Mem Cogn (2012) 40:844–860
DOI 10.3758/s13421-012-0198-0

Context change and retrieval difficulty in the list-before-last paradigm

Lili Sahakyan · Hannah E. Hendricks

Published online: 6 March 2012
© Psychonomic Society, Inc. 2012

Abstract Using the "list-before-last" paradigm (Jang & Huber, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 112–127, 2008; Shiffrin, *Science*, *168*, 1601–1603, 1970), we examined whether the difficulty involved in retrieving a previously studied list affects the recall of the current list. Participants studied three lists (L1, L2, and L3), and between L2 and L3 study they either engaged in retrieval of L1 or solved math problems for the same duration of time. After L3 encoding, all participants recalled L2. We examined accurate recall of L2, intrusions from L3, and also the first-response function from L2 across all four experiments. In Experiments 1 and 2, retrieval difficulty was manipulated by presenting participants with variable numbers of letter cues for the L1 words, with some participants receiving more cues than others. In Experiment 1, L1 and L2 words shared the first two letters in common to create potential item interference, whereas in Experiment 2, they did not share common cues. In Experiments 3 and 4, we manipulated retrieval difficulty using a delay manipulation—L1 was encoded 1, 24, or 72 h prior to the session during which L2 and L3 were encoded. In Experiment 3, L1 retrieval involved a cued recall test, whereas in Experiment 4, it involved a free recall test. The results of all of these experiments showed that, as compared to solving math problems, retrieving L1 led to forgetting of L2, reduced intrusions from L3, and reduced first-response functions from L2. However, all of the dependent measures were invariant across the various manipulations of retrieval difficulty. We discuss the results in terms of the context-change interpretation of prior-list retrieval.

Keywords Context change · Retrieval · Forgetting

L. Sahakyan (✉) · H. E. Hendricks
Department of Psychology,
University of North Carolina at Greensboro,
296 Eberhart Building, P.O. Box 26170, Greensboro,
NC 27402-6170, USA
e-mail: l_sahaky@uncg.edu

Imagine that at the end of a Wednesday class, you present a pop quiz in which students are required to write down the most important points from a lecture that you delivered a few days ago, during Monday's class. This pop quiz places demands on the memory system, which has to be able to reinstate the context of the previous lecture in order to selectively focus the search processes on the subset of relevant items (e.g., J.R. Anderson & Bower, 1972; Hintzman, 1988; Howard & Kahana, 2002; Lehman & Malmberg, 2009; Mensink & Raaijmakers, 1988; Raaijmakers & Shiffrin, 1981). The success with which one may be able to reinstate the previous context could vary from situation to situation. For example, if the pop quiz provides no retrieval cues for the previously delivered lecture, it would make the reinstatement of the previous context substantially more difficult than if it contained some retrieval cues (e.g., the topic of the lecture). Likewise, it would be more difficult to reinstate the previous context if the quiz required retrieving facts from a lecture delivered a week ago rather than a day ago. In the experiments reported in this article, we examined how reinstatement of the context of previously learned information affects memory for more recently learned information. More specifically, we examined whether the degree of difficulty in reinstating the previous context differentially impairs memory for more recently learned information.

In addressing these questions, we utilized a variant of the *list-before-last* paradigm (e.g., Jang & Huber, 2008; Shiffrin, 1970). In this paradigm, participants encode several lists, and after study of each list (except the first), they are told to retrieve not the current, but rather the previous list. The prior list is therefore termed the *target* list, whereas the current list is termed the *intervening* list. Previous research with this paradigm has suggested that the act of retrieving the prior lists in between the encoding of other lists creates an internal context change, leading to list isolation. Specifically, Jang and Huber intermixed two types of trials into the list-before-last paradigm —on some trials, participants were asked to retrieve the prior list, whereas on others they simply waited to encode the next list without being asked to retrieve the prior list. In addition to manipulating the type of trials, the authors also crossed the lengths of the target and the intervening list, which were either short or long. The results showed that on retrieval trials, correct recall of the target list was mainly driven by the length of the target list (with shorter lists leading to greater recall), whereas there was no effect of the intervening list length. In contrast, on trials in which participants were not asked to retrieve the prior list, correct recall was influenced not only by the target list length, but also by the intervening list length with participants recalling more items when the target list was short and also when the intervening list was short versus long. These findings suggest that on retrieval trials, participants were able to exclude the intervening list during the recall period and focus mainly on the target list, explaining why there was no effect of the intervening list length. In contrast, during the no-retrieval trials, they were less able to exclude the intervening list, and hence the length of the intervening list mattered. Interestingly, the retrieval trials not only produced the benefits of reduced interference from the intervening list, but also led to costs. Specifically, as compared to the no-retrieval trials, the retrieval trials lowered the probability of recall of the target list, and this was particularly true when comparing the conditions in which the delay between the target list and the time of test was controlled for. Thus, Jang and Huber found that prior-list retrieval led to reduced interference from the intervening list along with forgetting of the target list—two outcomes that are typically associated with external context change (e.g., Bjork & Richardson-Klavehn, 1989; Smith & Vela, 2001).

To explain these findings, Jang and Huber (2008) proposed that the retrieval trials produced an internal/endogenous context change. The idea behind their explanation is that during no-retrieval trials, context drifts from one list to the next in a gradual fashion, so that the contexts between the two adjacent lists are somewhat similar to each other. Retrieval trials, on the other hand, disrupt the similarity of the list contexts, creating a change in endogenous context at the time of memory storage, thereby isolating and segregating the lists. During the no-retrieval trials, when participants attempt to retrieve the previous list, they might either rely on the context of the recent list as they search their memory, because it is sufficiently similar to the previous list, or attempt to directly reinstate the context of the previous list. Note that although the context associated with either list would be an appropriate cue to allow accessing items from the previous list, the contextual similarity between the two lists could make the list discrimination more difficult during recall (as if it were one long list), and therefore the recent list items would be likely to intrude during recall of the previous list. In contrast, a retrieval trial between the two lists disrupts their contextual similarity, and therefore if participants relied on the recent context as a retrieval cue, it would be a poor retrieval cue for the previous list, leading to a lower probability of recall of the previous list (i.e., forgetting would be observed). Likewise, directly reinstating the context of the previous list might be more difficult on the retrieval trials than on the no-retrieval trials, which would also lower the probability of recall of the previous list. On the upside, participants might be less likely to have intrusions from the recent list during recall of the previous list, because contextual dissimilarity between those lists might allow for better discrimination between them.

