Where is the forgetting with list-method directed forgetting in recognition?

LILI SAHAKYAN AND EMILY R. WALDUM
University of North Carolina, Greensboro, North Carolina

AARON S. BENJAMIN
University of Illinois at Urbana-Champaign, Urbana, Illinois

AND

SAMUEL P. BICKETT
University of North Carolina, Greensboro, North Carolina

Despite the fact that list-method directed forgetting instruction leads to decreases in memory performance on tests of free recall, there are to date no published reports of comparable effects in recognition testing. In the present article, we evaluated whether conditions that promote the importance of context in recognition, either via stimulus selection (Experiments 1 and 2) or by test choice (Experiment 3), elicit directed forgetting impairment. In all three experiments, we obtained reliable recognition deficits, suggesting that the typical conditions of recognition, rather than recognition itself, underlie the discrepancy between the tests of recognition and free recall in list-method directed forgetting.

In directed forgetting studies, participants are instructed to either forget or remember certain earlier studied items. Instructions to forget or remember can be delivered either on an item-by-item basis or after an entire list of items has been presented (known as the *item method* and the *list method* of directed forgetting, respectively). Directed forgetting instructions reduce memory for to-be-forgotten (TBF) items as compared with to-be-remembered (TBR) items (for reviews, see E. L. Bjork, Bjork, & Anderson, 1998; MacLeod, 1998; and Johnson, 1994) on tests of free recall and recognition when the item method is used (see, e.g., Basden, Basden, & Gargano, 1993; MacLeod, 1999), but only on recall tests when the list method is used (Basden et al., 1993; Benjamin, 2006; E. L. Bjork & Bjork, 2003; Block, 1971; Conway, Harries, Noyes, Racsma'ny, & Frankish, 2000; Elmes, Adams, & Roediger, 1970; Geiselman, Bjork, & Fishman, 1983; Gross, Barresi, & Smith, 1970; Sahakyan & Delaney, 2005; Whetstone, Cross, & Whetstone, 1996). In the present article, we will argue that the list-method directed forgetting impairment reflects the deleterious effect of a study–test context mismatch, and we will demonstrate that such effects emerge in recognition when stimuli are used or when tests are employed that enhance the utilization of contextual characteristics.

List-Method Paradigm and Basic Findings

In a typical list-method directed forgetting experiment, participants study two lists of items for a memory test. After the first list, half of them are told to forget that list because it was only for practice or because it was the wrong list, and that there is therefore no need to remember those items. The remaining participants are told to keep remembering the list. Both groups proceed to study a second list; then, they are given a memory test for all items (including any items that the participants were instructed to forget). In free recall, the directed forgetting effect consists of two components—the *costs* and *benefits* of directed forgetting. The costs refer to impaired recall of List 1 items in the forget group relative to the remember group, whereas the benefits refer to enhanced recall of List 2 items in the forget group relative to the remember group. Our focus in the present study was the enigmatic absence of directed forgetting costs in recognition, despite the fact that the benefits appear to be robust (see, e.g., Benjamin, 2006; Sahakyan & Delaney, 2005). The present strategy for understanding why the costs are not apparent involved exploring conditions under which those costs could be made to reappear. To do so, we took a lesson from studies of the role of global environmental context in recognition. First, however, we will briefly review how the findings from recognition of intentionally forgotten material are problematic for historical and contemporary theoretical accounts of directed forgetting.

Recognition Testing and Theoretical Accounts of the Directed Forgetting Effect

Single-process accounts of list-method directed forgetting invoke the same underlying mechanism to explain the

---

L. Sahakyan, l_sahaky@uncg.edu

costs and the benefits of directed forgetting, and they have included selective rehearsal (see, e.g., Benjamin, 2006; R. A. Bjork, 1970, 1972; Sheard & MacLeod, 2005) as well as retrieval inhibition (e.g., R. A. Bjork, 1989; E. L. Bjork & Bjork, 1996, 2003; Geiselman et al., 1983). According to the selective rehearsal account, the presentation of a forget cue terminates rehearsal of all preceding items and extra rehearsal is therefore devoted to postcue items, leading to the costs and the benefits, respectively. According to the inhibitory account, the costs of directed forgetting reflect a temporary state of inhibition of the TBF items. The benefits arise from reduced proactive interference acting on postcue items from inhibited TBF items in the forget group.

The dual-process account of directed forgetting relies on two different mechanisms to handle the directed forgetting effect (Sahakyan & Delaney, 2005). The first factor explains the costs by the context-change mechanism of Sahakyan and Kelley (2002). This hypothesis was based on the notion that during encoding, people not only encode the meaning of the item and the interitem relationships, but also encode the context in which the item occurred. Contextual information refers to the environmental, spatial–temporal, emotional, or mental states in which the item was experienced. Sahakyan and Kelley (2002) proposed that directed forgetting instruction encourages participants to change their mental context between the lists, thereby segregating the two lists as separate events. An example of a mental context cue would be the thoughts people experienced during List 1 learning, which in part include how people represent and think about the experimental episode connected with List 1 encoding. According to the contextual change hypothesis, directed forgetting impairment arises because the retrieval context better matches the List 2 encoding context than the List 1 encoding context, producing reduced recall of List 1 items. The second factor of the dual-process account explains the benefits by the strategy-change mechanism (Sahakyan & Delaney, 2003). According to this explanation, the benefits emerge because the forget cue encourages participants to adopt a more elaborate encoding strategy for postcue items than for the remember group. Thus, in the dual-factor account, the costs reflect a mismatch between the study and the test contexts, whereas the benefits reflect a strategic encoding enhancement.

Each of these accounts has difficulty with the finding that the costs emerge in free recall but not in recognition. It is problematic for the selective rehearsal account because both recognition and recall are sensitive to elaborative rehearsal (Geiselman & Bjork, 1980) and recognition is even more sensitive than recall to rote rehearsal (Benjamin & Bjork, 2000; Craik & Watkins, 1973). In other words, regardless of the nature of the rehearsal process involved in directed forgetting, terminating rehearsal of TBF items should have consequences for recognition memory. However, empirical findings to date attest to the contrary. Proponents of the inhibition account explain the lack of directed forgetting effects in recognition tests through the release of inhibition (see, e.g., E. L. Bjork & Bjork, 1996, 2003), by suggesting that re-presenting the TBF items on a

recognition test releases their inhibition, thereby eliminating the directed forgetting effect. The inhibitory account is not fully satisfactory in large part because of the lack of theoretical specificity by which inhibition and release from said inhibition operate. Finally, in the dual-factor account, the contextual explanation handles the absence of directed forgetting costs in recognition by pointing out that many environmental context manipulations do not affect recognition (Fernandez & Glenberg, 1985; Godden & Baddeley, 1980; Humphreys, Pike, Bain, & Tehan, 1988; Jacoby, 1983; Russo, Ward, Geurts, & Scheres, 1999; S. S. Smith, Glenberg, & Bjork, 1978; S. S. Smith, Vela, & Williamson, 1988). However, under certain conditions, environmental context effects do emerge in recognition (Canas & Nelson, 1986; Dalton, 1993; Krafka & Penrod, 1985; Malpass & Devine, 1981; Russo et al., 1999; S. M. Smith, 1985, 1986; S. M. Smith & Vela, 1992) and may lead to the reemergence of costs.

The present experiments employed characteristics that are thought to increase the effects of environmental context on recognition. If such contextual effects truly underlie the costs of directed forgetting, then the costs should be apparent under those conditions.

Environmental Context Effects in Recognition

There is a large body of literature showing that memory suffers when there is a low correspondence between the background context present during study and during test (see S. M. Smith & Vela, 2001, for a meta-analysis of these findings). These context-dependent memory effects have been observed more often in free recall than in recognition tests, in which results have been more inconsistent.

