Journal of Experimental Psychology:
Learning, Memory, and Cognition
2008, Vol. 34, No. 2, 408–414

0278-7393/08/$12.00 DOI: 10.1037/0278-7393.34.2.408

BRIEF REPORTS

Intentional Forgetting Is Easier After Two "Shots" Than One

Lili Sahakyan, Peter F. Delaney, and Emily R. Waldum
University of North Carolina, Greensboro

Three experiments evaluated whether the magnitude of the list-method directed forgetting effect is strength dependent. Throughout these studies, items were strengthened via operations thought to increase context strength (spaced presentations) or manipulations thought to increment the item strength without affecting the context strength (processing time and processing depth). The assumptions regarding which operations enhance item and context strength were based on the "one-shot" hypothesis of context storage (K. J. Malmberg & R. M. Shiffrin, 2005). The results revealed greater directed forgetting of strong items compared with weak items, but only when strength was varied via spaced presentations (Experiment 3). Equivalent directed forgetting was observed for strong and weak items when strengthening operations increased item strength without affecting the context strength (Experiments 1 and 2). These results supported the context hypothesis of directed forgetting (L. Sahakyan & C. M. Kelley, 2002).

Keywords: directed forgetting, context change, list-strength effect, spacing effect

Are some things easier to deliberately forget than others? We will propose that the answer to this question is yes and that the kinds of things easier to forget may provide important clues as to how intentional forgetting occurs. Such efforts are timely because there has been increasing interest in directed forgetting as part of a broader trend toward investigating inhibitory abilities. To date, published research has not addressed how recallability of an item influences the magnitude of directed forgetting—that is, if some things are easier to remember, how does the ease of remembering influence the magnitude of deliberate forgetting?

The current article seeks to explore the theoretical mechanisms of one of the two most frequently-used directed forgetting paradigms: list-method directed forgetting (invented by R. A. Bjork, LaBerge, & LeGrand, 1968). In the list method, participants encode two lists of items. Between the lists, some participants are told to try to forget everything up to that point. Typically, directed forgetting has a dual effect on memory; compared with a remember control group that is not told to forget anything, the forget group shows impaired recall of List 1 items (known as the *costs*) but enhanced recall of List 2 items (known as the *benefits*).

Directed forgetting is not caused merely by withholding responses in the forget group, because monetary incentives to recall additional information do not elicit better recall (MacLeod, 1999), nor does the procedure completely erase the items from memory, because recognition tests and indirect memory tests do not show directed forgetting, suggesting that the items are available in memory but are simply inaccessible (Basden et al., 1993; E. L. Bjork & Bjork, 1996). Directed forgetting has also been observed in incidental learning (Geiselman, Bjork, & Fishman, 1983; Sahakyan & Delaney, 2005), suggesting that pure rehearsal-based explanations are insufficient to explain the phenomenon. Consequently, some researchers proposed that directed forgetting involves an inhibitory process (E. L. Bjork & Bjork, 1996; R. A. Bjork, 1989). According to this view, an inhibitory mechanism is invoked at the time of retrieval that reduces access to unwanted memories, producing lower recall of List 1 items (the costs of directed forgetting). Because List 1 items are less accessible, they are correspondingly less likely to interfere with List 2 items, producing the benefits of directed forgetting (e.g., E. L. Bjork & Bjork, 1996).

An alternative account suggests that directed forgetting is a context effect (Sahakyan, 2004; Sahakyan & Kelley, 2002). Our interpretation of the directed forgetting effect was inspired by the search of associative memory (SAM)/retrieving effectively from memory (REM) memory models (e.g., Gillund & Shiffrin, 1984; Shiffrin & Steyvers, 1997) and their variants (e.g., Mensink & Raaijmakers, 1988). Like many memory theories, these models distinguish between item information and context information. They assume that processing an item constructs an *image* in memory, which is an interconnected set of features that represents information about the item content (such as its lexical/semantic representation) and the context in which that item was learned. Typically, item content information is the focus of attention during study, whereas context information is more peripheral information referring to the physical, spatial–temporal, environmental, physiological, or emotional states in which the item was experienced (Murnane, Phelps, & Malmberg, 1999). During recall, a person

Lili Sahakyan, Peter F. Delaney, and Emily R. Waldum, Department of Psychology, University of North Carolina, Greensboro.