Overall, Jang and Huber (2008) explained performance in the list-before-last paradigm by invoking the importance of episodic context—a theoretical construct that is part of memory models (e.g., J. R. Anderson & Bower, 1972; Hintzman, 1988; Howard & Kahana, 2002; Lehman & Malmberg, 2009; Mensink & Raaijmakers, 1988; Murdock, 1982; Raaijmakers & Shiffrin, 1981; Rohrer & Wixted, 1994). The postulation of this construct is necessary in part because we do not experience events in a vacuum; instead, events take place in a certain environment, directly tied to our internal states. The joint contribution of the temporal–spatial, social–emotional, and internal factors is termed *context*, and it is assumed that during encoding, contextual information gets stored in the memory trace along with the items that we are aiming to memorize. Furthermore, the similarity of the context between encoding and test affects the probability of recall, with context change between encoding and test being one of the causes of forgetting (e.g., Estes, 1955; McGeoch, 1932; Mensink & Raaijmakers, 1988).

The process of context change is defined differently in some models, and we have described it briefly because it has implications for our experiments. Some models represent context as a vector of elements (e.g., Estes, 1955; Mensink & Raaijmakers, 1988). A subset of these elements are perceived by the participant at any given time, thus becoming active elements. Over time, some active elements become inactive, and inactive elements become active, causing random contextual fluctuation. Thus, in these models contextual drift takes place independent of the items on the list. In contrast, in the temporal context model (TCM), contextual drift is assumed to be partially driven by the items themselves (Howard & Kahana, 2002). According to the TCM model, the presentation of each new item during study leads to the retrieval of its preexperimental context, which is used to update the current state of the context of the list. This constant updating leads to a gradual shift in the context of the study list. A similar process takes place during retrieval, where both the preexperimental context and the studied context are retrieved along with an item. Thus, successful retrieval of an item causes the current context to be updated to reflect prior experiences with the retrieved item, thereby causing contextual drift. Overall, the assumptions regarding whether contextual change is driven by, or independent of, the items on the list bear relevance to our experiments, which are described below.

Motivation for the present experiments

Our experiments were inspired by a recent investigation from our lab on the mental context change, where we utilized a *diversionary thought* paradigm that is similar in spirit to the list-before-last paradigm. In that study, participants studied two lists (L1 and L2), and they were asked either to daydream between the two lists or to solve math problems as the control task. Following L2 encoding, we tested memory for L1 items with free recall. Daydreaming between the lists involved retrieving events from either the temporally distant past or spatially distant places (Delaney, Sahakyan, Kelley, & Zimmerman, 2010). Specifically, in the first experiment, some participants were told to imagine their current home, which they presumably had visited in the last few hours, whereas others were told to imagine their parents' home, which they had not visited for quite some time. We hypothesized that the farther back in time participants were to mentally travel to, the more likely that context would be different from the current context, and thus the more forgetting of L1 items we might observe in those conditions. Indeed, we obtained more forgetting of L1 in the group who were told to imagine their parents' home, as compared to those who imagined their current home, with both groups recalling fewer L1 items than the control math group. Interestingly, the longer it had been since participants had visited their parents' home, the more forgetting of L1 items they showed.

Our interpretation of these findings was that imagining temporally distant events creates a greater mental context change between the two lists than does imagining more nearby contexts, leading to greater forgetting. Thus, we explained the different degrees of forgetting of L1 items by assuming that various conditions of daydreaming led to different degrees of context change between the lists. An alternative explanation of these findings (something that was expressed to the first author of the present study during the presentations of that earlier work) is that it may be harder to retrieve more temporally distant contexts, and therefore the results could be due to greater mental effort involved in imagining the parents' home versus the current home. We refer to this as the *difficulty* hypothesis. Interestingly, informal conversations with college students, who were the typical participants in that study, revealed quite the opposite sentiment: They expressed that it was much easier to imagine their parents' home than their current home, in part because their parents' home was more familiar and full of rich memories than was the current home (e.g., dorm room), which changed on a yearly basis as students changed their residences.

Motivated by a retrieval difficulty interpretation of our data, we conducted several experiments to more broadly test the notion of whether more difficult reinstatement of some prior context creates greater forgetting of more recently learned information. In all of our experiments, we utilized an abbreviated version of the list-before-last paradigm. Participants studied three lists (L1, L2, and L3), and between encoding of L2 and L3, they either engaged in retrieval of L1 or solved math problems. The ease with which L1 could be retrieved was manipulated by providing different numbers of letter cues to L1 items. Some participants received the first two letters, whereas others received either the first or the second letter of the L1 words. Upon encoding L3, all participants were asked to recall L2 without cues (i.e., free recall). The experimental procedures are shown in Fig. 1.

Theoretical predictions and dependent measures

On the basis of prior research, we expected to observe impaired recall of L2 in all retrieval conditions, relative to the math condition. The more critical question was whether more difficult retrieval of L1 items between L2 and L3 would lead to more forgetting of L2 items at the time of final test. Different theoretical accounts make different predictions regarding the outcomes of this study.

According to the TCM model, context change is partly driven by the retrieval of items. Therefore, if more L1 items are retrieved (e.g., when the first two letters are provided as a cue), the more the context should change between L2 and L3, and therefore more forgetting of L2 should be observed in the easy retrieval condition than in the difficult retrieval condition of L1 items. In contrast, other models do not tie context change to the retrieval of items from the list, and instead assume that contextual drift takes place independently of the items themselves (e.g., Mensink & Raaijmakers, 1988). Therefore, according to those models, L2 should be less accessible in the retrieval conditions than in the math condition because retrieval of the prior list would instigate a context change; however, the magnitude of L2 forgetting should be invariant across the conditions of L1 retrieval