Manipulations of background context may include a change in the *global* context, such as the physical environment, or a change in the *local* context, such as the text color, location of the word on the screen, the voice of a speaker, and so on. The latter is characterized as local because it accompanies the presentation of each item and is typically fast changing. The former is characterized as global because it is slow changing and involves the encoding of an entire set of items in a single environmental context (Glenberg, 1979; S. M. Smith & Vela, 2001). Different types of contextual manipulations have produced different effects in recognition, with local context mostly affecting response strategy (see, e.g., Feenan & Snodgrass, 1990; Murnane & Phelps, 1993, 1995), and global context affecting discriminability—albeit not consistently (for a review, see S. M. Smith & Vela, 2001). Our investigation was guided principally by the studies of global environmental context in recognition, because such a conceptualization of context is more similar to the notion of Sahakyan and Kelley's (2002) context change involved in directed forgetting. If directed forgetting instigates a mental context change after an entire set of items have been studied, then this type of context change is better described as the global list-wide context change rather than as a local context change that accompanies the presentation of each item.

The first two experiments reported in the present article were motivated by the observation of S. M. Smith

and Vela (2001)—that global environmental context effects in recognition were larger when encoding processes were primarily nonassociative, especially when the study materials involved unfamiliar faces or nonwords (Dalton, 1993; Krafka & Penrod, 1985; Malpass & Devine, 1981; Russo et al., 1999; S. M. Smith & Vela, 1992). More specifically, our first experiment was inspired by the study of Russo et al. (1999), the design of which is similar to the list-method directed forgetting paradigm; therefore, we will discuss it in greater detail.

Participants encoded two lists of nonwords in two different rooms (A and B). After a brief delay in a waiting room, they returned to Room A and received a yes/no recognition test on half of the nonwords studied in Room A, along with half of the nonwords studied in Room B, plus an equal number of unstudied nonwords. Then, they went to Room B and completed the recognition test on the remaining nonwords studied in Room B or Room A, along with new distractors. Russo et al. (1999) reported reduced discrimination accuracy for nonwords tested in the mismatching context. A similar experiment failed to obtain the effect of context when the study materials consisted of words.

Russo et al. (1999) proposed several hypotheses to explain their findings, and we will revisit them in part because they motivated our directed forgetting studies with nonwords. Broadly speaking, these explanations emphasized the association between the item and its context established during encoding, with nonwords forming stronger item-to-context associations than words. This could have occurred because prior history of occurrence of words in multiple contexts might have created contextual competition and have led to less specific contextual information being encoded in the memory trace when the words were processed in a specific experimental context. Thus, context cues might be better retrieval cues for nonwords than for words because nonwords have no prior "real world" contextual associations to interfere with experimental item-to-context associations. Research shows that concepts experienced in fewer preexperimental contexts suffer more when environmental context changes between the study and test, presumably because they form stronger item-to-context associations during encoding than do concepts that appear in many preexperimental contexts (Marsh, Meeks, Hicks, Cook, & Clark-Foos, 2006). Similarly, both recognition (Gorman, 1961) and source memory (Guttentag & Carroll, 1997) are superior for uncommon than for common words, which tend to be experienced in a lower variety of contexts (Marsh, Cook, & Hicks, 2006; Steyvers & Malmberg, 2003).

Another related explanation proposed by Russo et al. (1999) is that context associated with nonwords may act differently from the context associated with words. This reasoning relies on Baddeley's (1982) distinction between independent and interactive contexts. The former refers to context being independently associated with the item without significantly modifying its meaning, whereas the latter refers to context being integrated with an item and changing its meaning (e.g., the meaning of *jam* is different in the context of *strawberry* and *traffic*). Because nonwords do not have a stable meaning, context could become interactively encoded with them and elicit a temporary meaning, whereas it may be independently associated with words and have no detrimental effect on the perception of their meaning. A change in interactive context could reduce discrimination accuracy of nonwords because the meaning elicited at test may not match the meaning elicited by context during encoding (see, e.g., Light & Carter-Sobell, 1970; Tulving & Thomson, 1973). In contrast, changes in independent context are not typically hypothesized to change semantic meaning; hence, words may be more resistant to context effects. Note that when items are deliberately integrated with incidental context (even without being modified interactively by their context), they become more vulnerable to changes of context than they do when these two sources of information are not integrated with each other and are instead independently encoded (see, e.g., Murnane, Phelps, & Malmberg, 1999).

Finally, in addition to considering the processes engaged during encoding of nonwords, it is also important to discuss retrieval processes involved in recognition judgments because they will play a role during the test. Using the dual-process framework of recognition (see, e.g., Jacoby, 1991; Mandler, 1980), some researchers have argued that familiarity-based recognition is unaffected by environmental context changes, whereas recognition accompanied by conscious recollection is impaired when the context changes at test (e.g., Macken, 2002). Familiarity-based recognition might be sufficient to recognize words, but discriminating nonwords might require the retrieval of more specific details (e.g., specific letter-order combinations) because nonwords have fewer alternative bases for recognition due to their having less semantic content. If context change hurts retrieval-based recognition more than it does familiarity-based recognition, nonwords should show more directed forgetting than should words.

The present experiments employed manipulations of stimulus type (words vs. nonwords) and the particular demands of the recognition test (plurality discrimination vs. standard recognition) to see whether the costs of directed forgetting emerge under circumstances that increase the significance of contextual information.

EXPERIMENT 1

We implemented the list-method directed forgetting design, but had participants study two lists of nonwords rather than of words. Memory was tested with an old–new recognition test.

Method

Participants. Fifty-six University of North Carolina, Greensboro, undergraduates participated in the experiment in exchange for course credit. They were tested individually. Half of them were assigned to the forget group, and the remaining half were in the remember group.

Materials. The stimuli were 112 nonwords that were created by a group of undergraduates for a related classroom project (the stimuli are shown in the Appendix, Table A1). They were variations on Bermudian slang words that were altered to reduce their similarity to English words. In pilot testing, a separate group of 18 undergradu-

ates were given the list of nonwords and asked to write down the first English word that came to mind in response to the nonwords. If a given nonword elicited the same association in two or more people, it was not included in the pool of final stimuli.

From a pool of 112 items, we constructed four lists of 28 items (Lists A, B, C, and D). For half of the participants, Lists A and B served as targets and Lists C and D served as distractors, whereas for the remaining participants, Lists C and D served as targets and Lists A and B served as distractors. Each target list was assigned equally often to appear as the first or the second study list. The presentation order of the items within each study list was randomized for each participant. Test lists were constructed by randomly mixing the target items from both lists with an equal number of distractors from the unstudied lists.

Design. The design of the study was a cue (forget vs. remember) $\times$ list (List 1 vs. List 2) mixed factorial design, with lists as the within-subjects factor.

Procedure. Participants were given two lists of 28 nonwords per list to study for an unspecified memory test. They were told to read the nonwords aloud and to try to memorize them. Items were presented one at a time on the computer screen at a rate of 4 sec per item. After encoding List 1, half of the participants were told that the first list was just for "practice," and that there was no need to remember it. They were told to try to forget it. The remaining participants were told to remember that list because it was only "the first half of the items." Then all participants studied List 2, which was followed by a remember instruction for all participants.

Upon completion of the study phase, participants received an old–new recognition test. They were told that they would be presented with items, some of which they had learned during the experiment, and some of which would be new and had not been presented in the experiment. They were told to press OLD if they remembered studying the nonword on either List 1 or List 2 (in the forget group, the instructions referred to the lists as the "practice" list or the "experimental" list). Otherwise, they were told to press NEW if they did not remember studying the item in the experiment. For each participant, the computer generated a test list by randomly mixing all 56 nonwords from List 1 and List 2 with 56 new distractors.

Results and Discussion

We replaced 4 participants whose hits rates were below false alarm rates. The hits for List 1 and List 2 items as well as the false alarms are reported in Table 1, along with statistical comparisons across the forget and the remember conditions. The finding that neither false alarms nor List 2 hits differed between the forget and the remember conditions suggests that participants in these two groups did not employ different decision strategies. In this section, we will report the analyses of the discrimination accuracy (d').

