Journal of Experimental Psychology:
Learning, Memory, and Cognition
2009, Vol. 35, No. 4, 970–988

0278-7393/09/$12.00 DOI: 10.1037/a0015728

A Global Theory of Remembering and Forgetting From Multiple Lists

Melissa Lehman and Kenneth J. Malmberg
University of South Florida

Forgetting is frustrating, usually because it is unintended. Other times, one may purposely attempt to forget an event. A global theory of recognition and free recall that explains both types of forgetting and remembering from multiple list experiments is presented. The critical assumption of the model is that both intentional and unintentional forgetting are often due to contextual interference. Unintentional forgetting is the natural result of contextual changes between study and test. Intentional forgetting is accomplished by a rapid, metacognitively instigated change in mental context that renders to-be-forgotten information relatively inaccessible and renders to-be-remembered information more accessible (L. Sahakyan & C. M. Kelley, 2002). This occurs for both recognition and free recall. Implications for item-method directed forgetting, exclusion recognition, source memory, and encoding operations are discussed.

Keywords: episodic memory, directed forgetting, memory models, free recall, recognition

Forgetting is frustrating when it is unintended. However, forgetting also can result from attempts to control metacognitively the accessibility of memories. For instance, one might attempt to temporarily render inaccessible a memory for a traffic ticket or a rejection letter received prior to reviewing a journal manuscript (cf. Wenzel, Pinna, & Rubin, 2004). This is *intentional forgetting*, and in these cases, forgetting is welcome. In this article, we ask how these forms of forgetting are related. Toward that end, we present a model of intentional and unintentional forgetting.

Unintentional forgetting has been investigated in countless ways (Anderson & Neely, 1996; Crowder, 1976; for reviews). Free recall and recognition are two of the most common memory tasks, and unintentional forgetting characterizes the performance of both. Intentional forgetting, on the other hand, requires specialized experimental procedures. For the *list method of directed forgetting*, two lists are studied, and participants are instructed to remember both lists (the remember condition) or are instructed, after studying the first list, to forget the first list (the forget condition). Contrary to the instructions, both lists are tested. In prior list-method experiments, depending on how memory has been tested, intentional forgetting has been somewhat less robust than what is typically observed in standard memory experiments. For free recall, memory is worse in the forget condition than in the remember condition for words from the to-be-forgotten list, and is better for words from the to-be-remembered list (Basden & Basden, 1998; Basden et al., 1993; Bjork, 1972, 1978; Bjork & Geiselman, 1978; Bjork, LaBerge, & Legrand, 1968; Bjork & Woodward, 1973; Block, 1971; Geiselman & Bagheri, 1985; Geiselman et al., 1983; MacLeod, 1998; MacLeod, Dodd, Sheard, Wilson, & Bibi, 2003; Malmberg, Lehman, & Sahakyan, 2006; Roediger & Crowder, 1972; Sahakyan, 2004; Sahakyan & Delaney, 2003; Sahakyan & Kelley, 2002; Sheard & MacLeod, 2005; Weiner, 1968; Weiner & Reed, 1969; Woodward & Bjork, 1971). These effects are the *costs* and *benefits*, respectively, of directed forgetting. It is important to note that concurrent costs and benefits have not been observed for recognition memory (some report no effects, see Basden, Basden, & Gargano, 1993; Block, 1971; Elmes, Adams, & Roediger, 1970; Geiselman, Bjork, & Fishman, 1983; others report partial effects, see Benjamin, 2006; Loft, Humphreys, & Whitney, 2008; Sahakyan & Delaney, 2005). Thus, intentional forgetting is ubiquitous for free recall, but not for recognition.

This conclusion suggests that forgetting is at times under the control of the participant, although never completely so. While there are many—not necessarily mutually exclusive—hypotheses about specific intentional forgetting findings, a comprehensive explanation is currently lacking. One reason is that the relationship between recognition and recall has not been investigated rigorously, and a comprehensive explanation will account for both tasks (Gillund & Shiffrin, 1984; Humphreys, Bain, & Pike, 1989; Malmberg & Shiffrin, 2005; Shiffrin & Steyvers, 1997). On that note, we identified several methodological issues that make interpreting the results from prior experiments difficult, we conducted several new experiments that addressed these problems, and we analyzed the data at a level of detail that illuminates several results that strongly constrain theory. Additional constraints were imposed by consid-

Melissa Lehman and Kenneth J. Malmberg, Department of Psychology, University of South Florida.

This research served as partial fulfillment of Melissa Lehman's master thesis at the University of South Florida. Parts of this research were presented at the 2006 Cognitive Science Society Meeting in Vancouver, British Columbia, Canada; the 2006 Annual Meeting of the Society for Mathematical Psychology in Vancouver, British Columbia, Canada; the 2006 Annual Summer Interdisciplinary Conference in Andulnase, Norway; the 4th Annual Context and Episodic Memory Symposium in Tampa, Florida; the XXIX International Congress of Psychology in Berlin, Germany; and the 49th Annual Meeting of the Psychonomic Society, Chicago, Illinois. Portions of Experiment 1 were previously reported in the *Proceedings of the Annual Meeting of Cognitive Science Society* (2006).

We thank Mark Goldman, Doug Nelson, and Jon Rottenberg for their contributions. We also thank Marc Howard and Simon Dennis for their thoughtful comments on a prior version of the article.

Correspondence concerning this article should be addressed to Melissa Lehman, Department of Psychology, University of South Florida, 4202 East Fowler Avenue, PCD 4118G, Tampa, FL 33620. E-mail: mlehman@mail.usf.edu

ering how accounts of intentional forgetting relate to accounts of unintentional forgetting. Thus, we sought to understand both types of forgetting phenomena by relating them to a single model.

Classical Forgetting Hypotheses

Differential Rehearsal

Rehearsal plays a well-documented role in many memory models as a mechanism that maintains an item in an accessible state, thereby also increasing the amount of information that is encoded about that item (Atkinson & Shiffrin, 1968; Rundus, 1971). Rehearsal has also been proposed to play an important role in intentional forgetting. According to the *differential-rehearsal hypothesis* (Bjork et al., 1968; Sheard & MacLeod, 2005) instructions to forget alter the allocation of limited resources during study and, hence, the extent to which some items are encoded. Accordingly, participants stop rehearsing words from the to-be-forgotten List 1 (i.e., L_1) after the forget instruction is given and devote all further rehearsals to the following List 2 (i.e., L_2).[1] In contrast, participants in the remember condition covertly rehearse items from L_1 while they study L_2. This reduces the average number of rehearsals allocated to L_2 items and increases the average number of rehearsals allocated to L_1 items. Because items on L_1 receive more rehearsals after an instruction to remember, compared with L_1 items in the forget condition, items on L_1 are encoded better, and they are more likely to be remembered. This explains the costs of directed forgetting. Because items from L_2 compete with items from L_1 for limited rehearsals in the remember condition, these L_2 items are not remembered as well as L_2 items in the forget condition, and this produces the benefits of directed forgetting. Indeed, the instruction to forget affects the form of free-recall serial position curves (MacLeod, 1998; MacLeod et al., 2003; Sheard & MacLeod, 2005). For L_2, there is a pronounced primacy effect in the forget condition and an almost absent primacy effect in the remember condition. Thus, most of the L_2 benefits are associated with enhanced memory for items in the early serial positions. For L_1, the instruction to forget has a smaller effect on the form of the serial position curves, although performance is greater in the remember condition than in the forget condition, of course.

According to the differential-rehearsal hypothesis, the advantage for the items studied early on L_2 in the forget condition is due to a reduced competition for resources used for rehearsal. This differential-rehearsal hypothesis, however, is unlikely to provide a complete explanation of list-method directed forgetting for several reasons. The instruction to remember should enhance memory for L_1 items presented at the end of the list, on the assumption that they are the L_1 items given extra rehearsals during L_2 in the remember condition. However, this has not been observed at times (Sheard & MacLeod, 2005), and thus, directed forgetting can be observed even when the recency portion of L_1 is unaffected by the instruction to forget.[2]

In addition, there is little support for the assumption that there is a significant amount of cross-list rehearsals in the remember condition. When two lists were studied, according to Ward and Tan (2004), participants sometimes rehearsed items from L_1 during the presentation of L_2. However, the number of times that an item from L_1 was rehearsed was very small, ranging from 0.33 to 0.09 for items at different L_1 serial positions. The mean number of rehearsals for an L_1 item was 0.16, which means that on average there were 3.2 cross-list rehearsals. Even so, this might overestimate cross-list rehearsals in a directed-forgetting experiment. First, participants in these experiments were instructed to rehearse aloud, and this is not done in list-method directed forgetting experiments. Second, participants were always required to recall from L_1 and never from the most recent list (cf. Jang & Huber, 2008; Shiffrin, 1970), and the cogent participant would be expected to use a cross-list rehearsal strategy. Indeed, Ward and Tan's participants were well practiced at recalling from L_1, as they participated in 20 study–test cycles during the course of the experiment. Thus, given the task demands and the amount of practice participants received, it is surprising that so little cross-list rehearsal occurred, if it plays a major role in producing the cost and the benefit of list-method directed forgetting.

Other findings are problematic for the differential-rehearsal hypothesis. The differential-rehearsal hypothesis predicts that directed forgetting should not be observed when rehearsal is discouraged, but directed forgetting is observed (Bjork et al., 1968; Block, 1971 Geiselman et al., 1983; Sahakyan & Delaney, 2005). Moreover, every theory of memory predicts that altering the extent of item encoding, via enhanced rehearsal or other means, should improve both free recall and recognition (cf. Malmberg, 2008a). The fact that directed forgetting has been observed rarely for recognition memory is problematic for these models. Acknowledging this, Sheard and MacLeod noted that serial position effects might be smaller for recognition, and thus, it might be difficult to observe directed forgetting because prior experimental designs were not suitable for observing reliable effects.[3] We discuss, in detail, other reasons why it has been difficult to observe list-method directed forgetting for recognition in subsequent sections of this article.

Inhibition

The role of inhibition in episodic memory is under active investigation, most notably as it relates to unintentional forgetting in the domain of retrieval-induced forgetting (Anderson & Bjork, 1994; Norman, Newman, & Detre, 2007). However, the

[1] The models that we discuss focus on the list method because it is for this procedure that the interactions between recall and recognition have been observed. For the item method of directed forgetting, items are presented with a subsequent cue to remember or forget each item. Recognition and free recall for to-be-remembered words is better than for to-be-forgotten words (Roediger & Crowder, 1972; MacLeod, 1975; Woodward & Bjork, 1971; Woodward, Park, & Seebohm, 1974). Thus, the differential-rehearsal hypothesis assumes that upon the presentation of the remember instruction, participants engage in an elaborative rehearsal process that is not invoked after the instruction to forget (MacLeod, 1975; Woodward et al., 1973).

[2] One finding suggests a small effect of the forget instruction on the recency portion of L_1, but the reliability of this result is not known because formal analyses were not reported (Geiselman et al., 1983).

[3] Sheard and MacLeod (2005) also note that different remember and forget serial position curves are a necessary consequence of the instruction to forget, and they should be observed for recognition memory if the instruction to forget is successful, but this has not been observed.

possibility that inhibition is used to intentionally forget has also been investigated; inhibition of to-be-forgotten items produces the costs, and a concomitant reduction in interference produces the benefits (Elmes, Adams, and Roediger, 1970; Weiner, 1968; Weiner & Reed, 1969). For instance, participants might mentally group the to-be-forgotten and to-be-remembered material separately and then inhibit the to-be-forgotten set during retrieval (Geiselman, Bjork, & Fishman, 1983). Because they are inhibited, these items create less proactive interference, leading to the benefits.