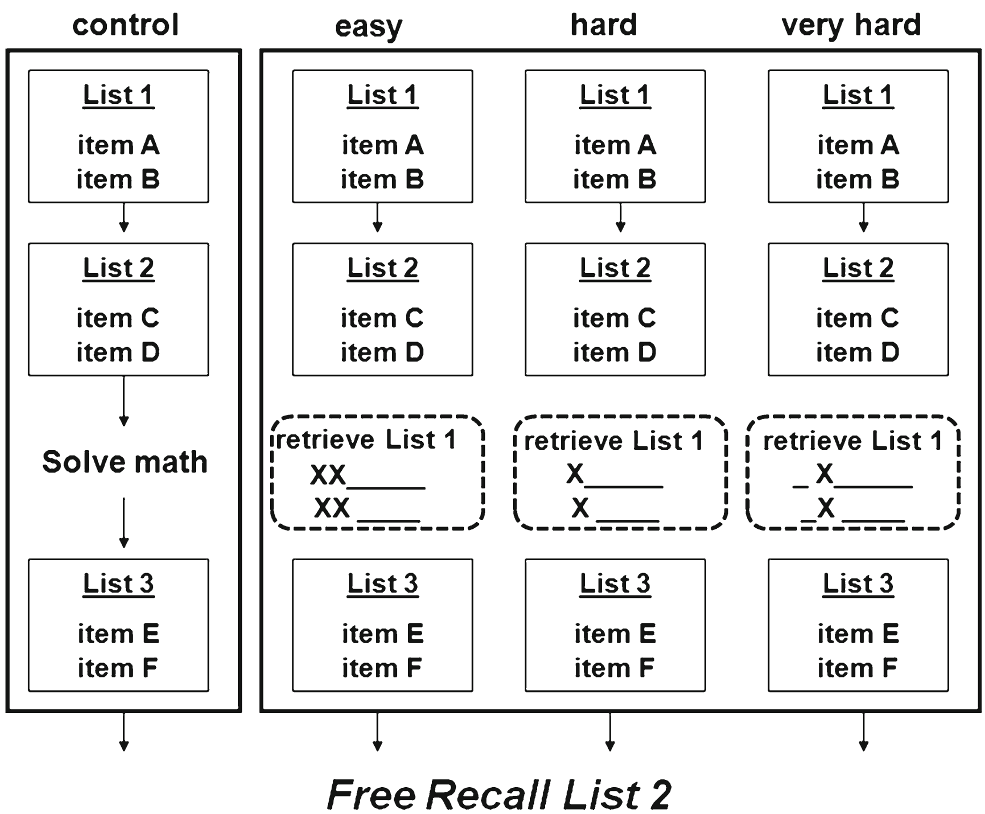

Fig. 1 Experimental design and procedures of Experiments 1 and 2

difficulty, because the degree of context change is not a function of retrieved items.

In addition to examining overall L2 recall, we also planned to examine a new dependent measure that had not been investigated by Jang and Huber (2008) in their research. We proposed to analyze the first-response probability from L2 recall, which is a measure of how participants initiate their recall. It indicates the responses that in theory have the strongest overlap with the cues being used during retrieval (e.g., Howard & Kahana, 2002; Lehman & Malmberg, 2009). Because the first attempt to search memory in free recall relies on the context cues (e.g., Raaijmakers & Shiffrin, 1981), the very first item that participants recall from L2 can indicate how accessible L2 context is in the different conditions of the experiment. If retrieval creates context change between the lists, then we should observe reduced first-response functions between the math group and the retrieval groups. Furthermore, if context change is driven in part by the retrieval of items, as suggested the TCM model, retrieving more L1 items between L2 and L3 should lead to a greater context change between L2 and L3 than would retrieving fewer L1 items. Therefore, the first item recalled from L2 should be less accessible in the easy than in the hard retrieval condition. In contrast, if context change is not driven by the recalled items, the number of items recalled from L1 should have no major impact on the magnitude of context change, and therefore the first response functions from L2 should be similar across the various conditions of L1 retrieval difficulty.

We also planned to examine intrusions from L3 during L2 recall. If retrieval isolates the list from interference of the intervening list, we should expect more L3 intrusions in the math condition than in the conditions involved in L1 retrieval between L2 and L3. If the degree of context change is driven by the number of retrieved L1 items, the easy retrieval condition should produce more list isolation, and hence fewer intrusions from L3 than the hard retrieval condition. However, we may not observe any differences in intrusions across the various conditions of retrieval difficulty, which would suggest that context change takes place independent of the number of retrieved items (if at all). Finally, failing to observe reduction in L3 intrusions between the math and retrieval groups would be inconsistent with the entire context-change interpretation of prior-list retrieval.

All of the predictions that we have considered thus far treat each list in the list-before-last paradigm as the unit of information and invoke context change between the lists as an explanatory mechanism. However, we would be remiss not to consider potential competition between the *items* across the lists at the time of retrieval. For example, one could think of L2 recall at the time of the final test suffering from retrieval-induced forgetting (RIF) induced by the earlier retrieval of L1 items during the experiment. The RIF phenomenon refers to the finding that practicing retrieval of some items from memory creates forgetting of other items that are in competition with the retrieved information (e.g., M. C. Anderson, Bjork, & Bjork, 1994). Different theoretical mechanisms have been postulated to explain what causes forgetting in RIF. According to the *inhibitory account*, to allow for successful retrieval of items during the retrieval practice stage, inhibitory processes are invoked to overcome competition from the remaining items that are linked to the same retrieval cue. Therefore, when the competing information is strong, it requires more inhibition in order to be suppressed, leading to greater forgetting of those items on a later test (for reviews, see M. C. Anderson & Levy, 2010; Storm, 2010). In our experiment, if L1 items are retrieved more easily due to the provision of two-letter cues, they may require less inhibition of L2 items, and therefore less forgetting of L2 items might be observed on the final test, as compared to the condition in which L1 items are harder to retrieve. An exactly opposite prediction is made by the *interference* account of RIF (e.g., J. R. Anderson, 1983; Mensink & Raaijmakers, 1988; Raaijmakers & Shiffrin, 1981). According to the this view, retrieved items are strengthened as a result of retrieval practice, and at the time of final test, they come to mind more easily, blocking access to the remaining items. Therefore, this account predicts that the more L1 items are retrieved between the two lists, the more they should block access to L2 at the time of final test, and therefore more forgetting of L2 should be observed in the easy L1 retrieval condition than in the hard retrieval condition.

The examination of L3 intrusions during L2 recall could be one way to differentiate between the predictions of context models and the inhibitory account of RIF. If items on a given list must be suppressed in order to allow for retrieval of another list (as assumed by the inhibitory account), L3 items must be suppressed at the time of final test in order to allow for retrieval of L2 items (much as L2 items should be suppressed to allow for retrieval of L1 items between L2 and L3). Since L2 items are tested after L3 in all conditions of the experiment, then according to the inhibitory account, the math group should suppress L3 items in order to permit retrieval of L2 items, much like the remaining conditions that engage in retrieval of L1. In other words, there should be no differences in the magnitude of L3 intrusions during L2 recall at the time of final test between any of the experimental conditions. In contrast, the context-change accounts predict that L3 items should be less likely to intrude during L2 recall in the retrieval conditions than in the math condition, because retrieval of L1 between L2 and L3 should isolate L2 from interference from L3 items.

Finally, according to the *difficulty* explanation, the more mental effort that is required to retrieve L1 items, the more forgetting of L2 we might observe, as compared to conditions in which L1 retrieval is easier.

Experiment 1

Participants studied three lists, and after L2 encoding, some solved math problems, while others were told to retrieve L1 using a variable number of letter cues provided to them. After L3 encoding, all participants recalled L2. Our main concern was to examine whether the retrieval difficulty of L1 impacted the magnitude of forgetting of L2 recall, intrusions from L3 during L2 recall, and the first-response functions from L2. To enhance the chances of observing potential item-interference effects across the lists at the time of retrieval, L1 and L2 words were selected such that they always started with the same first two letters. Therefore, during retrieval of L1, there were always two items that could fit the provided cue, leading to potential RIF-style competition. The number of provided letter cues not only would influence retrieval difficulty, but could also be viewed as a manipulation that varied the degrees of item interference versus context interference (i.e., with fewer letters, the task would become closer to free recall, based on context alone).