In calculating d' scores, hits and false alarms were calculated by adding .5 to each frequency and dividing by $N + 1$, where N represents the number of old items (for calculating hits) or new items (for calculating false alarms). This procedure was recommended by Snodgrass and Corwin (1988) to avoid

estimation problems arising from hit rates of 1 and false alarm rates of 0. We conducted a cue (forget vs. remember) by lists (List 1 vs. List 2) mixed factorial ANOVA on d' scores.

There was a significant list $\times$ cue interaction [$F(1,54) =$ 4.27, $MS_e = .098$, $p < .05$]. Follow-up tests showed that there was a significant difference in discrimination accuracy of List 1 items between the forget and the remember groups [$t(54) = 2.91$, $p < .05$]. However, there were no differences between the groups on List 2 items ($t < 1$). In other words, we observed the costs, but not the benefits of directed forgetting with nonword stimuli.

This is the first demonstration of significant memory impairment in directed forgetting using a recognition test. Prior studies have failed to obtain this effect with word stimuli, whereas the use of nonwords allowed us to demonstrate directed forgetting costs in recognition. Given the importance of such findings, we aimed to replicate the costs with more "conventional" nonword stimuli in Experiment 2, because the nonwords used in Experiment 1 were created in our lab, and it is possible that our findings were specific to those stimuli.

Two prior studies reported recognition benefits, but they did not obtain the costs (Benjamin, 2006; Sahakyan & Delaney, 2005). In contrast, Experiment 1 revealed the opposite: significant costs, but no benefits with nonwords. If the directed forgetting benefits arise from more elaborative encoding strategies employed by the forget group (Sahakyan & Delaney, 2003), then the absence of benefits with nonwords may be due to difficulty applying such strategies to nonassociative stimuli, which are considerably harder to interrelate and link to each other than are words. However, because such an explanation is post hoc, we sought to evaluate it in Experiment 2 by including a word condition in which we anticipated obtaining the benefits.

EXPERIMENT 2

In Experiment 2, some participants learned nonwords, whereas others learned words in order to allow direct comparison across the stimulus type. The nonwords were drawn from prior published studies (McCann & Besner, 1987). We also varied the length of the study lists, with some participants studying two short lists (16 items/list) and some participants studying two long lists (36 items/list).

There were reasons to include a list-length manipulation; specifically, in Experiment 1, the study lists included 28 nonwords/list, which is longer than the lists used in

Table 1
List-Specific Hits, False Alarms (FAs), List-Specific Discriminability Statistic (d'),
and Response Bias (C_L) Across the Forget and Remember Groups in Experiment 1

	Measures											
	List 1 Hits		List 2 Hits		FA		d' List 1		d' List 2		Bias (C_L)	
Cue	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE
Forget	.68	.02	.70	.02	.36	.03	0.91	0.07	0.99	0.10	0.40	0.09
Remember	.73	.02	.67	.02	.31	.02	1.21	0.08	1.02	0.09	0.55	0.09
Significance	$p = .08$		n.s.		n.s.		$p < .05$		n.s.		n.s.	

the majority of the list-method directed forgetting studies. Therefore, it could be that impaired recognition with nonwords was mediated by the length of the lists. When Sahakyan and Delaney (2005) first documented directed forgetting in recognition (they found the benefits, but not the costs), they noted that the benefits were obtained with long lists but not with short lists. Given that Experiment 1 was the first demonstration of directed forgetting *costs* in recognition, it could be that list length in general is involved as a factor in recognition differences in directed forgetting. By including a short-list condition in this experiment, we would be able to evaluate whether the costs would also emerge with 16-item lists of nonwords, which is a more common list length for the list-method directed forgetting studies. In the long-list condition, we expected to replicate the costs with nonwords that we obtained in Experiment 1. In addition, we wanted to further examine whether the absence of benefits obtained in Experiment 1 was specific to nonword stimuli. If the difficulty of using elaborative encoding with nonwords is what precluded us from obtaining the benefits in the previous experiment, then the directed forgetting benefits should be observed in a condition with words, which can be more elaborately encoded. To maximize the chances of obtaining the benefits with words, we included the long-list condition because prior studies examining the benefits in recognition have also detected them with longer lists (see, e.g., Benjamin, 2006; Sahakyan & Delaney, 2005). One explanation for why longer lists contribute to detecting the benefits is that longer lists might create a larger list-length effect in the remember group than in the forget group, because increasing the list length impairs recollection (e.g., Cary & Reder, 2003; Yonelinas & Jacoby, 1994). However, if List 2 items are better encoded in the forget group, then more elaborate encoding can provide support for recollection and offset the list-length effect in the forget group. To summarize, on the basis of prior studies, we expected to obtain the benefits of directed forgetting with long lists in the word condition, and we wanted to examine whether they would also emerge with the nonwords.

Method

Participants. A total of 144 University of North Carolina, Greensboro, undergraduates participated in this experiment in exchange for course credit. None of them had participated in Experiment 1. They were tested individually.

Materials. The stimuli were 144 words (mean frequency, 60.7 per million; Kučera & Francis, 1967) and 144 nonwords. Nonwords were taken from the study of Russo et al. (1999), who in turn selected their stimuli from the materials created by McCann and Besner (1987). The word stimuli are shown in the Appendix (see Table A2).

For each participant, two study lists of pure words or pure nonwords were constructed. The study lists were either short or long, with short lists containing 16 items and long lists containing 36 items/list. From a pool of 144 items (words or nonwords), we constructed four lists of 36 items (Lists A, B, C, and D). For half of the participants, Lists A and B served as targets and Lists C and D served as distractors, whereas for the remaining participants, Lists C and D served as targets and Lists A and B served as distractors. Each target list was assigned equally often to appear as the first or the second study list. The presentation order of the items within each list was randomized for each participant. Participants in the long-lists condition studied

the entirety of Lists A and B (or Lists C and D) during the study session and received a test list containing 72 studied items that were randomly mixed with equal numbers of distractors from the unstudied lists. Participants in the short-lists condition studied two lists of 16 items that were randomly selected from Lists A and B (or Lists C and D). The short recognition test lists contained 32 studied items mixed with 32 distractors selected from the unstudied lists. Half of the distractors were randomly drawn from one unstudied list, and the remaining half were drawn from the second unstudied list.

Procedure. The encoding and the testing procedures followed those of Experiment 1.

Design. The design was a cue (forget vs. remember) × list (List 1 vs. List 2) × item type (words vs. nonwords) × list length (short vs. long) mixed factorial, with list as the only within-subjects factor.

Results and Discussion

We replaced 2 participants in the nonword condition whose hit rates were lower than the false alarm rates. The hit rates for List 1 and List 2 items as well the false alarm rates are reported in Table 2. In this section, we will focus on the analyses of the discrimination accuracy (d') and response bias (C_L).

Discrimination accuracy. In calculating d' scores, hits and false alarms were transformed according to the procedures recommended by Snodgrass and Corwin (1988), which were described earlier. We conducted a mixed factorial ANOVA on d' scores using cue (forget vs. remember), list length (short vs. long), item type (words vs. nonwords), and list (List 1 vs. List 2). The results are summarized in Table 3.

There was a main effect of list length [$F(1,136) = 24.17$, $MS_e = 0.657$, $p < .001$], with better discriminability of short lists (2.12) than of long lists (1.65). There was also a main effect of item type [$F(1,136) = 76.08$, $MS_e = 0.658$, $p < .001$], with better discriminability of words (2.30) than of nonwords (1.47). There was a significant list × cue interaction [$F(1,136) = 5.55$, $MS_e = 0.112$, $p < .05$], signifying a directed forgetting effect. However, this effect depended both on the item type and on the length of the list, as was evidenced by a significant four-way interaction [$F(1,136) = 5.90$, $MS_e = 0.112$, $p < .05$]. The remaining effects were not significant. To follow up the interaction, we analyzed the words and the nonwords conditions separately.