We thank Sarah Thrower, Greg Spillers, Bridget Glover, Kimberly Ward, Ethan Smith-Murray, Rafe Kielar, Jennifer Brown, Jennifer Hynes, Elizabeth Lancaster, Cathy Sawyer, Virginia Wilmoth, Liliana Camara, Krycya Flores-Rojas, Vincent Woolfolk, Jared Cook, and Heather Coby for their assistance in data collection and scoring.

Correspondence concerning this article should be addressed to Lili Sahakyan, Department of Psychology, University of North Carolina, Greensboro, 296 Eberhart Building, P.O. Box 26170, Greensboro, NC 27402. E-mail: l_sahaky@uncg.edu

assembles a set of retrieval cues that are used to activate and select images in long-term memory—a process called sampling. The model samples images, and whenever an image is sampled, a separate recovery process attempts to retrieve information out of that image. Successful sampling does not necessarily guarantee successful recovery. In free recall, the first sampling attempt relies strictly on context cues. If an item is successfully recovered, then that item serves as an additional retrieval cue (along with the context cues) to guide subsequent retrieval attempts. However, if the attempt to recall was unsuccessful, then context cues alone continue to guide search. The success of the context cue when accessing relevant images in long-term memory depends on the overlap between context at the time of storage and the test context.

Given the importance of context cues in free recall, Sahakyan and Kelley (2002) proposed a context-based explanation for directed forgetting. They suggested that forget instructions encourage establishing a new mental context for List 2 encoding. One strategy for deliberately forgetting List 1 would be to stop maintaining the contextual elements that were present at the time of List 1 encoding and instead sample new contextual cues for List 2. For example, the forget instruction may induce some people to deliberately think of something unrelated to the experiment, leading to changes in mental context between the two lists. At the time of final recall, the context cues will match the List 2 learning context better than the List 1 context, producing forgetting of List 1 items. Recall is poorer when there is a low correspondence between the context at the time of the study and test (see Smith & Vela, 2001, for a meta-analysis of experimental findings on context-dependent forgetting). Empirical support for the context-based explanation of directed forgetting came from studies that produced effects similar to directed forgetting without instructing people to forget but, instead, by engaging participants in a diversionary thought intended to change their mental context between the two lists (Sahakyan & Delaney, 2003; Sahakyan & Kelley, 2002).

If directed forgetting costs arise from a context mismatch between study and test, then items that are more strongly linked to List 1 context should be more forgettable when context changes (as in the Forget condition). Imagine two events (A and B) that were experienced in the same context, but Event A was more strongly linked to that context than was Event B. When context is used as a retrieval cue during recall, it will be a better cue for A than for B. However, if context changed between the study and test, it should impair retrieval of A more than that of B. Therefore, directed forgetting should hurt items that were more strongly linked to List 1 context more than items that were weakly linked to List 1 context.

To examine this hypothesis, one needs to know what strengthens the relationship between an item and its context as well as what simply increases item strength without enhancing context storage. Malmberg and Shiffrin (2005), building on SAM/REM, suggested that a fixed amount of context information is stored when an item is first studied. Extra study time and/or deeper processing do not substantially enhance the storage of context information; they merely increment the item strength. In contrast, spaced presentations of the study items strengthen both item information and context information. Their conclusions were based on the presence or absence of the list-strength effect (LSE) in memory. While a list of strong items is recalled better than a list of weak items, the recall difference between strong and weak items is even larger when they appear on the same list—known as the LSE (Ratcliff, Clark, & Shiffrin, 1990; Tulving & Hastie, 1972). In SAM/REM, the LSE in free recall is attributed to a sampling advantage of items with more context storage. When context is used as a retrieval cue at the time of recall, images containing strong context become activated to a greater extent than images with weak context and are therefore sampled preferentially. Thus, on mixed lists there is a tendency to sample the strong images at the expense of sampling weaker images. The sampling difference is the cause of LSE in these models. Once an image is sampled, the item content features contained in that image determine the success of the recovery process. For example, on pure lists, the memory advantage of strong items over weak items is attributed to differences in the recovery process because the sampling of all images is approximately equivalent on pure lists.