Pilot data

Prior to the experiment, we collected pilot data from $N = 60$ undergraduates at the University of North Carolina at Greensboro (UNCG) who were never involved in the subsequent experiments. They were given two lists of 12 unrelated English nouns to study for a later memory test; the order of the lists was counterbalanced during presentation. During encoding, these participants were instructed to perform a pleasantness rating task on each word using a 5-point Likert scale. The words were presented one at a time on the computer screen, at a rate of 4 s/word. Following L2 encoding, participants were given 60 s to recall the words from L1 when provided with a set of cues. The cues were typed on a sheet of paper with blank lines next to them. One third of the participants were provided with the first two letters of each L1 word and instructed to use those cues to retrieve L1. The remaining participants were given only either the first or the second letter of each L1 word. Afterward, the participants were asked to use a scale from 1 to 7 to answer the question "How much overall effort did you put in the retrieval task—that is, how hard did you try?" As expected, L1 recall increased gradually with the number of provided retrieval cues, $F(2, 57) = 35.76$, $MSE = .020$, $p < .01$. Importantly, participants indicated putting more effort into the task when only the second letter of the L1 words was provided to them ($M = 5.95$, $SD = 1.00$), relative to those who received the first letter of the L1 words ($M = 4.90$, $SD = 1.37$), who in turn gave higher effort ratings than did participants receiving the first two letters of the L1 words ($M = 3.90$, $SD = 1.41$), $F(2, 57) = 12.94$, $MSE = 1.624$, $p < .01$. Overall, these data suggest that with an increased number of provided cues, the retrieval task got easier.

Method

Participants The participants in the first experiment were 144 UNCG undergraduates who received extra credit in their classes for participation. They were tested in small groups of no more than five participants at a time.

Materials We created three different sets of words (A, B, and C) containing 12 unrelated nouns per set. The words were 5–6 letters long, with an average frequency of 81.68 (Kučera & Francis, 1967). The Set A and Set B words shared their first two letters in common, and they did not overlap with Set C words. During encoding, Set C was always the third list in the presentation order, whereas the presentation order of Sets A and B was fully counterbalanced. There were two different presentation orders of the words per list.

Procedure The participants were informed that they would be presented some words to study for a later memory test, without being told how many lists they were going to study. Then three lists of words were presented to participants (L1, L2, and L3). The words were presented one at a time on the computer screen, at a rate of 4 s/word, and participants performed a pleasantness rating task on each word. After L1, they were told to get ready to study another list, which was encoded in the same fashion, using the same pleasantness rating task. Following L2, some participants were given 60 s to solve three-digit multiplication problems (e.g., *control* condition), whereas the remaining participants were asked to recall L1 words from a set of cues for the same duration of time (e.g., *retrieval* conditions). The cues were typed up on a sheet of paper with blank lines next to them. In the *easy* retrieval condition, participants were provided the first two letters and were told to use them to retrieve L1 words. In the *hard* retrieval condition, participants were given only the first letter, whereas in the *very hard* retrieval condition, they received only the second letter of each L1 word. Following retrieval of L1 or the math task, the experimenter collected the response sheets and informed participants to get ready to study another list (e.g., L3) using the same orienting task. After L3 encoding, all participants were given 60 s and were told to free recall the words from L2 on a new sheet of paper.

Results and discussion

First, we assessed L1 recall as a function of retrieval difficulty, then we analyzed L2 recall as a function of L1 retrieval condition, and finally we examined intrusions from L3 and the first-response probabilities from L2.

List 1 recall The results of L1 recall are summarized in Fig. 2 (top panel). A one-way ANOVA of L1 recall by groups (easy retrieval vs. hard retrieval vs. very hard retrieval) revealed significant differences between the groups, $F(2, 105) = 16.70$, $MSE = .024$, $p < .001$. The easy-retrieval group remembered more items than did the hard-retrieval group, $t(70) = 3.69$, $p < .001$, which in turn had better memory than the very-hard-retrieval group, $t(70) = 1.96$, $p = .05$.

List 2 recall There were also significant differences in L2 recall between the four groups (math vs. easy retrieval vs. hard retrieval vs. very hard retrieval), $F(3, 140) = 7.51$, $MSE = .024$, $p < .001$ (see Fig. 2, bottom left panel). Recall was higher in the math group than in the easy-retrieval group [$t(70) = 3.16$, $p < .01$], the hard-retrieval group [$t(70 = 4.17$, $p < .001$], or the very-hard-retrieval group [$t(70) = 3.69$, $p < .001$]. If the math group was excluded from the analyses, the remaining conditions did not differ from each other, $F < 1$.

These findings suggest that solving math between the lists does not create a context change nearly as much as does the act of retrieving L1 between L2 and L3 does. This was evident in the fact that L2 recall was impaired in the retrieval groups as compared to the math group. However, the degree of context change does not appear to be driven by the number of items retrieved from L1, because all three retrieval conditions suffered to the same extent from prior-list retrieval. These findings are inconsistent with the TCM model, which predicts that with the retrieval of each item, the current context would get updated, and therefore the more items that are retrieved from L1, the more context change should be observed between L2 and L3. The results are more consistent with an interpretation that does not tie context change to the retrieval of items (e.g., Mensink & Raaijmakers, 1988). It appears that the act of thinking back to L1 is sufficient to instigate an internal context change, regardless of how many items are retrieved from that list. Given that all three retrieval conditions retrieved at least one item from L1, it is plausible that retrieval of a single item from an earlier list is sufficient to instigate the maximum extent of context change. It could also be that successful retrieval of prior-list items has nothing to do with context change, but that merely thinking back to the previous experience, without necessarily retrieving specific items from that list, is sufficient to create context change. Finally, L2 recall invariance across conditions of L1 retrieval is inconsistent with the retrieval difficulty interpretation, and it is also inconsistent with various RIF explanations, including inhibition and item interference.

List 3 intrusions The proportions of L3 words intruding during L2 recall were analyzed with a one-way ANOVA as a function of the four groups. The results revealed significant