Words. A list × cue × list length ANOVA on d' scores revealed a significant effect of list length [$F(1,68) = 8.18$, $MS_e = 0.972$, $p < .01$], confirming better discriminability of short lists (2.54) than of long lists (2.07). None of the remaining main effects were significant ($Fs < 1$), and neither was the cue × length interaction ($F = 1.13$). However, there was a significant three-way interaction [$F(1,68) = 7.22$, $MS_e = 0.123$, $p < .05$], indicating that the directed forgetting effect varied as a function of list length. The results are displayed in Figure 1.

To follow up the interaction, we examined each list-length condition separately. There was no directed forgetting with short lists: There was neither a list × cue interaction [$F(1,34) = 1.59$, $p = .22$] nor a main effect of cue ($F < 1$) nor a main effect of list ($F < 1$). In other words, there were neither directed forgetting costs nor directed forgetting benefits. In contrast, in the long-list condition,

Table 2
List-Specific Hits and False Alarms (FAs) by Item Type, Cue, and List Length in Experiment 2

| | | Hits List 1 | | | | Hits List 2 | | | | FA | | | |
| | | Words | | Nonwords | | Words | | Nonwords | | Words | | Nonwords | |
List Length	Cue	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE
Short 16-items	Forget	.86	.02	.71	.03	.82	.03	.74	.03	.11	.02	.21	.03
	Remember	.83	.03	.83	.03	.82	.04	.78	.03	.10	.02	.21	.02
Significance		n.s.		$p < .05$		n.s.		n.s.		n.s.		n.s.	
Long 36-items	Forget	.77	.03	.69	.03	.81	.04	.66	.03	.12	.02	.25	.02
	Remember	.79	.02	.75	.02	.74	.03	.65	.04	.15	.02	.24	.01
Significance		n.s.		$p = .09$		$p = .12$		n.s.		n.s.		n.s.	

Note—Statistical effects for List 1: item type, $p < .001$; list length, $p < .01$; cue, $p < .05$; item type × cue, $p < .05$. For List 2: item type, $p < .001$; list length, $p < .01$; list length × cue, $p = .14$. For false alarms: item type, $p < .001$; list length, $p < .05$.

there was a significant list × cue interaction [$F(1,34) = 6.68$, $MS_e = 0.115$, $p < .05$]. Follow-up tests revealed that there were no differences in discrimination of List 1 items between the forget and the remember groups ($t < 1$), implying that there were no costs of directed forgetting in the long lists. However, there was a significant difference in List 2 discrimination [$t(34) = 2.05$, $p < .05$], indicating directed forgetting benefits.

Nonwords. A list × cue × list length ANOVA in the nonwords condition revealed a main effect of list length [$F(1,68) = 23.19$, $MS_e = 0.343$, $p < .001$], indicating better discriminability of the short lists (1.70) than of the long lists (1.23). There was also a main effect of cue [$F(1,68) = 4.19$, $MS_e = 0.343$, $p < .05$] and a main effect of list [$F(1,68) = 4.96$, $MS_e = .101$, $p < .05$], which were qualified by a significant cue × list interaction [$F(1,68) = 6.62$, $MS_e = 0.101$, $p < .05$]. Follow-up comparisons revealed significant differences in the discrimination of List 1 items between the forget group (1.34) and the remember group (1.69) [$t(70) = 2.84$, $p < .01$], indicating directed forgetting costs with nonwords. However, there was no significant difference in the discrimination of List 2 items between the forget and the remember groups ($t < 1$). In other words, there were no directed forgetting benefits. These findings did not vary as a function of list length. There was neither a three-way interaction ($F < 1$) nor a two-way interaction involving the list length ($F = 1.18$). To summarize, in the nonwords condition, there were directed forgetting costs both with the short lists and the long lists. However, there were no benefits in either list length.

Response bias (C_L). Next, we analyzed the response bias measure, C_L (Rotello & Macmillan, 2008). Higher numbers in this measure indicate more conservative responding. A cue × item type × list length ANOVA on C_L

revealed a significant main effect of item type [$F(1,136) = 46.57$, $MS_e = 0.204$, $p < .001$], indicating a more conservative response bias in the words condition (1.33) than in the nonwords condition (0.81). There was also a significant main effect of list length [$F(1,136) = 7.29$, $MS_e = 0.204$, $p < .01$], indicating a more conservative bias in the short-list condition (1.17) than in the long-list condition (.97). There was neither a main effect of cue ($F < 1$) nor any interactions (all Fs < 1). To summarize, participants set a more conservative response criterion in the conditions in which the discrimination task was easier, as is typical (e.g., in short lists; with word stimuli; Benjamin, 2003; Benjamin & Bawa, 2004). However, directed forgetting did not affect response bias in any of the experimental conditions.

Summary. The results showed that directed forgetting impaired discriminability of List 1 items and that the results were not due to response bias. However, impaired recognition was found only with nonword stimuli, irrespective of the list length. Word stimuli, on the other hand, did not lead to impaired recognition, confirming the findings of earlier directed forgetting studies. Because directed forgetting lowered the hits without affecting the false alarms in the nonword condition, significant differences in the discriminability of List 1 items emerged between the forget and the remember groups. In contrast, in the word condition, directed forgetting did not impair discriminability: Neither the hits nor the false alarms were lowered. Finally, despite impaired recognition of List 1 items in the nonword condition, there was no associated memory enhancement of List 2 items. The benefits of directed forgetting emerged only with words, and only with long lists. We will defer further discussion of these findings to the General Discussion section.

Table 3
Discriminability and Response Bias by Item Type, Cue, and List Length in Experiment 2

| | | d' List 1 | | | | d' List 2 | | | | C_L Response Bias | | | |
| | | Word | | Nonword | | Word | | Nonword | | Word | | Nonword | |
List Length	Cue	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE
Short 16-items	Forget	2.58	0.15	1.48	0.14	2.43	0.17	1.58	0.15	1.39	0.13	92	0.12
	Remember	2.54	0.18	1.95	0.11	2.59	0.23	1.72	0.10	1.47	0.13	.91	0.11
Long 36-items	Forget	2.10	0.19	1.20	0.09	2.30	0.20	1.12	0.09	1.30	0.11	.69	0.08
	Remember	1.99	0.13	1.43	0.06	1.86	0.16	1.15	0.10	1.15	0.11	.73	0.05

Discriminability in the Word Condition
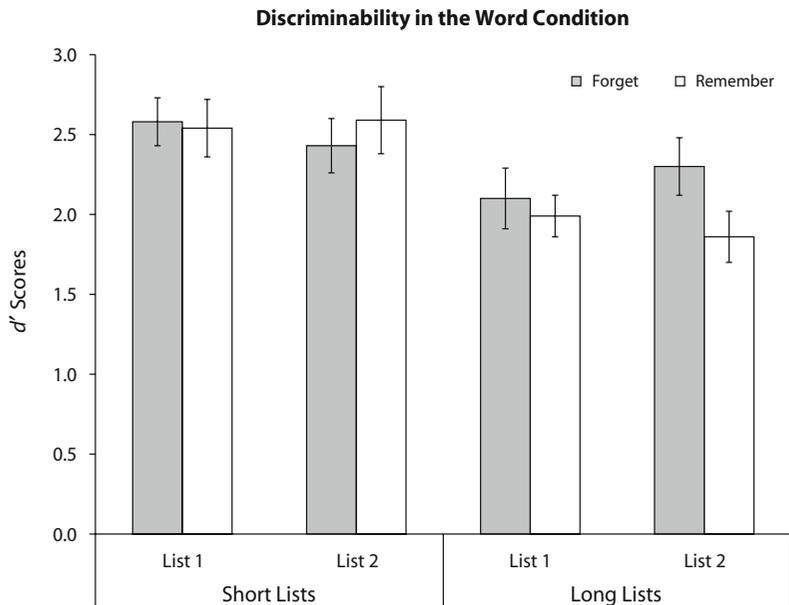
Figure 1. Discriminability scores (d') as a function of list length and cue in the word condition of Experiment 2. Error bars represent *SE*s of the means.