To summarize, in models like REM/SAM, the observation of the significant LSE implies that the study conditions must have enhanced the context strength, whereas the absence of LSE (despite the overall item strength effect on both mixed and pure lists) implies that context strength was not incremented by the study condition. Malmberg and Shiffrin (2005) found an LSE only for spaced repetitions, but not for other strengthening operations such as depth of processing or lengthening the study time. They concluded that each spaced presentation of the item leads to the storage of an additional "shot" of context, but other varieties of strengthening, such as additional processing time or depth of processing, do not enhance the context storage but merely increment the item strength. They termed this the *one-shot hypothesis* of context storage. For a full description and the formal account of the one-shot hypothesis and the REM/SAM model, refer to Malmberg and Shiffrin (2005).

Given the assumptions of the one-shot hypothesis, the context account of directed forgetting predicts that greater directed forgetting should be observed for strong items than for weak items when items are strengthened by an operation that produces an LSE (e.g., spacing). In contrast, strengthening manipulations that do not enhance context storage should lead to equivalent amounts of directed forgetting from strong and weak items. Therefore, in three experiments, we used different strengthening manipulations in conjunction with the list-method directed forgetting design, where List 1 words consisted of *strong items* and *weak items* for each participant.

Participants were instructed to perform judgments on two word lists and were told that their judgments would help create materials for a future study. Thus, words were encoded incidentally rather than intentionally. Incidental learning allows better experimental control over encoding manipulations because it circumvents rehearsal, which could redistribute study time inequitably across strong and weak items (e.g., Rundus, 1971). Additionally, some authors have suggested that directed forgetting costs arise partly from rehearsal differences between the forget and remember conditions (e.g., Benjamin, 2006; Sheard & MacLeod, 2005). Therefore obtaining directed forgetting with incidental learning would strengthen the argument that processes beyond selective encoding underlie directed forgetting.

Elsewhere we have argued that the main cause of directed forgetting benefits in intentional learning is that participants in the forget group adopt more elaborate encoding strategies during List 2 learning compared with participants in the remember group (Sahakyan & Delaney, 2003). Given incidental learning procedures, we did not expect to observe the benefits because incidental encoding prevents encoding strategy changes between the lists and produces no benefits (Sahakyan & Delaney, 2005).

Experiment 1: Depth of Processing

In Experiment 1, we manipulated strength by varying the depth of processing; that is, *strong* words were encoded with an orienting task involving a pleasantness judgment, whereas *weak* words were encoded with an orienting task that involved determining whether the word contained letter *E*.

Method

Participants. Sixty-four University of North Carolina, Greensboro, undergraduates participated for course credit and reported not expecting a memory test in their final verbal reports. They were randomly assigned to the forget or control group.

Materials. Two lists of 16 medium-frequency nouns were created. Half of the List 1 words contained the letter *E* and half did not. List 1 was split into two blocks of 8 items that were encoded with the same orienting task (either pleasantness judgment or letter *E*). Blocked presentation was employed to avoid frequent task switching and floor recall of List 1 weak items. The presentation order of the strong versus weak blocks was counterbalanced. Also, each word was assigned to the pleasantness task or the letter *E* task equally often. On List 2, half of the words rhymed with the word *seven* and half did not. All List 2 items were encoded with the same orienting task involving rhyming.