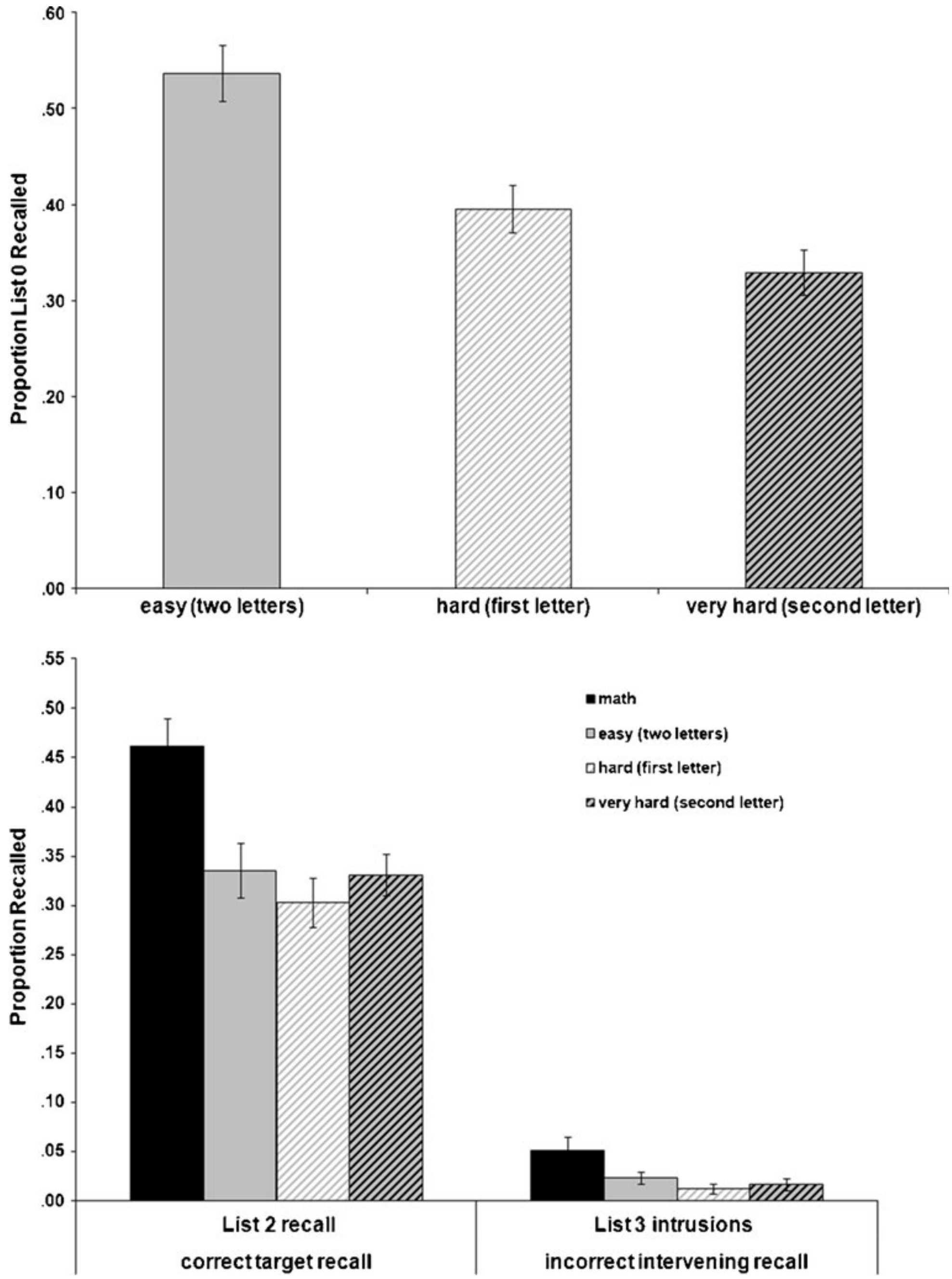

Fig. 2 Experiment 1: (Top) Proportions of List 1 recalled as a function of retrieval difficulty. (Bottom) Proportions of List 2 recalled (left panel) and proportions of List 3 intrusions (right panel) as a function of experimental condition. The errors bars represent±*SE*s of the means

differences among the conditions, $F(3, 140) = 4.36$, $MSE = .003$, $p < .01$ (see Fig. 2, bottom right panel). The math group had significantly more intrusions than did the hard-retrieval group [$t(70) = 2.76$, $p < .01$], the very-hard-retrieval group [$t(70) = 2.33$, $p < .05$], or the easy-retrieval group, although the latter comparison fell short of conventional significance [$t(70) = 1.88$, $p = .06$]. When the math group was excluded from the analyses, the three retrieval conditions did not differ from each other, $F < 1$. These findings are consistent with the context-change interpretation because they demonstrate that, relative to the math control condition, the retrieval conditions isolated L2 from interference with L3 items. The absence of differences in L3 intrusions among the three retrieval groups appears to be incompatible with the TCM model. If more items are retrieved from L1, the TCM model would assume that there should be a greater context change between L2 and L3, and therefore L2 should be more isolated from L3. Given that we found no differences in intrusions among the three groups, this could be a further indication that context change is not driven by the number of retrieved items. Note that the retrieval manipulation between the lists virtually eliminated L3 intrusions, which complicates direct comparisons between the various context-change models. As discussed earlier, the pattern of reduced intrusions in the retrieval groups as compared to the math group is incompatible with an inhibitory interpretation, which would predict equivalent intrusion rates across all conditions.

First-response probabilities from List 2 Finally, we examined how people initiated their L2 recall. This could indicate how accessible L2 context was in the different conditions of the experiment, because the first retrieval attempt in free recall relies on contextual information. The first-response function from L2 is shown in Fig 3. It demonstrates that the math group tended to start their recall from the first item on L2 much more than from anywhere else in the list. Specifically, the probability of recalling the item from the first serial position was around .36 in the math group, whereas the remaining serial positions produced much lower probabilities [$F(11, 385) = 6.60$, $p < .001$, for the effect of serial position in the math group]. This trend was also present in the three retrieval groups, but it appeared to be much more depressed [$F(11, 1177) = 1.96$, $p < .05$, for the effect of serial position in the retrieval groups]. The only significant difference between the four groups emerged in the recall of the item from the first serial position, $F(3, 140) = 2.65$, $p = .05$. When the math group was excluded from the analyses, the three retrieval conditions did not differ from each other ($F < 1$), so we collapsed the data across the L1 difficulty factor. The math group had significantly better access to the first item from List 1 than did the retrieval groups, $t(142) = 2.82$, $p < .01$. These results provide further evidence that prior-list retrieval—regardless of its difficulty level—creates a context change, making it difficult to reinstate the beginning of L2 context during the final test.

Experiment 2

In Experiment 1, two items could fit the provided cues at the time of L1 retrieval, potentially leading to RIF-style competition. However, the overall pattern of results across several dependent measures was more compatible with a context-change interpretation than with an item inhibition interpretation or an item interference/blocking interpretation. The latter approach would predict that the greatest degree of successful retrieval of competing information (i.e., the easy condition) should be associated with less retrieval of the second list. However, the results showed that L2 recall was not affected by the number of L1 items retrieved. The goal of Experiment 2 was to replicate the Experiment 1 findings by dropping the item interference manipulation, because this might allow for better observation of context effects across the various conditions of L1 retrieval difficulty. Differences in L2 recall across the various retrieval conditions might have been offset by the simultaneous manipulation of item interference across the lists. We therefore created a new set of stimuli, such that the items across L1 and L2 did not start with the same first two letters.

Method

Participants A group of 120 UNCG undergraduates participated for course credit. None of them had participated in the previous experiment.

Materials and design We prepared three new sets of words, with 12 nouns per set. The words were 5–6 letters long, with an average frequency of 69.64 (Kučera & Francis, 1967). Set A words started with different combinations of two letters than did the words from Set B or Set C. The remaining details were similar to those of Experiment 1.

Procedure The procedures were identical to those of Experiment 1.