EXPERIMENT 3

Across two experiments, we detected directed forgetting costs with nonwords. However, consistent with published findings in directed forgetting, we did not observe recognition impairment with words.

It could be that the reason we obtained the effect with nonwords was that episodic associations established between the nonwords and their context played a larger role in the recognition of nonwords. One reason why this could occur is because it is harder to create interitem associations with nonwords than with words. Therefore, it could be that when interitem associations were minimized, nonwords relied more heavily on context because they have fewer alternative bases for recognition. However, it could also be the case that directed forgetting with nonwords emerged because nonwords were more similar to their distractors than were the words, requiring more retrieval-based processes to discriminate them from each other. For example, Greene (2004) showed that participants rated nonwords to be more similar than words to the study list (comprised of a mixture of different nonwords and words). Nonwords may evoke a greater feeling of similarity to each other because they have less semantic content, and their structural features, such as letters and phonemes, may give rise to a greater feeling of similarity than they would when the assessment of similarity is mediated by meaning. In contrast, the semantic features of words make them more distinctive and unique than nonwords, and this may reduce their overall assessment of similarity. If nonwords are perceived to be more similar to each other than the words are, then in order to discriminate those from each other, more direct retrieval may be engaged than in recognition of words, which could be discriminated on the basis of familiarity alone. Participants may attempt to retrieve more specific details, such as whether they have seen a particular letter combination in that order and whether that combination was at the beginning, the middle, or the end of a nonword, instead of relying on familiarity, which may be less diagnostic.

Prior studies with words have shown that when targets and distractors are perceptually similar, direct retrieval strategies are often engaged during discrimination (see, e.g., Arndt & Reder, 2002; Rotello, Macmillan, & Van Tassel, 1999). Thus, we suspected that in order to observe directed forgetting in recognition, the test may need to engage more direct retrieval than has been typically captured by prior directed forgetting studies with words. Benjamin (2006) noted that familiarity-based recognition may be used when it can lead to acceptable levels of performance, and it may be unaffected by directed forgetting, whereas retrieval-based recognition may be impaired. However, many prior directed forgetting studies with words have created situations that were conducive for familiarity-based recognition, and more direct retrieval conditions are likely needed to obtain directed forgetting in recognition.

If direct retrieval was contributing to the directed forgetting impairment with nonwords, then we should be able to replicate these findings by using word stimuli and creating a test that places a greater demand on retrieval. One way to do this would be to manipulate the similarity of distractors to the targets. The presence of similar distractors on the test often elicits a *recall-to-reject process*, which capitalizes on the information about the recalled item to reject a similar distractor that could not be directly retrieved (see,

e.g., Rotello et al., 1999). For example, if during encoding participants studied *cloud* and on the recognition test were shown *clouds*, they might retrieve *cloud* from memory and use it to reject *clouds*.

In Experiment 3, we required participants to discriminate the targets from either similar or dissimilar distractors. Similar distractors were plurality-reversed versions of the target words (Hintzman & Curran, 1994), whereas dissimilar distractors were completely new words. We hypothesized that the plurality discrimination task would engage more direct retrieval during the test, thereby enabling us to detect directed forgetting, even with word stimuli.

Method

Participants. Ninety-six University of North Carolina, Greensboro, undergraduates participated in this experiment in exchange for course credit. None of them had participated in Experiments 1 or 2. They were tested individually.

Materials. The stimuli were 60 common nouns (mean frequency, 53.7/million; Kučera & Francis, 1967) and their plurals (e.g., *bean–beans*). The words are shown in the Appendix, Table A3. We chose only regular nouns so that their plural form could be created by adding an *s*, and so that the plural form would be of approximately the same word frequency as the singular form of the word. From the set of 60 singular–plural pairs, 40 pairs were used to create the study lists, whereas the remaining 20 pairs served as distractors in one of the conditions of the recognition test. Forty pairs were further divided into two sets of 20 pairs (Sets A and B). Each set was used to create two different study lists by selecting 10 singular items and 10 plural items from the word pairs in each set. Only a singular or a plural form of a word was selected from the word pair, not both. The singular–plural aspect of the selected words was counterbalanced. Thus, there were a total of four study lists (A1, A2 and B1, B2) containing 20 different items, half of which were singular words, and the other half of which were plural words. Each list was assigned equally often to serve as the first or the second study list. If A1 or A2 served as the first study list, then the second study list was either a B1 or B2 list (or vice versa). Two different lists constructed from the same set (e.g., A1 and A2) were never presented together. The presentation order of the words within each study list was randomized for each participant.

Memory was tested with an old–new recognition test. There were two testing conditions, varied between subjects. The plurality discrimination condition required participants to discriminate the studied words from their plurality-changed versions. The new discrimination condition required discriminating the studied words from the completely new distractors that were not presented during encoding.

Eight different test lists were constructed for each test condition. In the plurality discrimination condition, the old words consisted of half of the items from each study list (5 singulars and 5 plurals selected from each list). The plurality-changed versions of the remaining study list items were used as the distractors. Thus, the test lists contained 40 items, 20 of which were old words (10 from List 1 and 10 from List 2), and 20 of which were distractors that were plurality-changed versions of the remaining 10 words from List 1 and 10 words from List 2. In the new distractor condition, the test lists were created by the same principle, except that all plurality-changed distractors were replaced with completely new words (matched in singular–plural form to the plurality-changed distractors). The new distractors came from the set of 20 pairs that were not used during the study.

Thus, participants in both recognition conditions had the same study list (e.g., *rivers*, *book*, *town*, *windows*, etc.), but received a different test list depending on the condition. In the plurality discrimination condition, they were tested with *rivers*, *books*, *town*, and *window*, whereas in the new distractor condition, they were tested with *rivers*, *beans*, *town*, and *napkin*.

Design. The design of the study was a cue (forget vs. remember) × list (List 1 vs. List 2) × discrimination condition (plurality vs. new distractor) mixed factorial, with list as the only within-subjects factor.

Procedure. The encoding phase was similar to the word condition of Experiment 2, except that the study lists always contained 20 items. All participants were told to pay attention to the plurality of each item during encoding because it was going to be important later on. For participants in the new distractor condition, this was a minor deception, since it was not going to be helpful during the test. However, it was implemented in order to keep the encoding conditions constant. No mention was made of how memory was going to be tested.

After studying List 2, participants received an old–new recognition test. Half of them were given a plurality discrimination test list, whereas the remaining half received a new distractor test list. Participants were told to press OLD if they remembered studying the word on either list (including on the "practice" list, for the forget group participants), and to press NEW if they thought the word was not on either list. Participants in the plurality discrimination condition were told that some test items might be the plurality-changed versions of the study words, and that they had to respond positively only to the studied items and to reject their plurality-changed versions. They were told that only the singular or plural form of the word would appear on the test, but never both. Participants in the new discrimination condition were simply told to discriminate the studied items from unstudied items. An example was provided in both testing conditions.

Results and Discussion

The hits and the false alarms are presented in Table 4. Next, we will report the analyses of the discrimination accuracy and response bias.

Discrimination accuracy. Using a mixed factorial ANOVA, we analyzed d' scores with cue (forget vs. remember) × discrimination condition (plurality vs. new distractor) × list (List 1 vs. List 2) as factors. The results are summarized in Table 5. There was a main effect of discrimination condition [$F(1,92) = 49.05$, $MS_e = 1.31$, $p < .001$], indicating better accuracy in the new discrimination condition (2.14) than in the plurality discrimination condition (0.96). There was also a list × cue interaction [$F(1,92) = 6.79$, $MS_e = 0.313$, $p < .01$], which was further modified by a significant three-way interaction [$F(1,92) = 5.30$, $MS_e = 0.313$, $p < .05$]. We followed up this interaction with a separate list × cue ANOVA in each test condition.