However, the inhibition hypothesis does not well explain the null effects of intentional forgetting on recognition. To account for them, sometimes the inhibition hypothesis assumes that recognition testing releases the to-be-forgotten items from inhibition (Basden et al., 1993;Geiselman & Bagheri, 1985; Geiselman & Panting, 1985; MacLeod et al., 2003). This suggestion is circular, and usually, there is no evidence that the to-be-forgotten items were ever inhibited to start (Basden, Basden, & Wright, 2003; Bjork & Bjork, 1996; Geiselman et al., 1983). Inhibition accounts are further challenged to explain why some recognition experiments exhibit no costs, and yet the benefits remain (Benjamin, 2006; Sahakyan & Delaney, 2005). That is, under what conditions should a release from inhibition be observed, under what conditions should a release from inhibition not be observed, and why? Last, the inhibition hypothesis should explain how participants place the traces into two separate sets, inhibit one set, and activate the other. In this sense, inhibition accounts describe the data well, but they do not offer much insight into the operations of memory.

Contextual Differentiation

Changes in context play a primary role in forgetting, according to many theories (Dennis & Humphreys, 2001; Estes, 1955; Gillund & Shiffrin, 1984; Humphreys et al., 1989; Howard & Kahana, 2002; Jang & Huber, 2008; Mensink & Raaijmakers, 1989; Murdock, 1997; Murnane, Phelps, & Malmberg, 1999). As the difference between the context features encoded during study and context cues available at test increases, forgetting increases. According to Sahakyan and Kelley's (2002) variant of the set differentiation hypothesis of directed forgetting (Bjork et al., 1968; Bjork, 1970), study involves storage of information representing the studied items (i.e., item information) and the context in which the items occur (i.e., context information). L_1 and L_2 are associated with an overlapping set of contextual elements (e.g., Estes, 1955; Mensink & Raaijmakers, 1989). The instruction to forget causes an accelerated change in context between lists, and there is less interference between L_1 and L_2. When recalling from L_2, less interference from L_1 traces produces the benefits of the instruction to forget. The costs are the result of the relative inaccessibility of an effective L_1 context cue, due to the relatively rapid change in context that occurred between the list presentations. This is the *contextual-differentiation hypothesis.*

The logic behind the contextual-differentiation hypothesis is derived from the literature on context-dependent memory (e.g., Anderson, 1983; Eich, Weingartner, Stillmin, & Gillin, 1975; Godden & Baddeley, 1975; Goodwin, Powell, Bremmer, Hoine, & Stern, 1969; Macht, Spear, & Levis, 1977; Murnane et al., 1999; Smith, 1979; Smith, Glenberg, & Bjork, 1978). Sahakyan and Kelley (2002) compared standard directed forgetting conditions to a between-list context-change condition. In the context-change condition, some participants were given the remember instruction, followed by an instruction to imagine that they were invisible, in order to create a mental context change. Participants in the remember-plus-context-change condition performed almost identically to participants in the standard forget condition—showing both costs and benefits of context change. A strong prediction of the contextual-differentiation hypothesis is that the costs and benefits of directed forgetting are dependent on the ability of the participant to mentally reinstate appropriate context cues at test. Indeed, context effects are eliminated or reduced when appropriate context cues are available for both intentional and unintentional forgetting procedures. For instance, Smith (1979) showed that the mental reinstatement of the environmental context eliminates the costs of context dependent memory. In the intentional forgetting literature, Sahakyan and Kelley (2002) used standard remember and forget conditions, but after studying the second list, half the participants participated in a context-reinstatement procedure. Afterward, participants in forget and remember-plus-context-change groups showed reduced costs and benefits, compared with the groups that did not receive the reinstatement. Presumably, the remaining costs and benefits are due to the use of some contextual elements found at test. In any case, these findings revealed that context reinstatement has similar effects on intentional and unintentional forgetting.

We developed a model of the retrieval mechanisms supporting the contextual-differentiation model for free recall (Malmberg et al., 2006). To this point, however, it has not been applied to serial position analyses and other fine-grained aspects of the intentional forgetting. Moreover, the nature of context change and the model's ability to handle recognition simultaneously has not been explored. One problem for the contextual differentiation hypothesis concerns the lack of a recency effect observed in many free-recall directed forgetting experiments (Malmberg et al., 2006). All things being equal, any model of free recall that assumes that temporal context plays an important role during retrieval predicts that L_2 should be remembered better than L_1 in the remember condition because L_1 was learned prior to L_2 (Ebbinghaus, 1885). In contrast, sometimes L_1 is actually remembered better than L_2 in the remember condition (Geiselman, Bjork, & Fishman, 1983; Sahakyan & Kelley, 2002; Sahakyan, 2004). This reversed recency effect, better memory at longer retention intervals, suggested to us that the traditional designs used in list-method directed forgetting confound several variables with list order, such as the location of distractor tasks and the presence of proactive interference, and thus give L_1 an advantage over L_2. Another challenge for the contextual-differentiation hypothesis is the often observed null effect for recognition memory. Most models assume that context plays an important role in episodic recognition. The assumption is supported by findings that show context-dependent recognition performance (Dennis & Humphreys, 2001; Light & Carter-Sobell, 1970; Murnane et al., 1999). Thus, there should be an effect of the instruction to forget on recognition memory if the contextual-differentiation hypothesis is correct and if recognition depends on the use of mentally reinstated context.

Toward a Global Model of Forgetting

In the preceding discussion, we identified three issues that are critical for developing a coherent account of remembering and forgetting from multiple lists: the relation between retention interval and episodic memory, the effect of the instruction to forget as a function of serial position at study, and the relation between recognition and free-recall performance. In this section, we briefly discuss how we approached resolving these issues.

How does an increase in the retention interval affect episodic memory? The failure to find better memory for L_2 than L_1 in the remember condition is difficult to explain by most models, including the contextual differentiation model (e.g., Howard & Kahana, 2002; Jang & Huber, 2008; Mensink & Raaijmakers, 1989). However, the traditional list-method design confounds variables that benefit L_1 and harm L_2 (Malmberg et al. 2006). L_2 receives proactive interference from L_1, but L_1 does not receive proactive interference from another list. In addition, a distractor task is often used after L_2, but not after L_1. The distractor task discourages covert rehearsals for the last few items on a list and helps to ensure that retrieval is from long-term memory (e.g., Bjork & Whitten, 1974; Glanzer & Cunitz, 1966). Without a distractor task after L_1, additional rehearsals can easily be allocated to the last one or two L_1 items at the expense of the rehearsals allocated to the first L_2 items. Both confounds may benefit memory for L_1 relative to L_2, and thus, we attempt to control for these factors in present experiments.

We also noted that the differential-rehearsal hypothesis includes several predictions concerning the effect of the forget instruction on the serial position curves. If the costs and benefits are due to the covert rehearsal of the last few items on L_1 while participants study L_2, then there should be a pronounced recency effect for L_1 and a smaller primacy effect for L_2 in the remember condition, relative to the forget condition. This should occur for both free recall and recognition, if directed forgetting is possible for recognition. Failure to observe this pattern of data would disconfirm the present versions of the differential-rehearsal hypothesis and would point toward the context-differentiation account.

The nature of the recognition tests used to assess directed forgetting is another critical issue. The list method requires multiple study lists. Under these conditions, recognition experiments can use either an inclusion test or an exclusion test (Jacoby, 1991; Winograd, 1968). In an inclusion test, one should endorse any item studied during the experiment. Hence, context cues that differentiate the study lists are not required. In contrast, exclusion recognition requires that the participant endorse only words from a specified list. In this case, the participant must use a context cue that differentiates the study lists to accurately perform the task. Note that list-method free recall also requires a context cue for a particular list. Thus, the exclusion task is more similar to what is required for free recall than the inclusion task, and if the contextual-differentiation hypothesis is accurate, then we should see robust effects of directed forgetting on exclusion task performance. The effects should be similar to those observed for free recall, in which intrusion rates are reduced by the forget instruction (Malmberg et al., 2006). Thus, there should be costs and benefits on the hit rates and the recency advantage for L_2. There should also be more L_2 false alarms when a participant is attempting to recognize from L_1 than there would be L_1 false alarms when a participant is attempting to recognize from L_2, and false-alarm rates should be lower in the forget condition. It is interesting to note that most of the recognition experiments in the directed forgetting literature used an inclusion procedure rather than an exclusion procedure.

General Method

Having identified these critical issues, we used a three-list design with 30 s arithmetic tasks after each list, to better equate proactive interference and potential for covert rehearsals across the critical lists. Accordingly, the instruction to forget occurred for half the participants after completing the arithmetic task following the second list and immediately before studying the third list (memory was not tested for the first list, thus the first list is hereby referred to as L_0, the second list is referred to as L_1, and the third list is referred to as L_2, allowing for ease in comparing the critical lists with those in previous research). Because the free-recall and recognition methods that we used differed only in the nature of the memory test, we first present those methods and only discuss those aspects of the experiments that differed in subsequent sections.

Materials and Procedure

The entire experiment was completed on a computer in an individual participant room. For each participant, 48 medium frequency (frequency of occurrence between 20 per million and 50 per million) nouns were randomly chosen from the Francis and Kucera (1982) norms and were randomly divided into three lists of 16 words.

At the beginning of the experiment, participants were told that the experimenters wanted to see how well people could remember information and how well people could remember where that information came from. Participants were informed that they would see three lists of words and that they would be tested on only one of the lists but that they would not be told which list until later in the experiment, so they needed to remember all of the lists. The instructions were as follows:

> At the beginning of this experiment, you will study three lists of words. The words will appear on the screen one at a time for a few seconds each. Your task is to remember these words for a later memory test. Importantly, I will only ask you to remember the words from one of the lists, which will be chosen randomly, but **you will not be told which list until later in the experiment.** In between each list, there will be a short math task. This is involves adding digits in your head and entering the total into the computer. Once you have done so, the next list of words will be presented.

Once the participants indicated that they understood the instructions, they began the study phase. Participants were given a warning before each list that the study list was about to begin. Each list was presented one word at a time for 8 s. After each list, participants completed the distractor task, consisting of a series of two-digit addition problems. Participants were instructed to complete as many problems as they could in 30 s. Participants in the remember condition were shown each list and distractor task, followed by the test. In the forget condition, participants were shown the first two lists and distractor tasks, then given the forget instruction, followed by study of the third list and a third distractor task. The forget instruction was as follows:

> Next, you are going to receive the third study list. This is the list that you will be asked to recall, so you do not need to worry about the first two lists.

Participants in the forget condition were then shown the third list and distractor task, followed by the test.

Experiment 1: Free Recall

To relate intentional and unintentional forgetting using some version of the contextual-differentiation model, it is critically important to determine whether recency characterizes free-recall performance. If it does then intentional forgetting can be modeled within the frameworks of several global memory theories. If it does not then intentional forgetting would require a special model. The recall probabilities from this experiment were reported in Malmberg et al. (2006). The results of the prior analyses showed a reliable effect of recency: L_2 items were recalled more often than were L_1 items, in both remember and forget conditions. We report these analyses again here because they support new serial position analyses that are critical for differentiating between the contextual-differentiation hypothesis and the differential-rehearsal hypothesis. In particular, we sought to determine the loci of the effect of the instruction to forget because the differential-rehearsal hypothesis predicts that it should occur over the recency portion of L_1 and over the primacy portion of L_2. The serial position analyses are also critical for relating free-recall and recognition performance, as a coherent model ought to account for both.