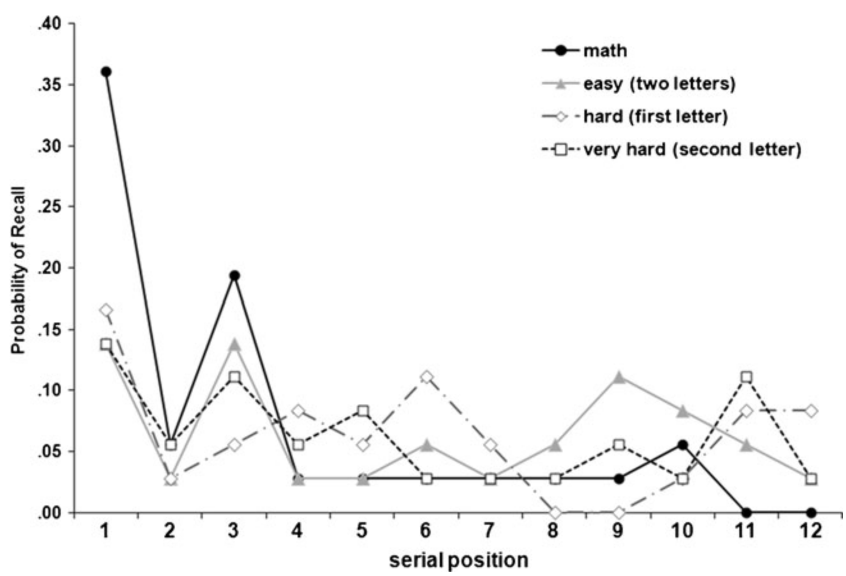

Fig. 3 First-item recall probabilities from List 2 as a function of experimental condition and serial position in Experiment 1

Results and discussion

List 1 recall The results of L1 recall are displayed in Fig. 4 (top panel). A one-way ANOVA on the proportions of L1 recalled with Group (easy retrieval vs. hard retrieval vs. very hard retrieval) as the factor revealed a significant effect of group, $F(2, 87) = 58.44$, $MSE = .024$, $p < .001$. The easy-retrieval group recalled significantly more L1 items than did the hard-retrieval group, $t(58) = 3.48$, $p < .001$. Also, the hard-retrieval group recalled significantly more items than did the very-hard-retrieval group, $t(58) = 7.33$, $p < .001$.

List 2 recall A one-way ANOVA on the proportions of L2 words recalled by each group (easy retrieval vs. hard retrieval vs. very hard retrieval vs. math) revealed a significant effect of group, $F(3, 116) = 9.79$, $MSE = .018$, $p < .001$. The results are summarized in Fig. 4 (bottom left panel). Specifically, the math group had significantly higher recall, as compared to the easy-retrieval group [$t(58) = 4.56$, $p < .001$], the hard-retrieval group [t $(58) = 3.78$, $p < .001$], and the very-hard-retrieval group [$t(58) = 4.74$, $p < .001$]. When the math group was excluded from the analyses, there was no longer an effect of group in the three retrieval groups, $F < 1$. These results fully replicate the previous experiment, despite the change in materials.

List 3 intrusions Next, we examined the proportions of L3 words that intruded during L2 recall using Group (math vs. easy retrieval vs. hard retrieval vs. very hard retrieval) as the factor in a one-way ANOVA. The results are displayed in Fig. 4 (bottom right panel). There was a significant effect of group, $F(3, 116) = 4.02$, $MSE = .004$, $p < .01$: There were more L3 intrusions in the math group than in the easy-retrieval group [$t(58) = 2.35$, $p < .05$], the hard-retrieval

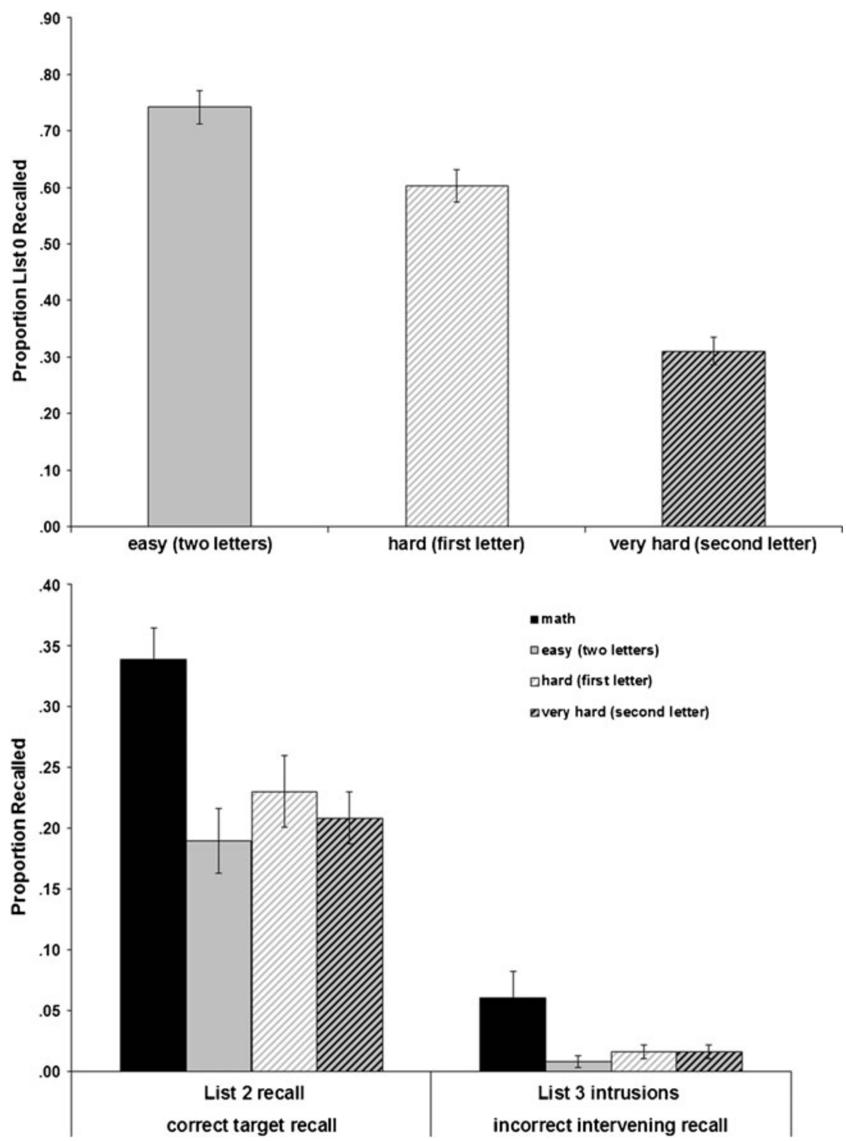

Fig. 4 Experiment 2: (Top) Proportions of List 1 recalled as a function of retrieval difficulty. (Bottom) Proportions of List 2 recalled (left panel) and proportions of List 3 intrusions (right panel) as a function of experimental condition. The errors bars represent ± *SE*s of the means

group [$t(58) = 1.97$, $p = .05$], or the very-hard-retrieval group [$t(58) = 1.97$, $p = .05$]. If the math group was excluded from the analyses, the remaining three conditions were not different from one another, $F < 1$. The intrusion results also replicate our Experiment 1 findings.