Procedure. Participants were informed that the purpose of the study was to design word norms for materials for a later experiment (cover stories are available on request). They were told that if a smiley-face icon appeared under a word, they were to indicate if it was pleasant or not (strong items), and if the letter *E* appeared under a word, they were to indicate if the word contained an *E* (weak items). Items were presented once at a rate of 4 s per word.

After the first list, half of the participants were told that they would need to rate one more list (further termed the *control group*). The remaining participants were told to forget the first list before moving on to rate the second list (further termed the *forget group*). They were told,

> The list of words you just rated was the first list that we need to collect ratings for. Could you please rate one more list of words? However, it is really important that you not be influenced by your prior ratings. Therefore, please make an effort to try and forget those words and pretend you did not rate anything. Try not to think of the earlier words in order not to contaminate the judgments on these new items.

Then the second list of 16 items was presented. List 2 words were judged for whether or not they rhymed with the word *seven*. The choice of the rhyming task was based on pilot testing because when both lists were processed with the same orienting task, the recall of the List 1 weak items was at floor. A surprise written free recall test was given afterwards on List 1, followed by List 2. Recall was carried out on separate sheets of paper for each list, with 90 s allotted for the recall of each list. After the memory test, participants were asked whether they suspected that their memory would be tested.

Results

The proportion of recalled List 1 words (including cross-list confusions) was submitted to a Cue (forget vs. control) × Item Strength (strong vs. weak) mixed analysis of variance (ANOVA). There was a significant main effect of cue, $F(1, 62) = 10.81$, $p < .01$, $MSE = .016$, $\eta^2 = .15$, indicating fewer List 1 items were recalled following a forget instruction (.29) than were recalled in the control condition (.37). Item strength was also significant, $F(1, 62) = 195.79$, $p < .001$, $MSE = .022$, $\eta^2 = .76$, indicating that strong items were better recalled (.51) than weak items (.15). However, there was no interaction ($F < 1$; see Figure 1), indicating that both types of items were equally forgettable. There were also no significant differences in List 2 recall ($F < 1$; .30 in forget and .31 in control).

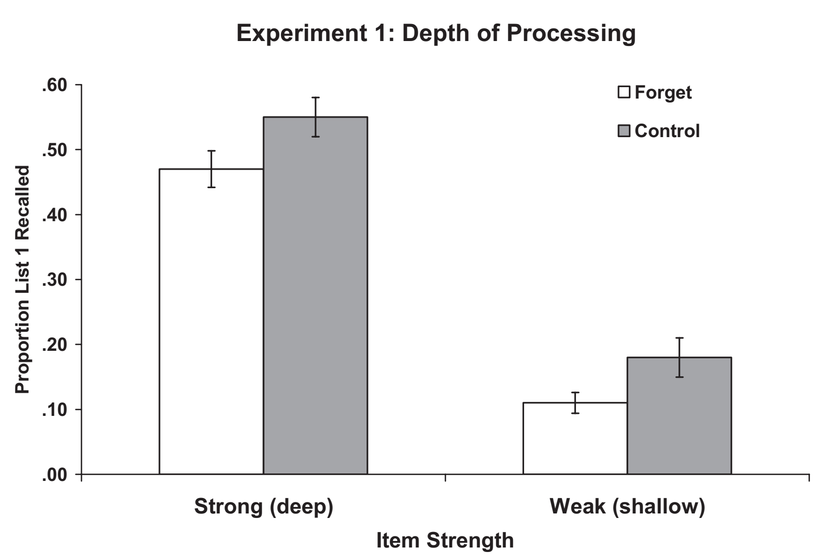

Figure 1. Mean proportion List 1 recall (±*SE*) by cue and item strength in Experiment 1.

Experiment 2: Processing Time

Experiment 2 varied strength by manipulating the processing time, with *strong* items processed for longer durations (two consecutive presentations) and *weak* items processed for shorter durations (single presentation).

Method

Participants. Sixty-four University of North Carolina, Greensboro, undergraduates participated for course credit. None of these participants expected a memory test as indicated by the retrospective questioning.

