability) from the complete pattern of activations. Without this, the number of items retrieved was too consistent to be at all comparable to participant data.

The results from 10,000 independent runs of list recall are presented in Figure 5. Items that were relatively less active show the same profile as intrusions. The probability of terminating search after an item with less activation was greater than for an item of greater activation, as shown in Figure 5A. Also, if items with lower activation were the final retrieved, it took longer to generate them (T_L is greater) and search was terminated more quickly after the final retrieval (E_L is smaller), as shown in Figure 5B. Therefore, the present results relating intrusions to search termination can be accounted for by

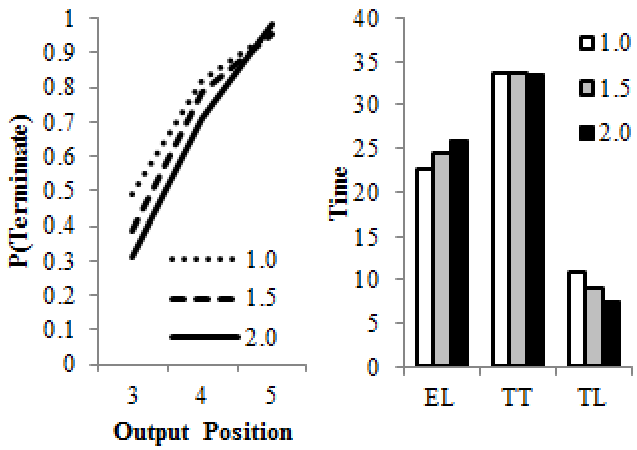

Figure 5: A) Probability of search termination and B) E_L, T_T, and T_L by item activation.

a simple sample-with-replacement model equipped with the cumulative failures stopping rule as long as it can be assumed that intrusions are relatively less active than words from the target list. This seems a reasonable assumption as models taking into account multiple sources of association (context, experimental word associations, and semantic word associations) often a assume a multiplicative use of search cues (Sirotin, Kimball, & Kahana, 2005; Kimball, Smith, & Kahana, 2007) and intrusions should at least have a lower association strength to the target context than list words.

In contrast, the model does not account for the observed impact of intrusions on subsequent retrieval. That is, it cannot account for the greater probability of retrieving a PLI when the immediately preceding item is a PLI. Therefore, it might be necessary to assume that the retrieval of an intrusion changes the contextual landscape which determined the probability of retrieving subsequent PLIs and target list items. Provided this extra assumption, the cumulative failure stopping rule should have no trouble accounted for the present results.

Conclusion

The results from three experiments using the open-interval design replicate the results from a large-scale re-analysis of closed-interval experiments (Miller et al., 2012) that participants are more likely to terminate search after an intrusion (either from previous list or from outside the experiment) than after a correct recall. Furthermore, search was terminated more quickly after an intrusion (smaller exit latency) and the time taken to generate an intrusion as the final output was greater (greater time-to-last retrieval). These results were found to be consistent with the predictions of a sample-with-replacement model equipped with a cumulative failure stopping rule. As such, the results suggest that the role of intrusions in search termination is an indirect one. Intrusions could lead to a greater probability of retrieval failures, with the total number of retrieval failures still being the direct cause of search termination.

Acknowledgements

We gratefully acknowledge NSF (grant BCS-1030831) for support of this research.

References

Dougherty, M. R., & Harbison, J. I. (2007). Motivatied to retrieve: How often are you willing to fo back to the well when the well is dry? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 1108–1117.

Harbison, J. I., Dougherty, M. R., Davelaar, E. J., & Fayyad, B. (2009). On the lawfulness of the decision to terminate memory search. *Cognition*, *111*, 416–421.

Harbison, J. I., Hussey, E. K., Dougherty, M. R., & Davelaar, E. J. (2012). Self-terminated vs. experimenter-terminated memory search. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th annual conference of the cognitive science society* (pp. 426–431). Austin, TX: Cognitive Science Society.

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

Kimball, D. R., Smith, T. A., & Kahana, M. J. (2007). The fSAM model of false recall. *Psychological Review*, *114*, 954–993.

Miller, J. F., Weidemann, C. T., & Kahana, M. J. (2012). Recall termination in free recall. *Memory & Cognition*, *40*, 540–550.

Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*, 93–134.

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

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.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2011). Factors that influence search termination decisions in free recall: An examination of response type and confidence. *Acta Psychologica*, *138*, 19–29.

Wilson, M. D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavioral Research Methods, Instruments and Computers*, *20*, 6–11.