Method

Participants. One hundred eighty undergraduate psychology students at the University of South Florida participated in exchange for course credit. Data for 12 participants were not used because they did not recall any words from any lists, leaving 168 participants (42 per condition).

Procedure. After all three study lists (and distractor tasks), participants were given a free-recall test lasting 90 s. They were told to type all of the words that they could remember from the specified list and enter them into the computer one at a time. Half of the participants in each condition were tested on L_1, and half were tested on L_2. Participants from the forget condition who were tested on L_1 (the to-be-forgotten list) were told that we wanted them to recall from this list, even though we had previously told them that they would not need to remember it. After being tested on the specified list, participants were tested on the other list (either L_1 or L_2); however, these data were only used to determine whether any participants failed to recall any words from either list—in which case the data were thrown away.

Results and Discussion

Correct recall. The statistical analyses are confined to the data obtained from L_1 and L_2. The results of a two-way analysis of variance show that for correct recall, there was a main effect of list, $F(1, 164) = 68.84$, $MSE = .021$, $p < .001$; for both the remember condition and the forget condition, probability of recall was significantly greater for L_2. This confirms the prediction that more recent lists would be better remembered than would less recent lists. There was no main effect of instruction, but there was a significant List $\times$ Instruction interaction, $F(1, 164) = 19.66$, $p < .001$. As shown in the left panel of Figure 1, recall of L_1 was better for the remember condition than for the forget condition, but the opposite was true for L_2 (the costs and benefits of directed forgetting, respectively). According to planned comparisons, all results shown here are significant to a .05 criterion.

Intrusions. As expected, intrusion rates were very low. These are shown in the left panel of Figure 1. For intrusions from L_1 when recalling L_2 and vice versa, there were no significant results; however, there were some interesting trends. Participants were

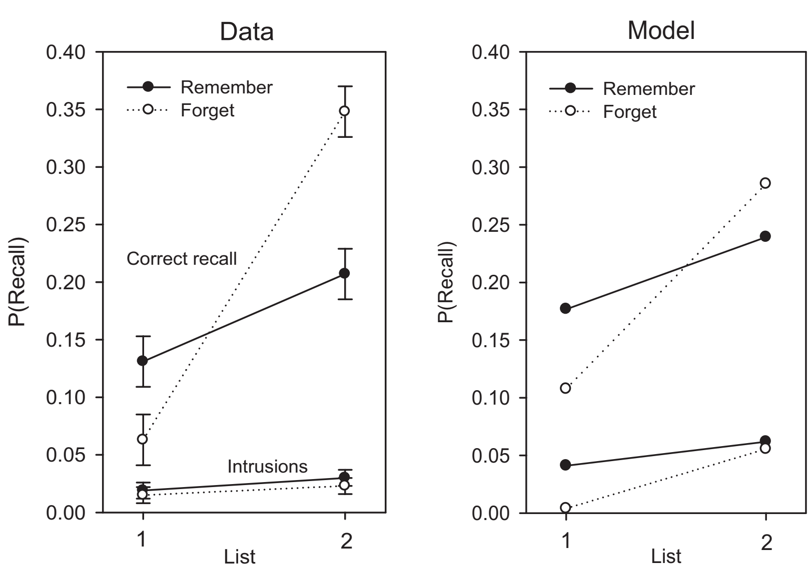

Figure 1. Probability of correct recall and intrusion errors for free recall in Experiment 1. The intrusions in this graph refer to intrusions that came from either List 1 or List 2. When recalling from List 1, any List 2 item that was output is referred to as an intrusion and vice versa. Error bars represent standard error. P(Recall) indicates probability of recall.

more likely to have intrusions from L_2 while being tested on L_1 than they were to have intrusions from L_1 while being tested on L_2. Further, the probability of either type of intrusion was lower for participants in the forget condition than for participants in the remember condition. For intrusions that came from L_0, there was a significant main effect of list, $F(1, 164) = 17.83$, $MSE = .002$, $p < .001$; participants were more likely to have intrusions from L_0 while they were recalling L_1 than while they were recalling L_2, for both remember and forget conditions, and again, intrusion rates were lower for participants in the forget condition. These are shown in Table 1.

Thus, the costs and benefits of directed forgetting were obtained, and L_2 items were remembered better, on average, than were L_1 items, which is expected if the context cue at test consists of features more similar to those during the study of L_2 than to those during the study of L_1. In addition, the intrusions from L_2 while being tested on L_1 were greater than were intrusions from L_1 while being tested on L_2, intrusion rates were lower in the forget condition than in the remember condition, and intrusions from L_0 were more likely when participants were recalling from L_1 than when participants were recalling from L_2.

Serial position. The serial position analyses allowed us to explore in detail the contribution of rehearsal to directed forgetting. The left panels of Figure 2 show the serial position data. The main effect of serial position, $F(15, 164) = 5.59$, $MSE = .125$, $p <$ .001 was significant; however, the List × Instruction × Serial Position interaction was not significant. To evaluate whether the forget instruction affected the primacy portions of L_1 and L_2 in opposite ways, we conducted two separate two-way analyses of variance (one for each list), with bin as a within-subjects factor (Bin 1 vs. Bin 4) and with instruction as a between subjects factor. For L_2, primacy was greater in the forget condition than in the remember condition, $F(1, 164) = 7.559$, $MSE = .368$, $p = .007$, which is consistent with the differential-rehearsal hypothesis (Sheard & Macleod, 2005). However, primacy was greater in the remember condition than in the forget condition for L_1, $F(1, 164) = 2.43$, $MSE = .435$, $p = .044$.

Although it is not immediately clear why the instruction to forget affected the primacy portion of L_1 according to the differential-rehearsal hypothesis, one possibility is that the distractor task was effective in eliminating L_1 items from a rehearsal buffer. If so, participants in the remember condition might have covertly recalled an item from L_1 immediately following the arithmetic task, as suggested by Ward and Tan (2004). Note that there is no effect of the instruction to forget on the recency portion of the L_1 serial position curve; this suggests that the last items on L_1 were not rehearsed during the beginning of L_2 (or during the arithmetic task). Thus, L_1 items rehearsed during L_2 might have come from earlier serial positions (cf. Ward & Tan, 2004). Additional analyses on the order in which items were output were conducted in order to further examine this hypothesis.

First-recall probabilities (FRPs). The left panels of Figure 3 show that the instruction to forget had a large impact on the FRPs, and L_1 and L_2 were impacted in opposite directions. There was a significant main effect of serial position, $F(15, 161) = 10.04$, $MSE = .057$, $p < .001$, moderated by a significant List × Instruction × Serial Position interaction, $F(15, 161) = 3.29$, $p < .001$. Thus, the first item recalled from L_1 was most likely to be the item in the first serial position only in the remember condition, whereas on L_2, the first item recalled was more likely to be the item in the first serial position for the forget condition than for the remember condition.

There was a strong tendency to output initially the first item on the study list. For L_1 in the remember condition, the probability of initially recalling the item from the first serial position was ~40%; no other serial position produced a probability of more than 10%. This effect was severely diminished in the forget condition, and the opposite pattern was observed for L_2. Thus, the effect of the forget instruction is apparent in the result of the first successful retrieval attempt. This suggests, according to the differential-rehearsal hypothesis, that participants in the remember condition might have covertly rehearsed the first item from L_1 when studying the first items on L_2 (cf. Ward & Tan, 2004). What remains unclear, however, is why rehearsing the first item from L_1 produces costs that extend to the first two thirds of L_1 (this finding replicates Sheard and MacLeod, 2005). To argue that more than one or two L_1 items were rehearsed during L_2, moreover, implies a very high intrusion rate when recalling from L_2, but Figure 1 shows that the probability of an L_1 intrusion was only about .025. Thus, the pattern of serial position data is difficult for the differential-rehearsal hypothesis to explain.

A more viable approach to the serial position and FRP data might be the contextual-differentiation hypothesis. Assume that the context cues at test are more strongly associated with the items studied at the beginning of each list. This would produce the observed FRP functions. The probability of initially recalling the item from the first serial position of L_2 would be greater in the forget condition due to less retrieval competition from L_1 items. In contrast, the opposite would be

Table 1
Intrusion Rates and False-Alarm Rates for Experiments 1, 2, 3a, and 3b

	Remember			Forget		
Experiment	L_1	L_2	Remember	L_1	L_2	Forget
Exp. 1: List 0 intrusions, free recall	.051	.016		.038	.010	
Exp. 2: False-alarm rates for unstudied foils, exclusion	.112	.083		.143	.073	
Exp. 3a: False-alarm rates for unstudied foils, inclusion			.079			.077
Exp. 3b: False-alarm rates for unstudied foils, inclusion			.100			.068

Note. In inclusion, L_1 and L_2 were tested together, thus there are only false-alarm rates for remember and forget conditions. For Exp. 2 and Exp. 3a, there was an 8 s study time; for Exp. 3b, there was a 4 s study time. Exp. = experiment; L_1 = List 1; L_2 = List 2.

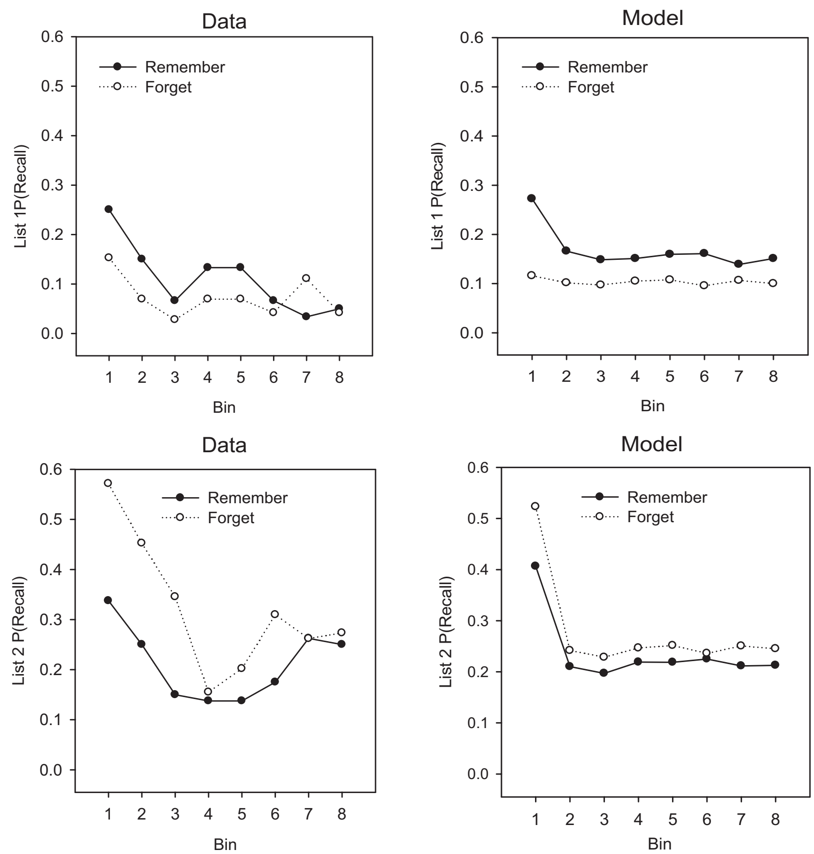

Figure 2. Serial position for free recall in Experiment 1. For the sake of clarity, the 16 item list was compiled into eight bins spanning two serial positions. For instance, bin *n* contains the data from serial positions $2n - 1$ and $2n$. P(Recall) indicates probability of recall.

true for L_1 because there is a smaller change in mental context between lists in the remember condition, and thus, the appropriate L_1 context cues are more readily available.