When participants had to discriminate the studied items from the new distractors, there were neither main effects nor an interaction (all Fs < 1). In other words, there was no directed forgetting when the distractors were new items. In contrast, when participants had to make plurality discriminations, there were no main effects (both Fs < 1), but there was a significant list × cue interaction [$F(1,46) = 8.29$, $MS_e = 0.455$, $p < .01$], implicating a directed forgetting effect. Follow-up tests revealed significant differences between the forget group and the remember group on List 1 items [$t(46) = 2.03$, $p < .05$]. However, there were no differences on List 2 items between the forget group and the remember group ($t < 1$). To summarize, in the new distractor condition, there was no directed forgetting effect: Neither the costs nor the benefits were significant. In contrast, in the plurality

Table 4
List-Specific Hits and False Alarms (FAs) by Cue and Discrimination Condition in Experiment 3

Condition	Cue	List 1 Hits		List 1 FAs		List 2 Hits		List 2 FAs		Pooled FAs	
		M	SE	M	SE	M	SE	M	SE	M	SE
Plurality discrimination	Forget	.70	.04	.44	.04	.80	.03	.43	.04	.39	.03
	Remember	.80	.03	.35	.04	.73	.03	.45	.04	.35	.03
Significance		$p < .05$		$p = .15$		$p = .18$		$t < 1$		$t < 1$	
New distractors	Forget	.81	.03	–		.84	.03	–		.14	.03
	Remember	.79	.03	–		.79	.03	–		.12	.03
Significance		$t < 1$				$p = .18$				$t < 1$	

Note—List-specific FAs in the plurality discrimination condition represent "yes" responses to distractor items, which were plurality-reversed versions of half of the List 1 words or List 2 words. Only pooled FAs are reported in the new distractor condition, because all distractors were new and unrelated to the study-list items.

discrimination condition, there were significant directed forgetting costs, but no benefits.

Response bias. We analyzed the response bias statistic, C_L, as a function of the cue and discrimination condition. Response bias was calculated from the pooled false alarms because there were no list-specific false alarms in the new distractor discrimination condition. There was only a main effect of discrimination condition [$F(1,92) = 61.97$, $MS_e = 0.305$, $p < .001$], indicating a more liberal response criterion in the plurality discrimination condition (0.36) than in the new distractor condition (1.25). Neither the main effect of cue nor the interaction were significant ($Fs < 1$). In other words, directed forgetting did not influence the response bias.

GENERAL DISCUSSION

In three experiments, we reported significant directed forgetting impairment in recognition accuracy by using either stimulus manipulations or distractor manipulations. In Experiments 1 and 2, we found significant differences between the forget and the remember groups in the discriminability of nonwords. No such impairment was observed with words in Experiment 2, or in prior work (Benjamin, 2006; Sahakyan & Delaney, 2005). Interestingly, these experiments reveal a double dissociation between stimulus type and the effects of directed forgetting: Memory for nonwords exhibits costs but no benefits, and memory for words exhibits benefits but no costs (the latter occurs only with longer lists; short lists produce neither the costs nor the benefits of directed forgetting). We believe that the difference in costs across stimuli reveals the greater effect of context on the recognition of nonwords (Russo et al., 1999) and that the difference in benefits demonstrates the diffi-

culty of engaging in elaborative encoding for stimuli that are devoid of meaning, such as nonwords. However, it is the former finding that is of central interest here. The claim that circumstances promoting contextual effects in recognition contribute to the emergence of directed forgetting costs also finds support in the results of Experiment 3. In that study, the costs of directed forgetting were found even with words when the distractors were similar to the targets and required more fine-grained discrimination. The occurrence of directed forgetting impairment with words in Experiment 3 can be attributed solely to the processes solicited by testing, because all participants studied the same study list but received a different test list requiring discrimination from plurality-reversed distractors or new distractors. Thus, the encoding conditions were kept constant and any differences that emerged between experimental conditions must have been due to the retrieval processes.

Overall, conditions that rely on precise recovery of details from the study episode appear to promote the emergence of directed forgetting impairment in recognition. Other recent findings provide additional support for this hypothesis. For example, Loft, Humphreys, and Whitney (2008) demonstrated that intentionally forgotten materials interfere more with ongoing task performance under exclusion than under inclusion conditions. Also, Gottlob and Golding (2007) found directed forgetting in source monitoring, where in addition to identifying the old–new status of the words, participants had to retrieve their case/color/list membership. In their study, directed forgetting affected memory for item details (e.g., color/case) more consistently than memory for the item itself. Taken together, these findings underscore the claim that engaging in more effortful retrieval processes at the time of test contributes to detecting directed forgetting impairment in recognition.

Table 5
Discriminability and Response Bias by Cue, Study List, and Discrimination Condition in Experiment 3

Condition	Cue	Discriminability d'				Response Bias C_L	
		List 1		List 2			
		M	SE	M	SE	M	SE
Plurality discrimination	Forget	0.72	0.20	1.05	0.18	.30	0.08
	Remember	1.26	0.18	0.80	0.18	.42	0.10
New distractors	Forget	2.12	0.19	2.19	0.19	1.20	0.13
	Remember	2.06	0.17	2.08	0.17	1.30	0.13

This is true both when attempting to retrieve the details about the item characteristics (such as their appearance) and in the exclusion instruction condition, where participants have to make a fine-grained discrimination at test by avoiding responding to the words from a certain list.

A number of related mechanisms could underlie the directed forgetting impairment with nonwords, and they all stem from the consideration of contextual factors. For example, nonwords could have relied more on context because they have no alternative bases for recognition. Because interitem associations are harder to establish with nonwords than with words, contextual cues could play a larger role for the recognition of nonwords. Also, nonwords could have formed stronger item-to-context associations during encoding because they have no prior history of occurrence in other contexts. Reduced competition from prior contexts could lead to richer encoding of contextual information in the memory trace of nonwords as compared with words. One potential criticism for such interpretation is that nonwords typically do not show superior memory for contextual details such as ink color, font type, or screen location as compared with words (Mulligan, Lozito, & Rosner, 2006). It is important to distinguish the cuing property of context from the ability to retrieve the contextual details, and it is the latter that was examined in the Mulligan et al. study. Thus, one might form stronger item-to-context associations during encoding and benefit from the cuing of those associations during retrieval, but might not necessarily demonstrate enhanced ability to retrieve contextual details when asked to specify the origin of remembered items. Furthermore, the Mulligan et al. study could be described as examining memory for the local context (i.e., changes that occur with the presentation of each item) rather than for the global context (i.e., list episode), and the two types of contexts may be encoded differently. Whether the processing of nonwords establishes stronger item-to-context associations or merely enhances the importance of contextual cues in recognition (because of the absence of better retrieval cues) remains to be further examined. However, both explanations can handle the larger directed forgetting impairment with nonwords than with words.

Finally, directed forgetting impairment with nonwords could emerge because of more direct retrieval processes involved in recognition judgments. Retrieval-based recognition has been shown to be more susceptible to context effects than has familiarity-based recognition (Macken, 2002). If recognition of nonwords more actively engaged retrieval processes, then this would explain larger directed forgetting with nonwords than with words. One criticism of such an interpretation is the claim that nonwords may actually elicit less recollection than words. For example, nonwords typically produce more know responses and fewer remember responses than do words in the remember–know procedure (see, e.g., Gardiner & Java, 1990; Greene, 2004; Macken, 2002; Rajaram, Hamilton, & Bolton, 2002). However, the present results demonstrate different response criteria for the recognition of words and nonwords, and such criteria are known to also affect the rate of remember responses (Benjamin, 2005;

Hirshman & Henzler, 1998). In other words, an alternative interpretation of the lower rate of remember responses with nonwords is that it reflects a response bias rather than a deficit in recollection (e.g., Dunn, 2004).