Materials. Two new lists of 16 medium-frequency nouns composed the stimuli for Experiment 2. Four versions of List 1 were created in order to counterbalance (a) whether a given word was presented once or twice and (b) whether the first word on the list was a once-presented word or a twice-presented word. Four more versions were created by exchanging the first four words of each type (once-presented or twice-presented) with the other four words of that type. A pleasantness judgment was performed on all items.

Procedure. The procedure was similar to Experiment 1, except that words were rated only for pleasantness. *Short* processing times involved a single 4-s judgment of pleasantness. *Long* processing times involved two consecutive 4-s presentations with two judgments of pleasantness (for a total processing time of 8 s). Participants were warned in advance that some items may be presented twice and that they should make a separate judgment for each presentation so that we could track the reliability of their ratings. After rating List 1 words, half of the participants were instructed to try to forget them with the instructions from the previous experiment. List 2 words were each presented for 4 s and were rated for pleasantness. Afterwards, a 90-s surprise recall test was given for List 1, followed by a 90-s recall of List 2 on separate sheets of paper.

Results

A Cue × Item Strength mixed ANOVA revealed a significant main effect of item strength, $F(1, 62) = 12.29$, $p < .01$, $MSE = .017$, $\eta^2 = .17$, indicating that strong items were better recalled (.37) than weak items (.29). There was also a significant main effect of cue, $F(1, 62) = 5.29$, $p < .05$, $MSE = .032$, $\eta^2 = .08$, indicating that fewer List 1 items were recalled following a forget instruction (.29) than were recalled in the control condition (.37). However, there was no interaction ($F < 1$; see Figure 2), suggesting that both types of items were equally forgettable. There were no significant differences in List 2 recall ($F = 1.00$; .32 in forget and .31 in control).

Experiment 3: Spacing

In Experiment 3, strength of List 1 items was varied by spacing repetitions of words. Weak items were repeated twice consecutively (*massed* presentation), while strong items were repeated twice but with several other words in between the two repetitions (*spaced* presentation).

Method

Participants. A total of 96 University of North Carolina, Greensboro, undergraduates participated for course credit. They were tested individually. Six participants were replaced because post-session questioning revealed that they expected a memory test.

Materials. Materials were the same as in Experiment 2. List 1 items were presented twice, with eight items presented through massed repetition and eight items through spaced repetition. Two types of judgments were performed on each word—one pleasantness judgment and one living/non-living judgment. Spaced items' first repetition occurred in the first half of List 1, and their second presentation was in the second half, with an average lag of 12.5 items. Half of the massed items appeared in the first half of List 1,

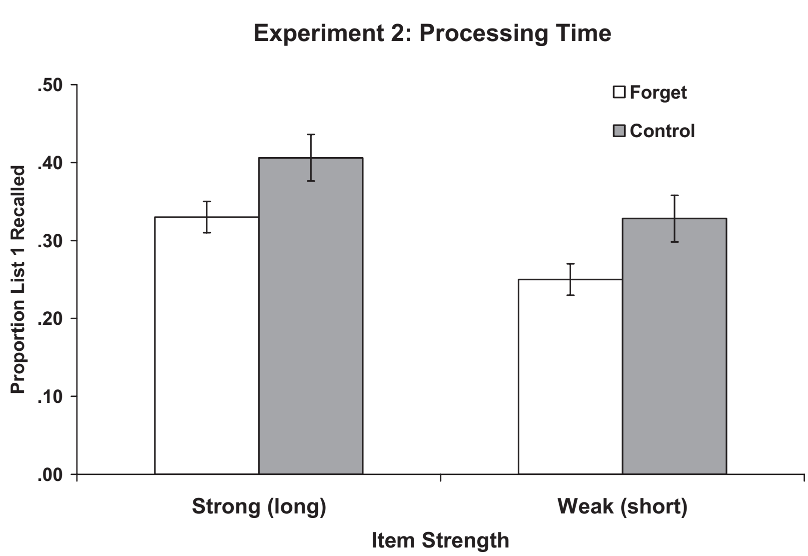

Figure 2. Mean proportion List 1 recall (±*SE*) by cue and item strength in Experiment 2.