To account for the full serial position curve with the contextual-differentiation hypothesis, we may further assume that after the first item is recalled it is used as part of a compound cue along with contextual elements to probe memory. If so, and if an episodic association between items is stored during study, the next item to be recalled will tend to be from the next or a nearby serial position. The probability of an unsuccessful retrieval attempt will obviously increase with each additional item output. This accounts for the low intrusion rates, and it explains the decrease in the probability of recall with increases in serial position over the primacy portion of the serial position curve.

Experiments 2, 3A, and 3B: Exclusion and Inclusion Recognition

The free-recall results of Experiment 1 are consistent with the contextual-differentiation hypothesis. However, this hypothesis also predicts costs and benefits for both inclusion and exclusion recognition. To assess these predictions, we conducted inclusion and exclusion recognition experiments in which the study portions were exactly the same as Experiment 1. That is, we used a three-list design with a 30 s distractor task after each list. The only difference was the instruction given to the participant directly before memory testing. In Experiment 2, participants received exclusion instructions, and in Experiments 3A and 3B, participants received inclusion instructions with 8 s and 4 s, respectively, of study time per item. The shorter study time was used in Experiment 3B to guard against possible ceiling effects that might be observed with the 8 s study time used for free-recall and exclusion recognition.

Method

Undergraduate psychology students at the University of South Florida participated in exchange for course credit. In Experiment 2 (exclusion), there were 148 participants (37 in each condition). The materials and study procedure were identical to those of Experiment 1. For the test, participants were told which list of words they would need to recognize (either L_1 or L_2). They were told to respond "yes" if the word shown was on the specified list and to respond "no" if the word shown was from a different list or if it was a new word. The test list consisted of L_1 and L_2 targets and 16 new words presented in a random order. In Experiment 3A (inclusion with 8 s study time), there were 60 participants (30 in each condition). The design, materials, and procedure were the same as Experiment 2, except participants were

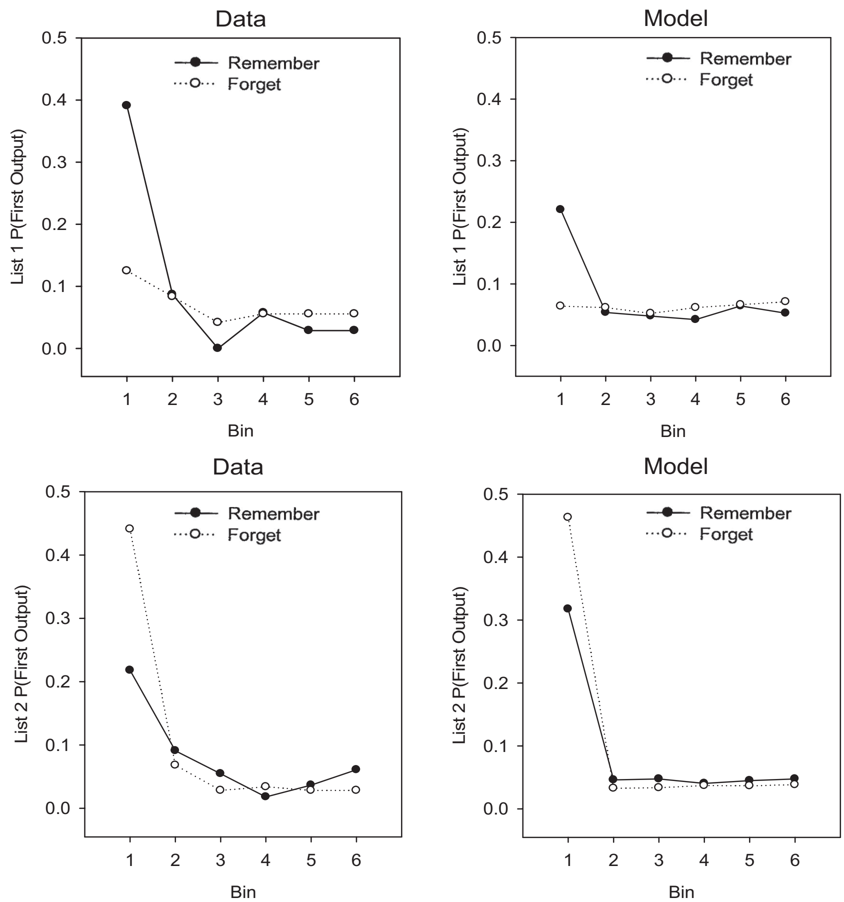

Figure 3. First recall probabilities for free recall in Experiment 1. For the sake of clarity, the 16 item list was compiled into bins. Bin 1 represents the first item on the list (since this is where differences are seen), and all other serial positions are grouped by three. P(First Output) indicates first recall probabilities.

told that if they had seen that word on any list in the experiment then they should respond by clicking *yes* and that if the word was a new word they were to respond by clicking *no*. In Experiment 3B (inclusion with 4 s study time), there were 86 participants (43 in each condition), and the design, materials, and procedure were exactly the same as Experiment 3A except that a 4 s study duration was used.

Exclusion Results

Hits. The data were analyzed as a two-way analysis of variance with list and instruction as between-subject factors. The hit rates were greater for L_2 than for L_1, $F(1, 144) = 22.06$, $MSE = .026$, $p < .001$. There was no significant main effect of instruction, but as shown in the left panels of Figure 4, there was a significant List × instruction crossover interaction, $F = 13.54$, $p < .001$, with hit rates in the remember condition greater than in the forget condition for L_1, and the opposite for L_2. Planned comparisons confirm the simple effects. Thus, in terms of hit rates, we found the costs and the benefits of directed forgetting and a recency effect. As displayed in the left panels of Figure 5, there was not a significant effect of serial position, $F(15, 144) = 1.05$, $MSE = .225$, $p = .399$; however, there was a significant List × Instruction × Serial Position interaction ($F = 1.49$, $p = .021$).

False alarms. The false-alarm rates for those test items that were not studied are shown in Table 1. There are no significant main or interaction effects. This indicates that average recognition performance (whether an item was studied) is captured solely by the differences that were observed in hit rates. The crossover interaction that was observed in hit rates, therefore, indicates that there are costs and benefits associated with the instruction to forget for recognition memory just as is the case for free recall.

The exclusion instructions were to respond negatively to items that were studied but that were not studied on the target list. The rate at which participants failed to do so can also be considered a false-alarm rate. As shown in the left panel of Figure 4, false-alarm rates are uniformly lower for participants in the forget condition than for participants in the remember condition, $F(1, 144) = 5.14$, $MSE = .049$, $p < .03$. Additionally, while the effect of list was not reliable, participants responded "yes" to L_2 items when being tested on L_1 more often than they responded "yes" to L_1 items while being tested on L_2. This trend is consistent with the assumption that the context used to probe memory is more similar to L_2 than to L_1. Thus, more L_2 items are mistakenly associated with the L_1 context when memory is probed and when the target list is L_1 and vice versa when L_2 is the target list.

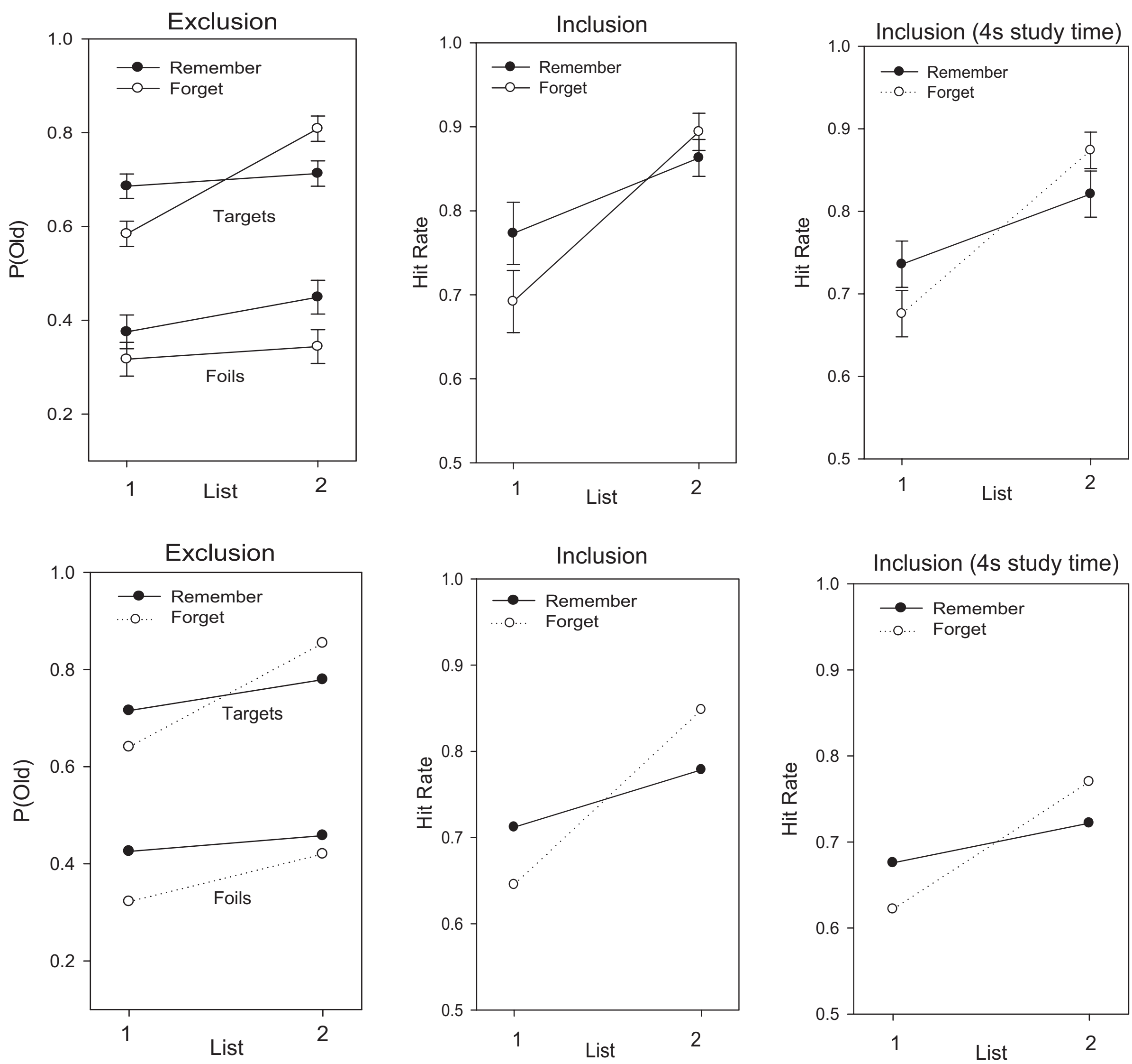

Figure 4. Recognition performance for Experiments 2, 3A, and 3B. Data from the recognition experiments are shown on the top row, and model predictions are shown on the bottom. In the exclusion experiment, participants were told to say "yes" only to items that were studied on a specific list. Thus, some items should have been rejected even though they were studied because they were studied on the to-be-excluded list and not studied on the to-be-endorsed list. The graphs on the left show hit rates for List 1 and List 2 in the remember and the forget conditions (targets), along with false-alarm rates for items that were studied on the to-be-excluded list (foils). For example, if the participant was instructed to positively endorse only items on List 1, any List 2 items that were positively endorsed counted as List 2 foils. In inclusion, participants were told to say "yes" to any studied word; thus, middle and right graphs show hit rates only. False-alarm rates for unstudied foils in exclusion are as follows: For data, R-L_1, 0.1120; R-L_2, 0.0830; F-L_1, 0.1430; and F-L_2, 0.0730; and for model, R-L_1, 0.0706; R-L_2, 0.0731; F-L_1, 0.0881; and F-L_2, 0.0750. Error bars represent standard error. R = remember; F = forget; L = list; P(Old) = probability that the item is judged as old.