Although our primary focus in the present article was with the costs of directed forgetting, we will briefly discuss the directed forgetting benefits, which emerged only with long lists of words in Experiment 2. They did not emerge with short word lists, or with nonwords. According to the strategy-change account, directed forgetting benefits reflect a strategic encoding advantage in the forget group. The use of nonassociative stimuli likely precluded the application of better encoding strategies, which can explain the absence of benefits with nonwords. Words, on the other hand, can benefit from organizational encoding, but the benefit of such encoding was only evident with longer lists. These results replicate prior reports of directed forgetting benefits, which were found only with the longer lists (Benjamin, 2006; Sahakyan & Delaney, 2005), and imply that the list-method directed forgetting benefit in recognition is mediated by the list length. In comparison with the short word lists, the long lists in Experiment 2 hurt overall memory in the remember group more than in the forget group, indicating a greater list-length effect in the remember group than in the forget group. This could have occurred because increasing the length of the list impaired recollection (see, e.g., Cary & Reder, 2003; Yonelinas & Jacoby, 1994), but more elaborate encoding of List 2 items in the forget group provided support for recollection and offset the list-length effect, allowing for the benefits to emerge. Shorter lists, on the other hand, may have placed less demand on recollection, and familiarity-based recognition provided a sufficient basis for discrimination. Therefore, with short study lists, performance in the remember group was in a range that masked the benefits in the forget group. The absence of benefits in Experiment 3 could either be linked to the shorter lists being employed in that study, or to the encoding instructions that emphasized to participants to pay attention to the plural/singular status of the words. We employed such instructions because prior research showed that drawing attention to the plurality aspect during encoding more extensively engages retrieval-based recognition than do the standard instructions (e.g., Rotello et al., 1999). However, such encoding demands may have precluded successful application of better encoding strategies because participants' resources were taxed as they were dual tasking. These questions require further investigation and are beyond the scope of the present article.

In conclusion, our studies have provided the first empirical demonstration that directed forgetting costs can be reliably obtained in recognition when stimuli are used or when tests are employed that enhance the utilization of contextual information. We hope that our findings will encourage more active investigation of this phenomenon, which has long eluded the field.

REFERENCES

ARNDT, J., & REDER, L. M. (2002). Word frequency and receiver operating characteristic curves in recognition memory: Evidence for a dual-process interpretation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 830-842.

BADDELEY, A. D. (1982). Domains of recollection. *Psychological Review*, **89**, 708-729.

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, & Cognition*, **19**, 603-616.

BENJAMIN, A. S. (2003). Predicting and postdicting the effects of word frequency on memory. *Memory & Cognition*, **31**, 297-305.

BENJAMIN, A. S. (2005). Recognition memory and introspective remember/know judgments: Evidence for the influence of distractor plausibility on "remembering" and a caution about purportedly non-parametric measures. *Memory & Cognition*, **33**, 261-269.

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

BENJAMIN, A. S., & BAWA, S. (2004). Distractor plausibility and criterion placement in recognition. *Journal of Memory & Language*, **51**, 159-172.

BENJAMIN, A. S., & BJORK, R. A. (2000). On the relationship between recognition speed and accuracy for words rehearsed via rote versus elaborative rehearsal. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 638-648.

BJORK, E. L., & BJORK, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness & Cognition*, **5**, 176-196.

BJORK, E. L., & BJORK, R. A. (2003). Intentional forgetting can increase, not decrease, residual influences of to-be-forgotten information. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 524-531.

BJORK, E. L., BJORK, R. A., & ANDERSON, M. C. (1998). Varieties of goal-directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 103-137). Mahwah, NJ: Erlbaum.

BJORK, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning & Verbal Behavior*, **9**, 255-268.

BJORK, R. A. (1972). Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 217-235). Washington, DC: Winston.

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.

BLOCK, R. A. (1971). Effects of instructions to forget in short-term memory. *Journal of Experimental Psychology*, **89**, 1-9.

CANAS, J. J., & NELSON, D. L. (1986). Recognition and environmental context: The effect of testing by phone. *Bulletin of the Psychonomic Society*, **24**, 407-409.

CARY, M., & REDER, L. M. (2003). A dual-process account of the list-length and strength-based mirror effects in recognition. *Journal of Memory & Language*, **49**, 231-248.

CONWAY, M. A., HARRIES, K., NOYES, J., RACSMA'NY, M., & FRANKISH, C. R. (2000). The disruption and dissolution of directed forgetting: Inhibitory control of memory. *Journal of Memory & Language*, **43**, 409-430.

CRAIK, F. I. M., & WATKINS, M. J. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **12**, 599-607.

DALTON, P. (1993). The role of stimulus familiarity in context-dependent recognition. *Memory & Cognition*, **21**, 223-234.

DUNN, J. C. (2004). Remember–know: A matter of confidence. *Psychological Review*, **111**, 524-542.

ELMES, D. G., ADAMS, C., III, & ROEDIGER, H. L., III (1970). Cued forgetting in short-term memory: Response selection. *Journal of Experimental Psychology*, **86**, 103-107.

FEENAN, K., & SNODGRASS, J. G. (1990). The effect of context on discrimination and bias in recognition memory for pictures and words. *Memory & Cognition*, **18**, 515-527.

FERNANDEZ, A., & GLENBERG, A. M. (1985). Changing environmental context does not reliably affect memory. *Memory & Cognition*, **13**, 333-345.

GARDINER, J. M., & JAVA, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition*, **18**, 23-30.

GEISELMAN, R. E., & BJORK, R. A. (1980). Primary versus secondary rehearsal in imagined voices: Differential effects on recognition. *Cognitive Psychology*, **12**, 188-205.

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.

GLENBERG, A. M. (1979). Component-levels theory of the effects of spacing repetitions on recall and recognition. *Memory & Cognition*, **7**, 95-112.

GODDEN, D. [R.], & BADDELEY, A. [D.] (1980). When does context influence recognition memory? *British Journal of Psychology*, **71**, 99-104.

GORMAN, A. M. (1961). Recognition memory for nouns as a function of abstractness and frequency. *Journal of Experimental Psychology*, **61**, 23-29.

GOTTLOB, L. R., & GOLDING, J. M. (2007). Directed forgetting in the list method affects recognition memory for source. *Quarterly Journal of Experimental Psychology*, **60**, 1524-1539.

GREENE, R. L. (2004). Recognition memory for pseudowords. *Journal of Memory & Language*, **50**, 259-267.

GROSS, A. E., BARRESI, J., & SMITH, E. E. (1970). Voluntary forgetting of a shared memory load. *Psychonomic Science*, **20**, 73-75.

GUTTENTAG, R. E., & CARROLL, D. (1997). Recollection-based recognition: Word frequency effects. *Journal of Memory & Language*, **37**, 502-516.

HINTZMAN, D. L., & CURRAN, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory & Language*, **33**, 1-18.

HIRSHMAN, E., & HENZLER, A. (1998). The role of decision processes in conscious recollection. *Psychological Science*, **9**, 61-65.

HUMPHREYS, M. S., PIKE, R., BAIN, J. D., & TEHAN, G. (1988). Using multilist designs to test for contextual reinstatement effects in recognition. *Bulletin of the Psychonomic Society*, **26**, 200-202.

JACOBY, L. L. (1983). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 21-38.

JACOBY, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, **30**, 513-541.

JOHNSON, H. M. (1994). Processes of successful intentional forgetting. *Psychological Bulletin*, **116**, 274-292.

KRAFKA, C., & PENROD, S. (1985). Reinstatement of context in a field experiment on eyewitness identification. *Journal of Personality & Social Psychology*, **49**, 58-69.

KUČERA, H., & FRANCIS, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

LIGHT, L. L., & CARTER-SOBELL, L. (1970). Effects of changed semantic context on recognition memory. *Journal of Verbal Learning & Verbal Behavior*, **9**, 1-11.

LOFT, S., HUMPHREYS, M. S., & WHITNEY, S. J. (2008). Control of access to memory: The use of task interference as a behavioral probe. *Journal of Memory & Language*, **58**, 465-479.