and the rest were in the second half of List 1. We created eight versions of the lists that fully counterbalanced (a) whether each word was presented as a spaced or a massed item, (b) whether the first word on the list was a massed or a spaced item, and (c) whether the first presentation of a word received a pleasantness judgment or a "living/non-living" judgment; the second judgment was always the opposite. Eight additional versions of the lists were created by exchanging the first four words of each type (spaced or massed) with the other four words of that type. List 2 items were presented once, with half of the items encoded via pleasantness judgment and the remaining items with living/non-living judgment.

Procedure. The procedure was similar to Experiment 2, except that two different orienting tasks were performed on each word. Participants were told that words might appear more than once. If a smiley-face icon appeared under a word, they were to indicate if it was pleasant or not, and if a yin–yang icon appeared under a word, they had to indicate if it represented a living or non-living thing. All List 1 words were presented twice for 4 s each time. Every word was judged once on each dimension, with the order of the ratings counterbalanced. After rating List 1 words, half of the participants were instructed to try to forget them with the instructions from the previous experiments. Then List 2 words were presented once for 4 s, with half of the words being rated for pleasantness and half for living/non-living. Afterwards, a 90-s surprise recall test was given for List 1, followed by a 90-s recall of List 2 on separate sheets of paper.

Results

A Cue × Item Strength mixed ANOVA on proportion List 1 recall revealed a significant main effect of item strength, $F(1, 94) = 37.92$, $p < .001$, $MSE = .029$, $\eta^2 = .29$, confirming that strong items were better recalled (.51) than weak items (.36). There was also a significant main effect of cue, $F(1, 94) = 9.17$, $p < .01$, $MSE = .032$, $\eta^2 = .09$, with fewer words recalled in the forget group (.40) than in the control group (.48). In addition, the interaction was significant, $F(1, 94) = 4.97$, $p < .05$, $MSE = .029$, $\eta^2 = .05$, showing that there was significant directed forgetting of spaced items, $t(94) = 3.95$, $p < .001$, but not of massed items ($t < 1$; see Figure 3). The lack of forgetting of massed items likely reflects reduced competition in the forget group—fewer strong items were available to compete with the recall of weak items than those in the control group, and hence weak items recovered. Finally, there were no significant differences in List 2 recall ($F < 1$; .27 in forget and .26 in control).

General Discussion

In three experiments, we evaluated the magnitude of directed forgetting following different strengthening manipulations. In each experiment, strong items were better recalled than weak items. Furthermore, while we obtained significant directed forgetting in all three experiments, we observed more forgetting of strong items than weak items only when strength was varied by spacing of presentations (Experiment 3). Depth of processing and longer study time led to equivalent amounts of directed forgetting (Experiments 1 and 2). Given the assumptions of the one-shot hypothesis, these results provide support for the context hypothesis of directed forgetting (Sahakyan & Kelley, 2002).

The one-shot hypothesis of Malmberg and Shiffrin (2005) was motivated by differences in the LSE across various strengthening manipulations. As stated previously, SAM/REM memory models attribute LSE to context effects. An alternative explanation for the LSE was proposed by Bäuml (1997), who argued that the LSE arises from output order biases. Recalling strong items earlier in the output sequence leads to output interference, reducing recall of weak items. When Bäuml (1997) controlled output order in cued recall, LSE was eliminated, but when recall order was unrestricted an LSE emerged. Although the present studies used free recall rather than cued recall, we examined the output order because the results of Bäuml (1997) implied that output order could contribute to the LSE.