Inclusion Results

For Experiment 3A, there was a significant main effect of list, $F(1, 58) = 35.43$, $MSE = .018$, $p < .001$; hit rates were higher for L_2 than for L_1. There was no main effect of instruction, but there was a significant List × Instruction interaction, $F(1, 58) = 5.271$, $p = .025$. As shown in the middle panels of Figure 4, hit rates were higher for the remember condition on L_1, but there was a minimal difference in recognition on L_2. Planned comparisons revealed that performance was significantly higher for L_1 in the remember condition but that the difference between conditions was not significant for L_2 ($p = .47$). There were no differences in the probability of responding *yes* to new words between remember and forget conditions, as shown in Table 1. As displayed in the left

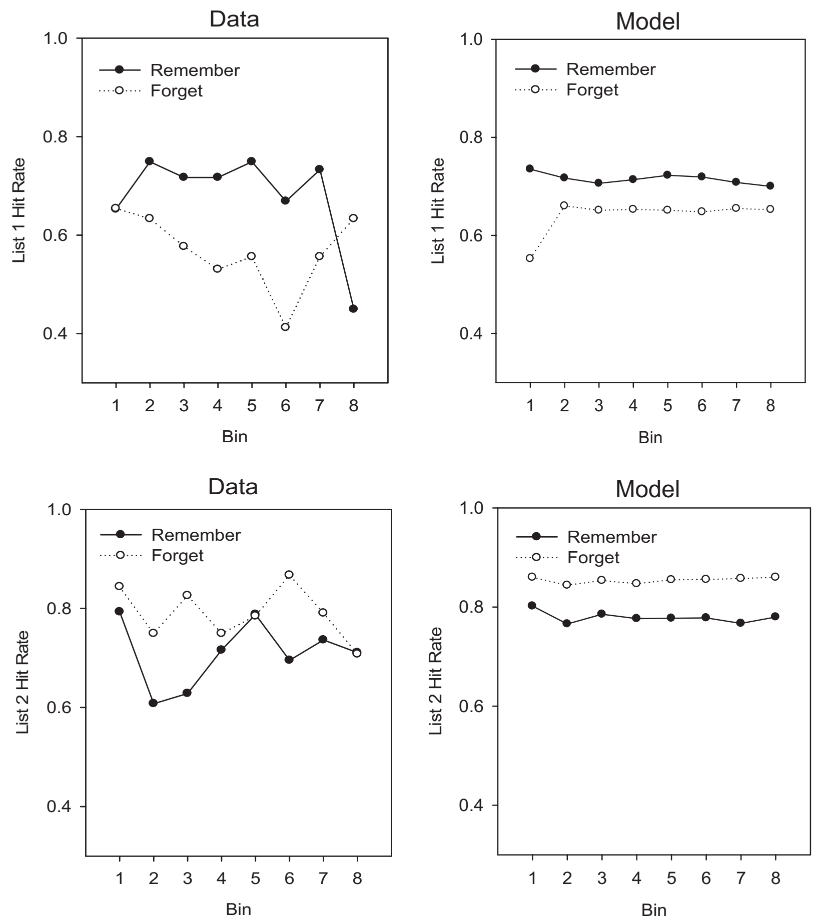

Figure 5. Exclusion serial position data for Experiment 2. For the sake of clarity, the 16 item list was compiled into eight bins spanning two serial positions. For instance, bin n contains the data from serial positions $2n - 1$ and $2n$.

panel of Figure 6, there was not a significant effect of serial position in the inclusion experiment, $F(15, 58) = 1.43$, $MSE = .147$, $p = .123$.

For Experiment 3B, there was a significant main effect of list, $F(1, 82) = 45.49$, $MSE = .019$, $p < .001$; hit rates were higher for L_2 than for L_1. There was no main effect of instruction, but there was a significant List $\times$ Instruction interaction, $F(1, 82) = 7.21$, $MSE = .019$, $p = .009$. As shown in the right panels of Figure 4, hit rates were better for participants in the remember condition on L_1 and the forget condition on L_2. Planned comparisons revealed that all differences shown here were significant. As predicted, we saw recency of L_2 and both costs and benefits of directed forgetting. As in Experiment 3A, there were no significant differences in the false-alarm rates in the remember and forget conditions (see Table 1).

Discussion

For exclusion, the costs and benefits of directed forgetting are observed in hit rates. The hit rate was greater for L_2 in the forget condition, and the hit rate was greater for L_1 in the remember condition. In addition, the false-alarm rates were greater in the remember condition than in the forget condition for both L_1 and L_2. This is consistent with the assumption that a mental change of context occurred after the instruction to forget that led to a decrease in the tendency to assign L_1 items to L_2 and L_2 items to L_1.

For the inclusion experiment, we observed a similar pattern of costs and benefits in the hit rates as we observed in the exclusion experiment. However, the benefits did not reach the standard of significance. As Sheard and MacLeod (2005) noted, however, such small effects could be difficult to detect when approaching a ceiling. For this reason, we replicated the inclusion experiment and reduced study time from 8 s per item to 4 s per item. Because of the reduced study time, the hit rates were lower than with an 8 s study time, and a significant crossover interaction in the hit rates was observed. Thus, costs and benefits of directed forgetting are observed for both recognition memory tasks. The effects are more robust for the exclusion task than for the inclusion task, but they are obtained for both. The left panels of Figures 5 and 6 show that differences in the serial position curves do not account for the differences in the remember and forget conditions. Hence, it appears that directed forgetting effects can be observed for recogni-

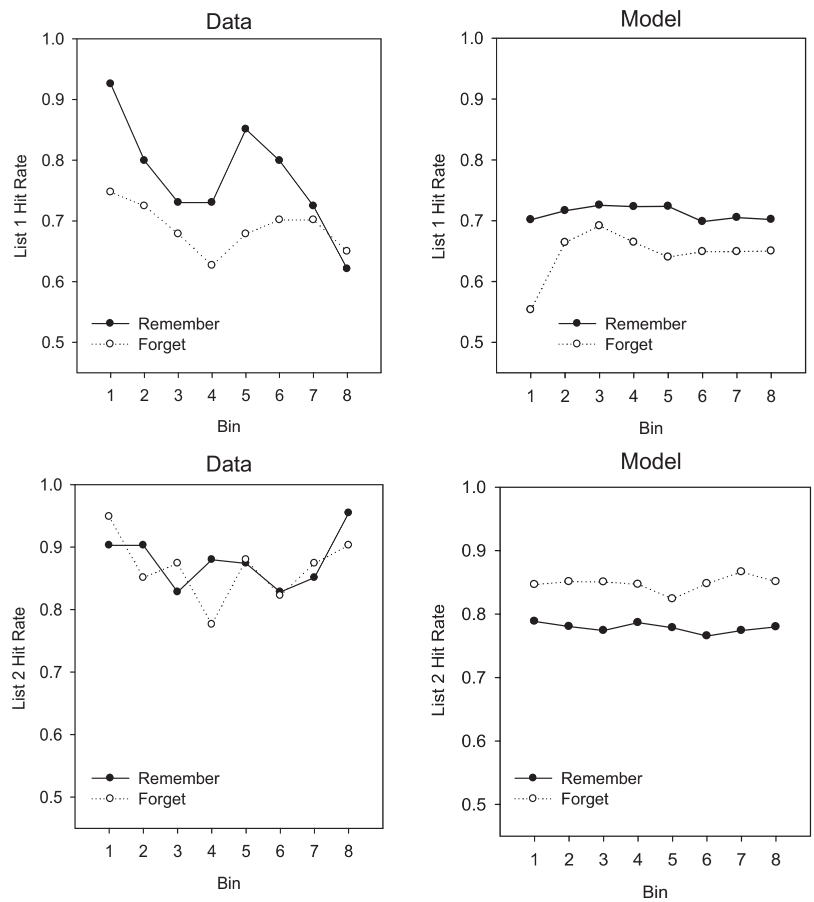

Figure 6. Inclusion serial position data for Experiment 3A. For the sake of clarity, the 16 item list was compiled into 8 bins spanning two serial positions. For instance, bin n contains the data from serial positions $2n - 1$ and $2n$.

tion memory in the absence of serial position effects (cf. Sheard & MacLeod, 2005).

Forgetting and Remembering From Multiple Lists

We developed a multilist memory model by extending the retrieving effectively from memory (REM) free-recall and context encoding models (Malmberg & Shiffrin, 2005) and the REM recognition models (Criss & Shiffrin, 2004; Malmberg, 2008b; Malmberg, Holden, & Shiffrin, 2004; Xu & Malmberg, 2007; Shiffrin and Steyvers, 1998). Within this framework, we implemented the contextual differentiation hypothesis as a major factor underlying intentional forgetting. We show that it accounts for multilist unintentional forgetting, the costs and benefits of intentional forgetting, serial position curves, and FRP functions.

Representation

According to REM, general knowledge of items is stored in lexical/semantic memory traces and information about past events is stored in episodic memory traces. Lexical/semantic traces are acquired over a lifetime. They contain information about how words are spelled and pronounced and what they mean. In addition, they contain information about the contexts or situations in which they have been encountered. As such, they are accurate, complete, and generalizable to the contexts in which they usually occur.

Two concatenated vectors represent these traces. One vector represents the item and the other represents the contexts in which it has been encountered. The w features comprising the vectors are generated according to a geometric distribution with the base rate parameter, g:

$$P[V = j] = (1 - g)^{j-1}g, j = 1, \ldots, \infty. \quad (1)$$

When a word is studied, the w item features of its lexical semantic trace are copied to form a new episodic trace that represents this occurrence. In addition, w features of the current context are stored.

Episodic encoding is an incomplete and error-prone process; a feature may be copied correctly, it may be copied incorrectly, or it may fail to be copied. The probability of storing a feature given a certain unit of time (t) is represented by the u_x^* parameter, where x indexes different aspects of memory in a manner that will become apparent. Given that a feature is stored, it is stored correctly with a probability c. An item will be stored incorrectly with a probability $1 - c$, in which case a feature will be randomly chosen according to the geometric distribution. The absence of a stored feature is represented by the value zero. Thus, the features representing the items will be randomly similar between each item.

Here, we begin to implement contextual differentiation in the model. When items are studied, context is stored in episodic traces in the same way. For the sake of simplicity, we assume that context features change between lists with a probability of β but do not change within lists (Criss & Shiffrin, 2005; Malmberg & Shiffrin, 2005). Thus, for each list, a single context vector is generated to represent the current context, and all items within that list are associated with the same context information, which is stored according to the rules for item storage outlined above. When a context feature value is changed, it is randomly sampled from the geometric distribution. We further assume that context features change after the final study list, in the same manner as they change between lists.