MACKEN, W. J. (2002). Environmental context and recognition: The role of recollection and familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 153-161.

MACLEOD, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1-57). Mahwah, NJ: Erlbaum.

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.

MALPASS, R. S., & DEVINE, P. G. (1981). Guided memory in eyewitness identification. *Journal of Applied Psychology*, **66**, 343-350.

MANDLER, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, **87**, 252-271.

MARSH, R. L., COOK, G. I., & HICKS, J. L. (2006). The effect of context variability on source memory. *Memory & Cognition*, **34**, 1578-1586.

MARSH, R. L., MEEKS, J. T., HICKS, J. L., COOK, G. I., & CLARK-FOOS, A.

(2006). Concreteness and item-to-list context associations in the free recall of items differing in context variability. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **32**, 1424-1430.

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 14-24.

Mulligan, N. W., Lozito, J. P., & Rosner, Z. A. (2006). Generation and context memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **32**, 836-846.

Murnane, K., & Phelps, M. P. (1993). A global activation approach to the effect of changes in environmental context on recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 882-894.

Murnane, K., & Phelps, M. P. (1995). Effects of changes in relative cue strength on context-dependent recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 158-172.

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

Rajaram, S., Hamilton, M., & Bolton, A. (2002). Distinguishing states of awareness from confidence during retrieval: Evidence from amnesia. *Cognitive, Affective, & Behavioral Neuroscience*, **2**, 227-235.

Rotello, C. M., & Macmillan, N. A. (2008). Response bias in recognition memory. In A. S. Benjamin & B. H. Ross (Eds.), *The psychology of learning and motivation: Skill and strategy in memory use* (Vol. 48, pp. 61-94). London: Academic Press.

Rotello, C. M., Macmillan, N. A., & Van Tassel, G. (1999). Recall-to-reject in recognition: Evidence from ROC curves. *Journal of Memory & Language*, **43**, 67-88.

Russo, R., Ward, G., Geurts, H., & Scheres, A. (1999). When unfamiliarity matters: Changing environmental context between study and test affects recognition memory for unfamiliar stimuli. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 488-499.

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

Smith, S. M. (1985). Environmental context and recognition memory reconsidered. *Bulletin of the Psychonomic Society*, **23**, 173-176.

Smith, S. M. (1986). Environmental context-dependent recognition memory using a short-term memory task for input. *Memory & Cognition*, **14**, 347-354.

Smith, S. M., & Vela, E. (1992). Environmental context-dependent eyewitness recognition. *Applied Cognitive Psychology*, **6**, 125-139.

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

Smith, S. S., Glenberg, A. [M.], & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, **6**, 342-353.

Smith, S. S., Vela, E., & Williamson, J. E. (1988). Shallow input processing does not induce environmental context-dependent recognition. *Bulletin of the Psychonomic Society*, **26**, 537-540.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, **117**, 34-50.

Steyvers, M., & Malmberg, K. J. (2003). The effect of normative context variability on recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 760-766.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, **80**, 352-373.

Whetstone, T., Cross, M. D., & Whetstone, L. M. (1996). Inhibition, contextual segregation, and subject strategies in list method of directed forgetting. *Consciousness & Cognition*, **5**, 395-417.

Yonelinas, A. P., & Jacoby, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology*, **48**, 516-534.

APPENDIX
Stimulus Materials From Experiments 1–3

Table A1
Nonword Stimuli From Experiment 1

AKLACE	AMORTA	LEBALI	ABEISUT
AMORAM	BANARA	MECIENTA	ANUHMI
ASEBIE	BANEPA	MEFFON	ASTEPA
BAYGON	BEDIROF	MELAP	BANDULU
BREFFA	BREGIN	MENIN	BEIDNOK
CHACKA	CANXORST	MONALI	CANNEVY
CHYLUP	DANALI	MUOLIN	DIENTO
DANDEA	DIENMOR	MUORIT	DINWORI
FURALI	DINOAP	NADRAL	DISIROF
HONIN	DORLIDON	NAFFORAM	FEIDOS
HURMY	FEBRA	NIFDARN	FIEORT
INURMA	FEITRAN	NOGWIN	FOERIT
JINTY	FIASTOR	NUMILAN	FURSTOR
KAMTA	FIEWORT	OINNEVY	INORIDS
MEPUT	FOTULAF	OMALAM	INOTSAR
MINDRAL	FURDNOK	REDAHN	INOWIRT
MONGIN	FURRININ	RESHUM	INTULAF
NAFFIN	FURVUL	SADOLE	KIALIN
NAWER	FUSHUN	SLINIROF	LEBARA
NURMER	HAMUNI	SLIWORI	MECULS
REDOLE	IBEADOS	TEMLY	MUDARN
RONDILAN	INODRAL	TIMKA	MUXORST
SASHUN	INOKAIN	TULVY	NIFALI
TEMNU	INOSUT	VULCKA	OIDRAL
VADDIN	INTONAR	WANCULS	SLINOAP
VULLY	INUBIE	WEGGON	WANTRAN
WEGNUM	JINNIN	WIDORT	WIDMOR
ZAPPININ	KAMDRAL	WONTIAN	WONKAIN

Table A2
Word Stimuli From Experiment 2

AIR	CAST	ELBOW	LEADER	PAPER	SNAIL
ANGLE	CATTLE	ENGINE	LEMON	PARTY	STOVE
ANIMAL	CELLAR	ESSAY	LETTER	PATH	STREET
APPLE	CHAIR	FACT	LATE	PAUSE	SIGN
ARMY	CHEESE	FINGER	LIGHTS	PEDAL	SUGAR
ARROW	CHESS	FISH	LOFT	PEPPER	TAXI
ARTIST	CIRCLE	FLOWER	LUNCH	PERSON	TEA
AUTHOR	CITY	FOREST	MANGO	PHONE	TEMPLE
AMBER	CLOCK	FRAME	MARKET	PHOTO	TICKET
BABY	CLOTH	FRIEND	MASK	PIANO	TIGER
BARN	COAT	FROG	METAL	PILLS	TISSUE
BEAVER	COFFEE	GARAGE	MOLE	PINE	TOWER
BENCH	COMET	GLASS	MONEY	POTATO	TRAP
BIRD	CREW	GLUE	MONK	PRISON	TREE
BLOCK	CRIME	GRASS	MOTH	PUPIL	VESSEL
BODY	CROP	GUITAR	MUSIC	RADIO	VEST
BOTTLE	CROWN	HAMMER	NAIL	RIFLE	VICTIM
BULB	DAWN	HONEY	NEST	RIVER	VOICE
BUTTER	DISH	HOTEL	NOVEL	SALARY	VOTER
CABIN	DOOR	INSECT	OCEAN	SALT	WATER
CAKE	DRESS	ISLAND	OLIVES	SCARF	WAVE
CANDLE	DRUM	JACKET	OPERA	SEASON	WEDGE
CANDY	ECHO	JOKE	OWNER	SHOES	WIND
CARPET	EGG	LAKE	PAINT	SHOP	YARD

Table A3
Stimuli From Experiment 3
(Only Singular Version of the Nouns Is Shown)

ARTIST	SALAD	BLANKET
RIVER	RIBBON	SCANDAL
TORNADO	CLOUD	LETTER
BOOK	WINDOW	WHISKER
COSTUME	VEHICLE	ANTIQUE
MUSCLE	BALLOON	FRIEND
QUILT	BROTHER	STATUE
TICKET	FROG	MILE
CABIN	LIQUID	KITE
ROAD	PAPER	BUTTON
NOODLE	WEAPON	CAMERA
PLANET	TOWN	BALLOT
LAMP	SUNSET	GIFT
SIGN	JEWEL	WIRE
BAND	MENU	NATION
FLOWER	BEAN	MINERAL
TOOL	FABRIC	DOLLAR
BELL	GIRL	BASKET
ISLAND	MOVIE	SQUARE
SCHOOL	CARD	SHADOW