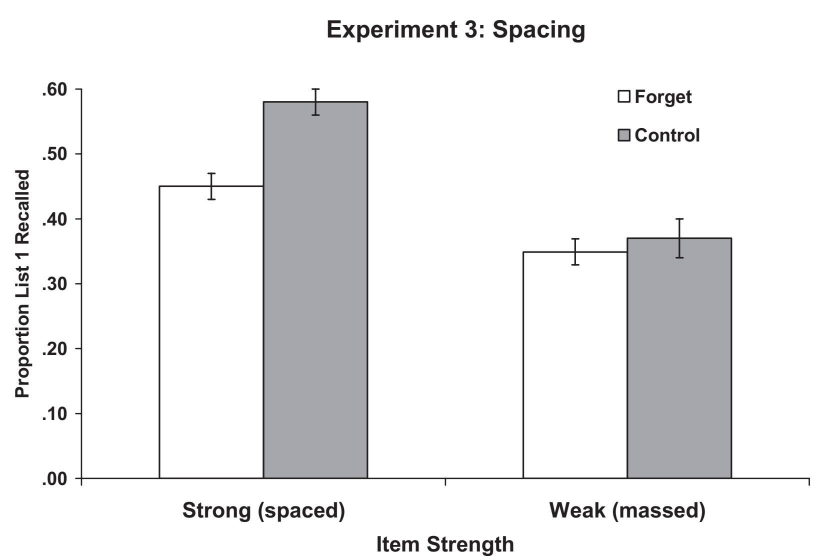

Figure 3. Mean proportion List 1 recall (±*SE*) by cue and item strength in Experiment 3.

We rank ordered the recall of strong and weak items in our experiments and calculated the average output percentile following the procedure recommended by R. A. Bjork and Whitten (1974). Smaller values of output percentile indicate earlier output in the recall sequence. For Experiment 3, we found that the control group recalled strong items earlier (.49) than weak items (.62), $t(45) = 3.25$, $p < .01$. However, in the forget group there was no significant difference in the output position of strong (.53) and weak items (.57; $t < 1$), implying that directed forgetting reduced the output order bias. This could be partly due to greater forgetting of strong items than weak items in the forget group in Experiment 3. Compared with the control condition, in the forget condition fewer strong items were available at the time of recall to cause output interference on weak items. Unlike in Experiment 3, in Experiment 2, there were no significant output order differences in the forget group (.58 for strong and .59 for weak) or in the control group (.57 for strong and .59 for weak; both ts < 1). Similarly, in Experiment 1, there were also no significant differences in the output order in either group, with the highest t value being 1.1.

In Experiment 3—where the task manipulation enhanced the context strength—we observed an output bias favoring strong items in the control group. Otherwise, we observed no preference in output order in the remaining studies (despite having higher recall rates for strong items). If output order was mainly driven by the item strength, we should have observed a similar pattern of output order across all experiments, regardless of the item strengthening operation. However, this was not the case in the current studies. These results suggest that the bias in output order may be driven by context strength, such that items with more context storage tend to be recalled earlier. Although output biases could contribute to the LSE, the current results suggest that output biases may be driven by context strength.

The results of these experiments have implications for the theoretical mechanisms of directed forgetting. Our position has been that the mechanism behind List 1 costs in directed forgetting involves mental contextual change. When participants receive the forget instruction, they stop maintaining List 1 context and instead encode List 2 items with new context cues. The context hypothesis of directed forgetting predicts that items that are strongly linked to their context will suffer more than items that are weakly linked to their context when context changes at the time of test. On the other hand, differences in the item strength alone would be insufficient to create differential amounts of directed forgetting for weak and strong items. The results of our experiments were fully consistent with these predictions.

If different strengthening operations differentially influenced the magnitude of directed forgetting across strong and weak items, then we should expect a three-way interaction when Experiment is included as a variable in the analyses. However, only Experiment 2 and Experiment 3 utilized the same materials and orienting tasks, and they were also run in the same semester. Therefore, it was more appropriate to compare the results across these two studies, especially because the same items (massed items) acted as the strong items in Experiment 2 and as the weak items in Experiment 3. In Experiment 2, they were strong relative to once-presented items because of their greater item strength; in Experiment 3, they were weak relative to spaced items because of their lower context strength and item strength. A Cue (forget vs. control) × Item Strength (strong vs. weak) × Experiment (processing time vs. spacing) ANOVA on proportion List 1 recall revealed that the three-way interaction was approaching significance, $F(1, 156) = 3.55$, $p = .06$, suggesting that directed forgetting was driven by context strength.