Buffer Operations

As a descendent of the Atkinson and Shiffrin theory (1968), the interaction of control processes and structural aspects of memory are used to model serial position data. Control processes operate on items located in a limited capacity rehearsal buffer during encoding. For present purposes, we chose a buffer capacity of two items, although larger capacities also work and would probably even help achieve more accurate predictions. However, buffer capacities larger than two require a relatively complex set of assumptions. For now, therefore, we keep the model as simple as possible while still allowing it to capture the major trends in the data.

To this point, we have made no new assumptions within the REM framework. Now we add assumptions to prior REM models of encoding to describe buffering operations that have never been needed in prior models. Upon the presentation of the first item on a list, its lexical/semantic item features enter the buffer, and two things happen. First, there are t attempts to copy the item's features to an episodic trace. The probability of storing an item feature is u_i^*. There are also t attempts to copy the current context features in an episodic trace. The probability of storing a context feature is u_{c1}^* for the first item on the list. For the first item, we assume $u_i^* = u_{c1}^*$.

Upon the presentation of the second item on the list, its lexical/semantic item features take the remaining slot in the buffer, and three things happen. New item and current context features are stored in the same way as before. In addition, some of the item features of the first item are stored in an additional concatenated vector (cf. Kimball, Smith, & Kahana, 2007). This represents the assumption that an episodic association is stored between the two items in the buffer (this loosely corresponds to strengthening an interitem association in search of associative memory (SAM) cf. Criss & Shiffrin, 2005; Xu & Malmberg, 2007). The probability of storing the associated item's features is u_a^*. We assume that participants tend to focus their attention on the most novel item in the rehearsal buffer. Thus, less information will be encoded about the older item than about the current item. This is represented by a greater u^* value for current-item information ($u_i^* > u_a^*$).

When two items occupy the buffer, its capacity is reached. Thus, the context that is stored is shared between the two item representations, and we assume that the probability of storing a context feature is less likely as the buffer capacity is taxed, $u_c^* < u_{c1}^*$. The assumption that context feature are more likely to be encoded for the first item than for subsequent items allows the model to account for the FRPs. When the next item is studied, the new item is added to the buffer, and this process repeats, with the oldest item being dropped with a probability δ.

Retrieval

The first step of the retrieval process is similar across all test conditions (recall, recognition-inclusion, and recognition-exclusion). A relevant subset of memory is created that consists of the items with the strongest association to the context used as the initial retrieval cue (cf. Shiffrin & Steyvers, 1997; REM.5). For creation of the relevant subset, the current context cue is matched against the context stored in the episodic images.[4] The matching process involves calculating a likelihood ratio for each trace, which takes into account both features that match and features that do not match. Matching features increase and mismatching features decrease the likelihood ratio; cases in which no features are stored do not contribute to the likelihood ratio either way. Likelihood ratios are calculated according to the following equation:

$$\lambda_j = (1 - c)^{n_{jq}} \prod_{i=1}^{\infty} \left[\frac{c + (1 - c)g(1 - g)^{i-1}}{g(1 - g)^{i-1}} \right]^{n_{ijm}}, \tag{2}$$

where g is the environmental base rate for the occurrence of features, c is the probability of correctly storing a feature, i is a feature value ranging from 1 to infinity, n_{jq} is the number of mismatching context features for episodic image j (regardless of their value), and n_{ijm} is the number of matches with value i in image j. The relevant subset of memory will consist of a certain percentage, ρ, of all traces in memory with the greatest likelihood ratios.

Free Recall

The free-recall task begins with the creation of the cue with which to probe memory. The initial cue consists of only context features; it is a combination of the current test context and reinstated list context. The proportion of reinstated list context features is represented by the γ parameter. The remaining context features in the cue are from the test context.

Free recall operates in REM cycles of sampling and recovery (Malmberg & Shiffrin, 2005). The initial context cue is matched against all traces in the activated subset in an attempt to sample an item from the given list. Likelihood ratios for all images are calculated according to Equation 2. The probability of sampling image, I_i, given the context retrieval cue, Q, is

$$P(I_i|Q) = \frac{\lambda_i}{\sum \lambda_k}. \tag{3}$$

[4] The assumption that only the images stored during study may be included in the relevant subset is probably overly simple. However, increasing the amount of noise coming from other traces will not affect the model's predictions because they are not associated with the reinstated context features, and hence, these images are unlikely to provide good matches to the retrieval cue.

Recovery does not play an important role in this model, so we simply assume that sampled traces are recovered successfully (cf. Malmberg & Shiffrin, 2005). Thus, when an item is sampled and recovered, and it comes from an incorrect list, the participant undertakes a monitoring process to determine whether it is an intrusion. We assume that items from the correct list are rarely withheld, and hence, if an item is sampled and if it is from the correct list, it is output with a probability of 1.0. The probability, η, of making an intrusion error, given that an item from the incorrect list is sampled and recovered, is a positive function of the overlap in context between lists (represented by this parameter). Thus, η is greater in the remember condition than in the forget condition.

If an item is output, the next cue used to probe memory will consist of both context and recovered item information. Again, the context portion of the cue consists of both current context features and context features associated with the given list. The item portion of the cue consists of the item vector from the last item recalled. Thus, it is most likely that co-rehearsed items, which share the current item's information, will be sampled next. If no item is output, then the original context cue is used for the next probe of memory. The sample-and-recovery process repeats κ times.

Recognition–Inclusion

For the inclusion task, the task is to positively endorse all studied items. For this reason, a simple global-matching process is used (Malmberg, 2008b; Malmberg, Zeelenberg, & Shiffrin, 2004; Shiffrin & Steyvers, 1997). In REM, a decision about whether an item is judged old is made on the basis of the likelihood ratios calculated for all items in the comparison set. The odds are calculated according to the following equation:

$$\Phi = \frac{1}{n}\sum_{j=1}^{n} \lambda_j, \tag{4}$$

and if the odds exceed 1.0, the item is judged old, otherwise it is judged new.

To create the relevant subset of items, we used the same context cue that was used to create this subset in free recall to probe memory. After this set of traces was created, a retrieval cue consisting only of the item information that represents the test word was used to probe memory, and the odds were calculated. In this case, n is the number of traces in the relevant subset of memory (Shiffrin & Steyvers, 1997).

Recognition–Exclusion

For the exclusion task, a participant positively endorses only items that came from a given study list. A global matching process is first used, as in the inclusion task, followed by a monitoring task, as in the free-recall task. After an item is identified as old, an output decision is made in the same manner as was used for free recall. That is, it is dependent on the overlap in context between the two lists, and this is captured by the η parameter at test. This is essentially a recall-to-reject process (Dosher, 1984; Humphreys, 1976; Malmberg, 2008a). For the sake of simplicity, however, we did not implement the sampling and recovery processes for the exclusion task because all of the water is carried by the overlap between the contexts: A large overlap in context means that it is harder to distinguish between the two lists, and the false-alarm rate will be increased. A description of these processes is found elsewhere (Malmberg, 2008a).

Effects of the Forget Instruction

The forget instruction has multiple effects on encoding at study and formation of context cues at test. Since it occurs prior to the presentation of the instruction of the memory tests, we assume that the forget instruction has the same effects on encoding for all tasks. First, it increases the rate of context change between lists (i.e., contextual differentiation), so that the two lists share fewer context features. In the model, the probability that a context feature changes, β, is greater in the forget condition than in the remember condition. This decreases the amount of interference from the list that is not the focus of recall or recognition when memory is probed with an appropriate context cue. Second, encoding of the context associated with first item on L_2 is enhanced in the forget condition. The enhanced encoding is reflected in a greater u^*_{c1} value. This produces the benefits of the forget instruction.[5] At test, the forget instruction decreases the probability of reinstating features for use in the L_1 context cue. It is more difficult to reinstate the L_1 context features because the forget instruction increases the context change that occurs between study and test. No change in contextual reinstatement is made for the L_2 context cue because the forget instruction accelerates the change in context before L_2.

Model Evaluation

The major modeling challenge is to simultaneously account for unintentional and intentional forgetting in a comprehensive and detailed manner. This was difficult, despite costs and benefits for free recall and recognition, because differences remain. For instance, intrusion rates are low for free recall, but false-alarm rates are relatively high for exclusion recognition, and the tasks produce different serial position curves. Another challenge was to model the FRP functions in a manner that made list discrimination possible and produced costs and benefits. The FRPs are critically important because most of the directed forgetting effect in free recall appears to be driven by them. Tulving (1983) described episodic memory as mental time travel, and these effects require one to "land on a dime" in a manner that is affected by intentional forgetting.

Our approach was to account for both tasks with a single, contextually driven mechanism. Hence, we refer to this as a global model because we are explaining these findings and the relationship between intentional and unintentional forgetting with just a

[5] This assumption might seem to be consistent with Sahakyan and Delany's (2003) two-factor hypothesis, which proposed that forget instructions increase the encoding of items on L_2 only after the instruction to forget (in addition to an increase in the change in mental context between lists). However, this assumption is disconfirmed by the fact that benefits are derived primarily from an advantage in initially recalling the first item from L_2 (see Figure 3). By definition, this must be due to probing memory with a context cue, and hence, there must be an enhancement of the binding of context to the first item on L_2 because of the instruction to forget.

few assumptions. The only differences between the models of free recall and recognition are the assumptions concerning retrieval, and they accounted for a wide variety of episodic memory phenomena (Criss & Shiffrin, 2004; Malmberg, in press, 2008; Malmberg, Holden, & Shiffrin, 2004; Malmberg & Murnane, 2002; Malmberg & Shiffrin, 2005; Malmberg & Xu, 2007; Malmberg, Holden, & Shiffrin, 2004; Shiffrin & Steyvers, 1997, 1998).

We did not attempt to find a best-fitting set of parameters; our focus was accounting for the patterns in the data. Parameter descriptions and their values are listed in Table 2. Twelve of the 16 parameters were fixed in all experimental conditions. Without exception, these scaling parameters are the same or almost the same as those used to fit other REM models to data (Criss & Shiffrin, 2004; Malmberg, in press, 2008; Malmberg, Holden, & Shiffrin, 2004; Malmberg & Murnane, 2002; Malmberg & Shiffrin, 2005; Malmberg & Xu, 2007; Malmberg, Holden, & Shiffrin, 2004; Shiffrin & Steyvers, 1997, 1998).

We have over 250 data points, and given these conditions, four parameters were allowed to vary between remember and forget conditions in accordance with the assumptions of the model: those parameters were u^*_{c1}, β, ρ_1, and η. With these parameters, we conducted a set of 1,000 Monte Carlo simulations for free recall, inclusion recognition, and exclusion recognition. The same set of parameter values were used to generate predictions for all of our experiments.

The right panel of Figure 1 shows the free-recall performance of the model. The model produces the correct patterns of costs, benefits, and intrusions. The predicted intrusion rates are slightly high, especially for L_2. This is because the intrusion rate in free recall is yoked to the false-alarm rate for exclusion recognition via the constant value for η. The lower left panel of Figure 4 shows the model's exclusion performance, and here, the false alarms are a little low, compared with the data. This suggests that the screening process is more rigorous for free recall than we are currently assuming and is less rigorous for exclusion recognition.