First-response probabilities from List 2 The first-recall function from L2 is shown in Fig. 5. As the figure demonstrates, participants in the math group tended to start their recall from the very first item of L2 much more than from anywhere else in the list. Specifically, the probability of initially recalling the item from the first serial position was around .33 in the math group, whereas the remaining serial positions produced probabilities no higher than .13 [$F(11, 319) = 4.41$, $p < .001$, for the effect of serial position in the math group]. In contrast, the three retrieval groups appeared to have lost access to the beginning of the L2 context, and their first recalled item was equally likely to come from any position in the list ($F = 1.18$, $p = .30$, for the effect of serial position in the retrieval groups). Significant differences between the four experimental groups emerged only in the first serial position, $F(3, 116) = 3.21$, $MSE = .127$, $p < .05$. When the math group was excluded from the analyses, the remaining three groups did not differ from each other ($F < 1$), and therefore we collapsed them together. The first item on L2 was significantly more accessible in the math group than in the retrieval groups, $t(118) = 3.13$, $p < .01$. Overall, this pattern of data replicates the previous experiment.

Experiment 3

In Experiment 3, we used a different method for manipulating L1 retrieval difficulty, by varying the delay separating L1 and the subsequent two lists. Some participants studied L1 only 1 h prior to the session in which they studied L2 and L3, whereas others studied L1 either 24 or 72 h prior to studying L2 and L3. We anticipated that it would be more difficult to retrieve L1 if it had been studied longer ago. Therefore, if the difficulty of prior retrieval has any differential effect on the magnitude of L2 recall, we would expect to observe those effects in the present experiment. The manipulation of L1 recency also has implications for the TCM model. According to this model, context changes during retrieval because successful retrieval of items causes the current context to be updated. Presumably, information retrieved from 72 h ago should entail a different context than would information studied very recently (e.g., 1 h ago). Therefore one might expect that retrieving information from longer ago should induce a greater context change.

Method

Participants The participants were 120 UNCG undergraduates who received extra credit in their psychology classes for participation. None of them had participated in the previous experiments.

Materials and design The materials were the same as in Experiment 2. The design was similar to that of the previous experiments, with a few minor changes involving the delay manipulation between L1 and the remaining two lists. As in the previous studies, we included a group that did not engage in retrieval of L1 between the presentation of L2 and L3, but instead solved math problems for the same duration of time.

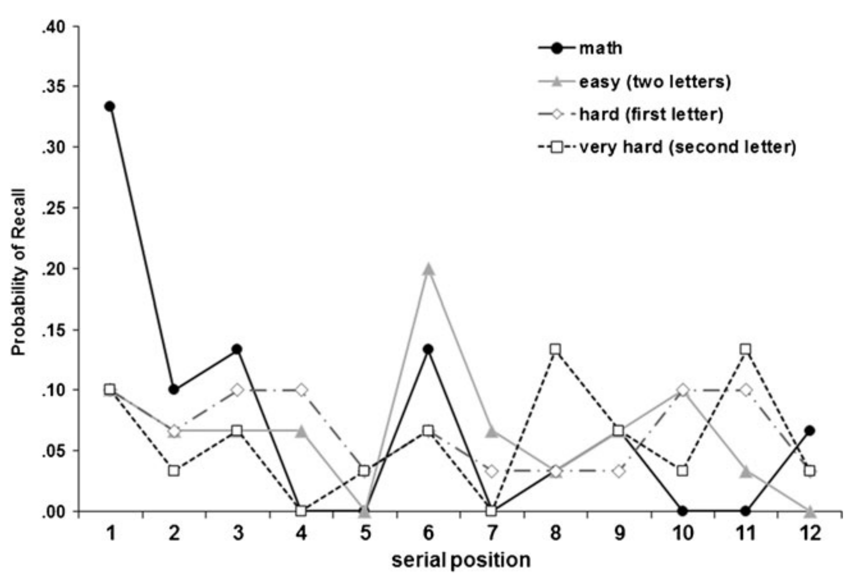

Fig. 5 First-item recall probabilities from List 2 as a function of experimental condition and serial position in Experiment 2

Procedure The participants were first shown L1 three times in different random orders each time, and they were asked to rate the pleasantness of words after each presentation. Words were presented at a rate of 4 s/word. During the first presentation of L1, participants were asked to provide a simple yes/no rating; during the second presentation, they provided a more graded response on a 1–7 scale, whereas during the third presentation, they provided the pleasantness rating on a scale from 0 to 100. We used different rating scales to ensure that participants would not automatically retrieve their earlier responses. After studying L1, the participants were told that they had to return to the second session of the experiment, which was to take place 1 h later, 24 h later, or 72 h later.[1] During the second session, participants studied only two lists (L2 and L3) using the yes/no pleasantness rating task used earlier. After L2, some participants were given 60 s to solve math problems, whereas others were asked to recall the words that they had seen during the first session (e.g., L1). Thus, there were three conditions of retrieval difficulty—*very hard retrieval* (e.g., L1 studied 72 h prior to the second session), *hard retrieval* (e.g., L1 studied 24 h prior to the second session), and *easy retrieval* (e.g., L1 studied 1 h prior to the second session). To avoid potential floor effects after a long delay, we provided the first two letters from L1 words as cues for everybody, in a same fashion as in the previous experiments. The remaining procedures followed those of the previous experiments.

Results and discussion

List 1 recall A one-way ANOVA on L1 recall using Group (easy retrieval vs. hard retrieval vs. very hard retrieval) as the factor produced a significant effect of group, $F(2, 87) = 28.11$, $MSE = .025$, $p < .001$. The easy-retrieval group remembered significantly more items from L1 than did the hard-retrieval group, $t(58) = 2.41$, $p < .05$, and the hard-retrieval group in turn retrieved significantly more items than did the very-hard-retrieval group, $t(58) = 5.08$, $p < .001$. The results are shown in Fig. 6 (top panel).

List 2 recall A one-way ANOVA on L2 recall using Group (easy retrieval vs. hard retrieval vs. very hard retrieval vs. math) as the factor showed a significant effect of group, $F(3, 116) = 9.21$, $MSE = .024$, $p < .001$. The results are shown in Fig. 6 (bottom left panel). Replicating the findings of the previous experiments, we observed higher L2 recall in the math group than in the easy-retrieval group [$t(58) = 3.59$, $p < .001$], the hard-retrieval group [$t(58) = 5.35$, $p < .001$], and the very-hard-retrieval group, [$t(58) = 4.67$, $p < .001$]. When the math group was excluded from the analysis, the remaining retrieval groups did not differ from each other, $F < 1$. These results fully replicated the findings from the previous experiments.