Throughout three experiments, we failed to observe enhanced List 2 recall in the forget condition (i.e., no benefits of directed forgetting). Because the reported studies utilized incidental learning, this null effect is consistent with Sahakyan and Delaney's (2005) finding that incidental learning did not lead to improved memory of post-cue items. In line with our previous research, the costs and the benefits of directed forgetting can be observed independently, suggesting that they likely have different underlying mechanisms.

The results from the reported directed forgetting experiments paralleled the findings obtained with the LSE paradigm. Strengthening operations that produced LSE effects also produced differences in the magnitude of directed forgetting. In contrast, strengthening operations that produced null LSE produced no differences in the magnitude of directed forgetting across the manipulations of strength. Taken together, these findings suggest that context mediates directed forgetting. Furthermore, given the parallels between the LSE and directed forgetting, the list-method paradigm could provide a new tool for investigating the effects of context on memory.

References

Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 603–616.

Bäuml, K. H. (1997). The list-strength effect: Strength-dependent competition or suppression? *Psychonomic Bulletin & Review, 4,* 260–264.

Benjamin, A. S. (2006). The effects of list-method directed forgetting on recognition memory. *Psychonomic Bulletin & Review, 13,* 831–836.

Bjork, E. L., & Bjork, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness & Cognition: An International Journal, 5,* 176–196.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger, III, & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309–330). Hillsdale, NJ: Erlbaum.

Bjork, R. A., LaBerge, D., & LeGrand, R. (1968). The modification of short-term memory through instructions to forget. *Psychonomic Science, 10,* 55–56.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6,* 173–189.

Geiselman, R. E., Bjork, R. A., & Fishman, D. L. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. *Journal of Experimental Psychology: General, 112,* 58–72.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–67.

MacLeod, C. M. (1999). The item and list methods of directed forgetting: Test differences and the role of demand characteristics. *Psychonomic Bulletin & Review, 6,* 123–129.

Malmberg, K. J., & Shiffrin, R. M. (2005). The "one-shot" hypothesis for context storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 322–336.

Mensink, G. J. M., & Raaijmakers, J. G. W. (1998). A model of interference and forgetting. *Psychological Review, 95,* 434–455.

Murnane, K., Phelps, M. P., & Malmberg, K. (1999). Context-dependent recognition memory: The ICE theory. *Journal of Experimental Psychology: General, 128,* 403–415.

Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). List-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 163–178.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology, 89,* 63–77.

Sahakyan, L. (2004). Destructive effects of 'forget' instructions. *Psychonomic Bulletin & Review, 11,* 555–559.

Sahakyan, L., & Delaney, P. F. (2003). Can encoding differences explain the benefits of directed forgetting in the list-method paradigm? *Journal of Memory and Language, 48,* 195–201.

Sahakyan, L., & Delaney, P. F. (2005). Directed forgetting in incidental learning and recognition testing: Support for a two factor account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 789–801.

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 1064–1072.

Sheard, E. D., & MacLeod, C. M. (2005). List method directed forgetting: Return of the selective rehearsal account. In N. Ohta, C. M. MacLeod, & B. Uttl (Eds.), *Dynamic cognitive processes* (pp. 219–248). Tokyo: Springer-Verlag.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM–retrieving effectively from memory. *Psychonomic Bulletin & Review, 4,* 145–166.

Smith, S. S., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8,* 203–220.

Tulving, E., & Hastie, R. (1972). Inhibition effects in intralist repetitions in free recall. *Journal of Experimental Psychology, 92,* 297–304.

Received June 15, 2007
Revision received October 8, 2007
Accepted October 18, 2007 ■