The free-recall model's serial position predictions are shown in the right panels of Figure 2. The model predicts the observed interaction between serial position and forget instruction. It also predicts that costs and benefits are greatest for the earliest serial positions. One departure from the data is that the model is predicting the L_1 cost or the L_2 benefit for all serial positions. This is because of our simplifying assumption that context does not change within lists. If reinstated context were to be more similar to the retrieval cue for the earlier serial positions than for the later serial positions, then we would produce a better fit of the model to the data. For now, the additional complexity associated with implementing this assumption does not seem warranted. The right panels of Figure 3 show that the model also provides accurate FRPs and their interaction with the instruction to forget.

The bottom panels of Figure 4 shows that the model accurately accounts for recognition hit rates and false-alarm rates. Finally, the right panels of Figure 5 and 6 show that the model describes the serial position curves for exclusion and inclusion recognition, respectively. The recognition serial position curves are relatively flat compared with those obtained for free recall. One deviation of the model from the data is for the first bin on L_1 in the forget condition, which shows that the hit rate for this bin is lower than the hit rate for other bins. This is due entirely to the predicted hit rate from Serial Position 1 being lower than the rest, and it is attributable to the assumption that reinstated beginning-of-the-list context is not used to probe memory for recognition. The context stored for the first item is stored more strongly than the context stored for the rest of the items on the list. This produces more mismatches with the current test context, and hence, the first item is less likely to be a member of the relevant subset of memory traces. If we used a context cue that consisted of a combination of reinstated context and current test context, this problem would be significantly attenuated.

Table 2
Parameter Values and Descriptions

Parameter	Value	Description
g	.4	Environmental base rate (standard value)
w	8	Number of item and context features
c	.8	Probability of correctly storing a feature
u^*_i	.5	Probability of storing an item feature
u^*_c	.2	Probability of storing a context feature
u^*_a	.1	Probability of copying a co-rehearsed item's feature
u^*_{c1}	.5[a]	Probability of storing a context feature for first item on a list
t	2	Number of storage attempts
κ	20	Number of sampling attempts
β	.2[a]	Probability of change for context features between lists
δ	.75	Probability of dropping the oldest item in the buffer
ρ_1	.2[a]	Probability of reinstating context features on L_1
ρ_2	.8	Probability of reinstating context features for L_2
σ	.8	Size of activated subset of items
η	.4[a]	Probability of outputting an intrusion

Note. L_1 = List 1 and L_2 = List 2.
[a] Parameter values that differ in the forget condition. For the forget condition, $u^*_{c1} = .75$; $\beta = .8$; $\rho_1 = .15$; $\eta = .2$.

General Discussion

Our findings challenge the prevailing views that the list method for directed forgetting does not affect recognition memory and that the list method does not produce a recency effect. After considering the designs of prior experiments, we eliminated several confounding factors in the present experiments. As a result, there were reliable effects of directed forgetting and recency effects for both recognition and recall (more recently learned lists were remembered better than lists learned earlier). We also presented a global model of remembering and forgetting from multiple lists. The model provides a unified account of recognition and free-recall performance under unintentional forgetting and intentional forgetting conditions. The model embraces traditional context-based accounts of unintentional forgetting (e.g., Estes, 1955; Gillund & Shiffrin, 1984; Mensink & Raaijmakers, 1989) and combines them with the newer contextual-differentiation account of intentional forgetting (Sahakyan & Kelley, 2002). The model predicts forgetting as the result of a natural change in context driven by factors that are not necessarily under the control of the participant and as the result of the change in context driven by an attempt to control metacognitively the accessibility of information in memory. Here, we discuss extensions of the present model to account for item-method directed forgetting and other phenomena and the relation between the present model and other models.

Critical Modeling Assumptions

The data suggest that a model based on the contextual-differentiation hypothesis would be best able to handle all of the data. The critical aspects of the current model are the following:

1. The enhanced change in mental context between lists in the forget conditions creates less overlap in contextual features between L_1 and L_2. This contributes to both costs and benefits of directed forgetting for both free recall and recognition; the costs occur because the context at test shares fewer context features with L_1 in the forget condition, and the benefits occur due to less competition from L_1 items when retrieving from L_2 in the forget condition.

2. The change in context that occurs with the forget instruction makes reinstating L_1 context features at the time of recall more difficult, which contributes to the costs.

3. Enhanced contextual differentiation reduces intrusion rates in the forget condition and false-alarm rates in exclusion recognition. Because L_1 and L_2 contexts share fewer features, the lists are more distinct, and list discrimination is enhanced.

4. The forget instruction affects some buffering operations. After the forget instruction, the context associated with the first item on L_2 is better encoded than in the remember condition. This contributes to the benefits because participants are better able to access the first item on L_2 and begin a series of item-plus-context cue retrieval cycles that aid in retrieval of other items from the list.

We know that more complex versions of the current model would provide better quantitative fits of the data, but the current set is the only one that we have identified that captures all of the major trends in the free-recall and recognition data. The current assumptions as a group are necessary to achieve accurate qualitative predictions, and none is sufficient by itself (see the Appendix regarding earlier modeling attempts).

Extensions of the Model

Item method. For the item method of directed forgetting, participants study a single list of items. After the presentation of each item, participants are told that they should remember the word or that they should forget the word. Item-method directed forgetting occurs for free recall and recognition, and there is strong support for this version of the differential-rehearsal hypothesis (Bjork & Geiselman, 1978; MacLeod, 1975; MacLeod, 1998; MacLeod et al., 2003; Sheard & MacLeod, 2005; Woodward, Bjork, & Jongeward, 1973). According to the differential-rehearsal hypothesis of the item method for directed forgetting, the current item is maintained in a rehearsal buffer until the instruction to remember or to forget is given. At that time, to-be-forgotten items are displaced and to-be-remembered items are processed in an elaborative manner to enhance future memory. To extend the current model, we assume that the number of attempts at storing item features, t, is greater for to-be-remembered items than for to-be-forgotten items (cf. Malmberg & Shiffrin, 2005, for a discussion of encoding during different orienting tasks). Enhanced item encoding will obviously enhance recognition, and it will enhance free recall because traces will be more likely to be recovered (Malmberg & Shiffrin, 2005).[6]

Context similarity and exclusion. The current model of exclusion recognition is a simple extension of the REM dual-process models (Malmberg, 2008b; in press; Malmberg, Holden, & Shiffrin, 2004; Malmberg & Xu, 2007; Xu & Malmberg, 2007). Accordingly, this model is appropriate when the recognition decision requires the discrimination of items based on the sampling and recovery of episodic details because the items themselves are both relatively familiar. In this version of the model, we assume that contextual elements corresponding to L_1 and L_2 are the episodic details that allow for accurate exclusion performance. Because the instruction to forget produces a greater change of context between L_1 and L_2 than does the remember instruction, exclusion accuracy is enhanced, but only for L_2. In fact, several experiments varied the similarity of the L_1 and L_2 contexts (Gruppuso, Lindsey, & Kelley, 1997; Mulligan and Hirshman; 1997); either the same orienting task or a different orienting task was used during study on L_1 and L_2. The consistent finding is that L_1 false-alarm rates increase as contextual similarity increases. We found the same increase in false-alarm rates when participants were given the remember instruction, compared with when they were given the forget instruction, which is consistent with the hypothesis that the forget instruction instigated an accelerated change in mental context.

Exclusion performance has been primarily investigated within the process dissociation procedure framework (PDP, Jacoby, 1991; compare with Loft et al., 2008). The designs of these experiments are different from the exclusion experiment that we conducted. We tested exclusion memory for L_1 and L_2, whereas the convention is to test memory only for the most recent list, and this allowed us to discover something important about intentional forgetting. There were greater false-alarm rates in the remember condition for both L_1 and L_2, and a greater L_2 hit rate and lower L_2 false-alarm rate in forget condition. There is clearly greater accuracy for L_2 in the forget condition. However, the change in hit rates and false-alarm rates is approximately equal and in the same direction for L_1, producing a d' of approximately .78 and .70 in the remember and forget conditions, respectively. Thus, the change in exclusion accuracy is limited to L_2, but it would be incorrect to conclude that the instruction to forget did not affect L_1.

Source memory. A typical source memory procedure involves presenting items from different people, books, television shows, and the like, and the task at test is to determine which source presented a given test item. The source memory task is also thought to be highly dependent on the ability to assign items to

[6] It is unclear, however, whether the interpolated remember and forget instructions will alter the buffer operations that we described earlier. It is possible that participant will not create episodic associations between to-be-forgotten and to-be-remembered words.

contexts. In these experiments, the sources can be considered a form of context. Consider that the sources can vary in similarity. For instance, Bayen, Murnane, and Erdfelder (1996) varied the similarity of the faces associated with auditory presentation of words, and they found that false-alarm rates (i.e., assigning words to the wrong source) increased when the faces were similar. An extension of the exclusion model that assumes that the retrieval of source information is the basis for reductions in source errors makes the same prediction.

Conditional recall probabilities. The present framework assumes that items are rehearsed in a limited capacity buffer. Whereas the prior models emphasized the effects of study on transfer of item information to a permanent store (Atkinson and Shiffrin, 1968; Gillund & Shiffrin, 1984; Raaijmakers, & Shiffrin, 1981), the present model emphasizes how study affects the binding of context to item information (as discussed by Malmberg and Shiffrin, 2005). Here, we further assume that the binding of context to item information is more effective under conditions in which the capacity of the buffer is not fully taxed. We also assumed that the capacity of the buffer was two items. In our view, however, the critical issue does not concern the capacity limitation per se, as clearly there are individual differences (Engle, Kane, & Tuholski, 1999). Rather, the present model departs more significantly from prior models with the assumption that item information is strengthened via the encoding of additional, context-dependent, associative traces. This assumption and prior assumptions are not mutually exclusive. For instance, item encoding might be strengthened in a single trace and additional context-to-item associations might be created during study.

Because the buffering operations support the creation of inter-item associations, when items are recovered as the result of a context probe, we assume that a recovered item is used in a joint probe with context on subsequent retrieval attempts. In addition, to the primacy portion of the serial position curve, this produces nonrandom conditional recall probabilities (CRPs), such that items that co-occur in the buffer tend to be recalled in adjacent output positions. Empirically, these tend to be asymmetric, with forward CRPs occurring more often than backward CRPs (Kahana, 1996). Our assumptions, that the capacity of the buffer is two items and that context does not drift within lists, severely diminish backward CRPs. However, relaxing these simple assumptions will increase this tendency. Due to the limited scope of the present project, we chose not to explore this in quantitative manner here. However, for modeling our findings, we found the present assumptions necessary and sufficient (see the Appendix regarding why other versions of the model fail).

Comparisons With the Bind, Cue, and Decide Model

The bind, cue, and decide model (BCDMEM, Dennis & Humphreys, 2001) also emphasizes the role context plays in episodic memory, by comparing mentally reinstated context with the context associated with items during study. For exclusion recognition, BCDMEM assumes that memory is probed with an item, and a composite representation of the contexts in which the item has occurred is matched to mentally reinstated context. The comparison is only with the mentally reinstated context of the most recent list, and the result is a continuous random variable that is compared with a criterion. If L_2 items should be called old and if the random variable associated with an item exceeds the criterion, the item is called old, otherwise it is called new. If items from L_1 should be called old and if the random variable associated with an item does not exceed the criterion, the test item is called old, otherwise it is called new. While BCDMEM accounts for a wide range of recognition phenomena and the present model might not correctly predict list-length effects, the present approach also has some advantages. The REM exclusion model assumes that a recollective process is used to reject items from the wrong study lists on the basis of the retrieval of context features from memory. This is consistent with a broader framework that assumes speed is sacrificed to enhance the accuracy of recognition tasks when foils are similar to targets (Malmberg, 2008a). The dynamic version of the REM model predicts the nonmonotonicities in the retrieval dynamics of exclusion performance that are not easily explained by BCDMEM (cf. McElree, Dolan, & Jacoby, 1999). In addition, BCDMEM does not well explain findings that suggest that item noise affects recognition performance (Criss & Shiffrin, 2004). Last, the REM framework accounts for recognition, free recall, and cued recall. BCDMEM would probably require different cueing assumptions to extend its scope to a global one.

Conclusions

According to the present model, changes in context play a major role in both intentional and unintentional forgetting. For unintentional forgetting, context changes between study and test produce context-dependent memory effects, namely recency and source confusion effects. For intentional forgetting, like that observed with the list method, context changes between lists produce costs and benefits of directed forgetting and enhanced list discrimination. In addition, changes in context make reinstating the study context relatively difficult, and this decreases the effectiveness of the retrieval cues used at test.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22,* 261–295.

Anderson, M. C., & Bjork, R. A. (1994). Mechanisms of inhibition in long-term memory: A new taxonomy. In D. Dagenbach & T. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 265–325). New York: Academic Press.

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 237–313). San Diego: Academic Press.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (pp. 89–195). New York: Academic Press.

Basden, B. H., & Basden, D. R. (1998). Directed forgetting: A contrast of methods and interpretations. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 59–102). Mahwah, NJ: Erlbaum.

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.

Basden, B. H., Basden, D. R., & Wright, M. (2003). Part-list re-exposure and release of retrieval inhibition in directed forgetting. *Consciousness and Cognition, 12,* 354–375.

Bayen, U. J., Murnane, K., & Erdfelder, E. (1996). Source discrimination, item detection, and multinomial models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 197–215.

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 and Cognition, 5,* 176–196.

Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and 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. (1978). The updating of human memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 12, pp. 235–259). New York: Academic Press.

Bjork, R. A., & Geiselman, R. E. (1978). Constituent processes in the differentiation of items in memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 347–361.

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.

Bjork, R. A., & Woodward, A. E. (1973). Directed forgetting in individual words in free recall. *Journal of Experimental Psychology, 99,* 22–27.

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

Criss, A. H., & Shiffrin, R. M. (2004). Context noise and item noise jointly determine recognition memory: A comment on Dennis & Humphreys (2001). *Psychological Review, 111,* 800–807.

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

Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic recognition memory. *Psychological Review, 108,* 452–478.

Dosher, B. A. (1984). Discriminating preexperimental (semantic) from learned (episodic) associations: A speed-accuracy study. *Cognitive Psychology, 16,* 519–555.

Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology.* New York: Teachers College, Columbia University.

Eich, J. E., Weingartner, H., Stillman, R. C., & Gillin, J. C. (1975). State-dependent accessibility of retrieval cues in the retention of a categorized list. *Journal of Verbal Learning & Verbal Behavior, 14,* 408–417.

Elmes, F. J., Adams, C., & Roediger, H. L. (1970). Cued forgetting in short-term memory: Response selection. *Journal of Experimental Psychology, 86,* 103–107.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* London: Cambridge Press.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review, 62,* 145–154.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Geiselman, R. E., & Bagheri, B. (1985). Repetition effects in directed forgetting: Evidence for retrieval inhibition. *Memory & Cognition, 13,* 57–62.

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

Geiselman, R. E., & Panting, T. M. (1985). Personality correlates of retrieval processes in intentional and unintentional forgetting. *Personality and Individual Differences, 6,* 685–691.

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

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning & Verbal Behavior, 5,* 351–360.

Godden, D., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and under water. *British Journal of Psychology, 66,* 325–331.

Goodwin, D. W., Powell, B., Bremer, D., Hoine, H., & Stern, J. (1969). Alcohol and recall: State dependent effects in man. *Science, 163,* 1358–1360.

Gruppuso, V., Lindsay, D. S., & Kelley, C. M. (1997). The process-dissociation procedure and similarity: Defining and estimating recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 259–278.

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

Humphreys, M. S. (1976). Relational information and the context effect in recognition memory. *Memory & Cognition, 4,* 221–232.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic and procedural tasks. *Psychological Review, 96,* 208–233.

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

Jang, Y., & Huber, D. E. (2008). Context retrieval and context change in free recall: Recalling from long-term memory drives list isolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 12–27.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory &* Cognition, *24,* 103–109.

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

Light, L. L., & Carter-Sobell, L. (1970). Effects of changed semantic context on recognition memory. *Journal of Verbal Learning and 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 and Language, 58,* 465–479.

Macht, M. L., Spear, N. E., & Levis, D. J. (1977). State-dependent retention in humans induced by alterations in affective state. *Bulletin of the Psychonomic Society, 10,* 415–418.

MacLeod, C. M. (1975). Long-term recognition and recall following directed forgetting. *Journal of Experimental Psychology: Human Learning and Memory, 1,* 271–279.

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., Dodd, M. D., Sheard, E. D., Wilson, D. E., & Bibi, U. (2003). In opposition to inhibition. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 163–214). New York: Elsevier Science.

Malmberg, K. J. (2008a). Recognition memory: A review of the critical findings and an integrated theory for relating them. *Cognitive Psychology, 57,* 335–384.

Malmberg, K. J. (2008b). Towards an understanding of individual differences in episodic memory: Modeling the dynamics of recognition memory. In A. Benjamin & B. Ross (Eds.), *The psychology of learning and motivation: Skill and strategy in memory use* (Vol. 48, pp. 313–349). London: Academic Press.

Malmberg, K. J., Holden, J. E., & Shiffrin, R. M. (2004). Modeling the effects of repetitions, similarity, and normative word frequency on old-new recognition and judgments of frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 319–331.

Malmberg, K. J., Lehman, M., & Sahakyan, L. (2006). On the cost and benefit of taking it out of context: Modeling the inhibition associated with directed forgetting. *Proceedings of the 28th Meeting of the Cognitive Science Society,* 549–554.

Malmberg, K. J., & Murnane, K. (2002). List composition and the word-frequency effect for recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 616–630.

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.

Malmberg, K. J., & Xu, J. (2007). On the flexibility and on the fallibility of associative memory. *Memory & Cognition, 35,* 545–556.

Malmberg, K. J., Zeelenberg, R., & Shiffrin, R. M. (2004). Turning up the noise or turning down the volume? On the nature of the impairment of episodic recognition memory by Midazolam. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 540–549.

McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 563–582.

Mensink, G. J. M., & Raaijmakers, J. G. W. (1989). A model for contextual fluctuation. *Journal of Mathematical Psychology, 33,* 172–186.

Mulligan, N. W., & Hirshman, E. (1997). Measuring the bases of recognition memory: An investigation of the process dissociation framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 280–304.

Murdock, B. B. (1997). Context and mediators in a theory of distributed associative memory (TODAM2). *Psychological Review, 104,* 839–862.

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

Norman, K. A., Newman, E. L., & Detre, G. J. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review, 114,* 887–953.

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

Roediger, H. L., & Crowder, R. G. (1972). Instructed forgetting: Rehearsal control or retrieval inhibition (repression)? *Cognitive Psychology, 3,* 244–254.

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

Sahakyan, L. (2004). The destructive effects of the "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,* 195–206.

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. (1970). Forgetting, trace erosion or retrieval failure? *Science, 168,* 1601–1603.

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

Shiffrin, R. M., & Steyvers, M. (1998). The effectiveness of retrieval from memory. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 73–95). London: Oxford University Press.

Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 460–471.

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

Tulving, E. (1983). *Elements of episodic memory.* Oxford, England: Clarendon Press.

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

Weiner, B. (1968). Motivated forgetting and the study of repression. *Journal of Personality, 36,* 213–234.

Weiner, B., & Reed, H. (1969). Effects of the instructional sets to remember and to forget on short-term retention: Studies of rehearsal control and retrieval inhibition (repression). *Journal of Experimental Psychology, 79,* 226–232.

Wenzel, A., Pinna, K., & Rubin, D. (2004). Autobiographical memories of anxiety-related experiences. *Behavior Research and Therapy, 42,* 329–341.

Winograd, E. (1968). List differentiation, recall, and category similarity. *Journal of Experimental Psychology, 78,* 510–515.

Woodward, A. E., & Bjork, R. A. (1971). Forgetting and remembering in free recall: Intentional and unintentional. *Journal of Experimental Psychology, 89,* 109–116.

Woodward, A. E., Bjork, R. A., & Jongeward, R. H. (1973). Recall and recognition as a function of primary rehearsal. *Journal of Verbal Learning and Verbal Behavior, 12,* 608–617.

Woodward, A. E., Park, D. C., & Seebohm, K. (1974). Directed forgetting as a function of explicit within-list cuing and implicit postlist cuing. *Journal of Experimental Psychology, 102,* 1001–1006.

Xu, J., & Malmberg, K. J. (2007). Modeling the effects of verbal- and non-verbal pair strength on associative recognition. *Memory & Cognition, 36,* 1351–1359.

(*Appendix follows*)

Appendix

Earlier Modeling Attempts

Various models were attempted before we settled on the current set of assumptions. Descriptions of models are in Table A1. Context-only models were originally attempted (Models 1–4); however, these were unable to account for all of the data. All of these models had the same basic process for directed forgetting as the current model—an increased context change between lists, given the forget instruction. Models 1–3 had context cues that changed at various rates, and different cues were attempted during recall, but each of these models failed to provide list-discrimination. It was only when list contexts were reinstated to be used at test (Model 4) that list-discrimination was possible; however, buffer operations were necessary to also produce serial position curves (Model 5).

Although Model 5 was able to produce list discrimination and serial position curves, it was unable to produce costs and benefits of directed forgetting. Model 6 included a component in which the forget instruction harms the ability to reinstate the context of L_1 to be used as a cue at test. This produced costs and benefits; however, intrusion rates were high. Adding a mechanism for reducing intrusions led to the current version of the model (Model 7).

Table A1
Description of Earlier Versions of the Present REM Models and the Problems They Have in Accounting for the Data

Model	Assumptions	Problems
1	Forget instruction only increases context change between lists	Produced recency; failed to produce list discrimination or FRPs
2	Context features changed at different rates; rapidly changing features not used in retrieval cue	Produced recency; failed to produce list discrimination or FRPs
3	Context features changed at different rates; subset of random features used to probe memory	Produced recency; failed to produce list discrimination or FRPs
4	Context features changed at same rate; context of to-be-recalled list was reinstated to use as a cue to recall	Produced recency; list discrimination; failed to produce benefits of directed forgetting or serial position curves
5	Same as Model 4 but with the added buffering operations	Produced recency; list discrimination; serial position curves; failed to produce costs and benefits of directed forgetting
6	Same as Model 5 but the instruction to forget harms ability to reinstate context of L_1 as a cue	Produced recency; list discrimination; serial position curves; costs and benefits of directed forgetting; intrusion rates too high
7	Current model; same as Model 6 but includes mechanism for reducing intrusions	Produced recency; list discrimination; serial position curves; costs and benefits of directed forgetting; reasonable intrusion rates

Note. FRP = first-recall probabilities; L_1 = List 1; REM = retrieving effectively from memory.

Received April 2, 2008
Revision received February 2, 2009
Accepted February 4, 2009 ■