List 3 intrusions Using the same Group factor described in the analyses above, we examined L3 intrusion errors during L2 recall. There was a significant effect of group, $F(3, 116) = 2.89$, $MSE = .002$, $p < .05$, with more L3 intrusions in the math group than in the easy-retrieval group [$t(58) = 2.02$, $p < .05$], the hard-retrieval group [$t(58) = 2.30$, $p < .05$], and the very-hard-retrieval group [$t(58) = 2.38$, $p < .05$]. If the math group was excluded from the analyses, the remaining three conditions were not different from each another, $F < 1$. The findings are summarized in Fig. 6 (bottom right panel), and they replicate those of the previous experiments.

First-response probabilities from List 2 The probabilities of first recall from L2 are shown in Fig. 7. The math group tended to initiate recall with the first item from L2 [$F(11, 319) = 4.98$, $p < .001$, for the effect of serial position in the math group]. This trend was also present in the three retrieval groups, albeit in a substantially more depressed fashion [$F(11, 957) = 3.13$, $p < .001$, for the effect of serial position in the three retrieval groups combined, because there was no effect of retrieval difficulty, $F < 1$]. The critical difference between the four conditions emerged only for the first item, in that the math group has better access to the first item from L2 than did the three retrieval groups combined, $t(118) = 2.51$, $p < .05$. These results replicate those of the previous studies, and they indicate that retrieving L1 between L2 and L3, regardless of how long ago L1 was learned, disrupts access to the L2 context during the test.

Experiment 4

In Experiment 1 and 2, we varied the number of letter cues provided during the retrieval of L1 in an attempt to vary retrieval difficulty, whereas in Experiment 3, the number of letter cues was fixed, but the delay separating L1 and L2 was varied. Thus, in all three experiments, L1 retrieval involved cued recall. In the present experiment, we aimed to replicate the retrieval difficulty manipulation of Experiment 3 (e.g., the delay/recency of L1 was varied), except that we asked the participants to retrieve L1 without the help of any external cues. The provision of letter cues in the previous experiments might have drawn too much attention to the cue itself, which could have dampened the impact of contextual information during retrieval. Thus, we switched to free

[1] The selected delay intervals were dictated by the constraints on the availability of the lab space and the participants at those time points.

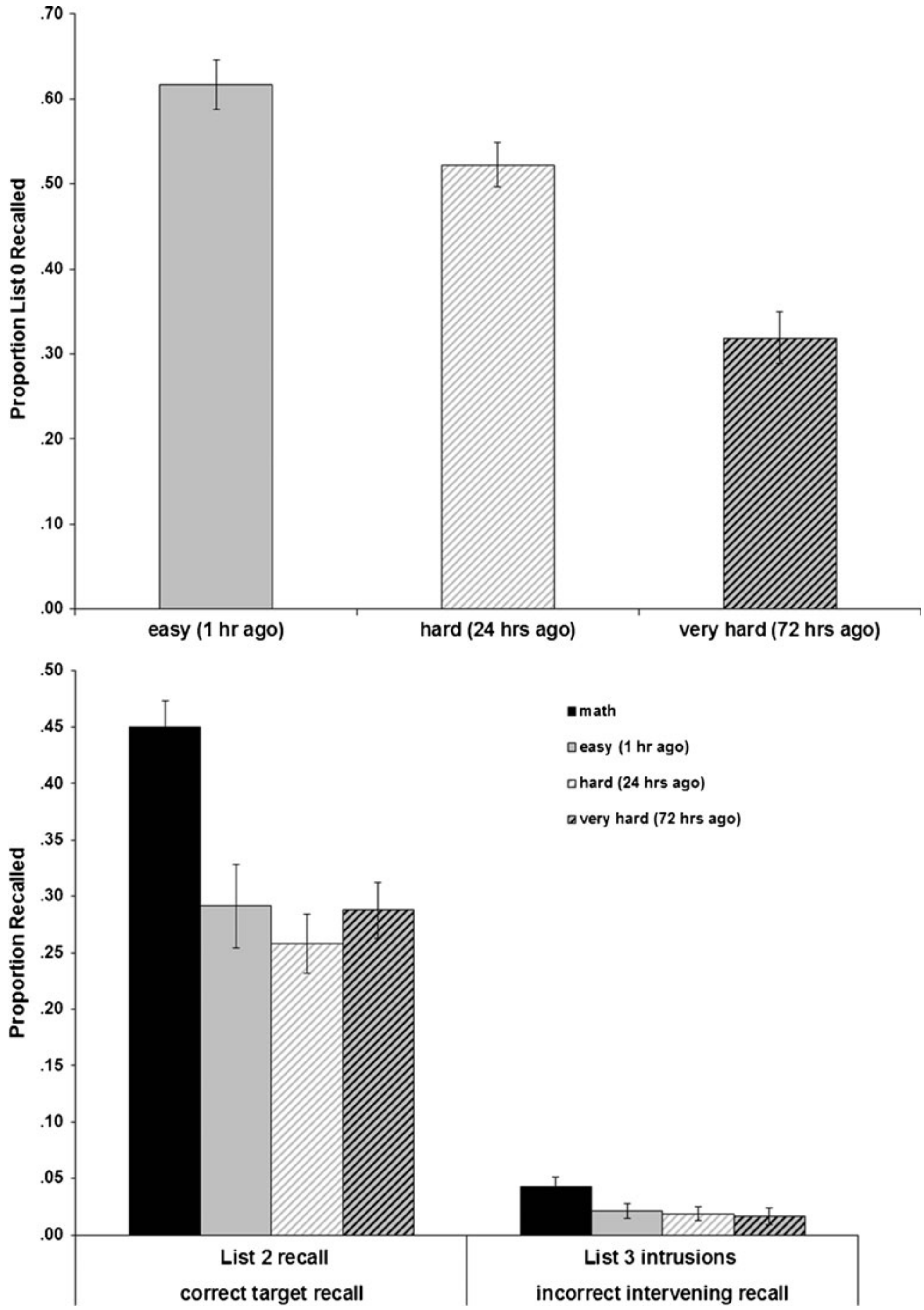

Fig. 6 Experiment 3: (Top) Proportions of List 1 recalled as a function of retrieval difficulty. (Bottom) Proportions of List 2 recalled (left panel) and proportions of List 3 intrusions (right panel) as a function of experimental condition. The errors bars represent ± *SE*s of the means

recall of L1 in the present experiment, because free recall relies on the context cues substantially more than does cued recall (e.g., Shiffrin, Ratcliff, & Clark, 1990).

Method

Participants A group of 120 UNCG undergraduates participated for course credit. None of them had participated in the previous experiments.

Materials and design The materials and design were the same as in Experiment 3.

Procedure The procedure was similar to that of Experiment 3, with a few changes. The first major change was that immediately after the third presentation of L1, participants were told to free recall L1. We used an immediate recall test as a way of offsetting potential floor recall after long delays (see, e.g., Roediger & Karpicke, 2006), since we were not planning on providing any letter cues for L1 items later on. The second

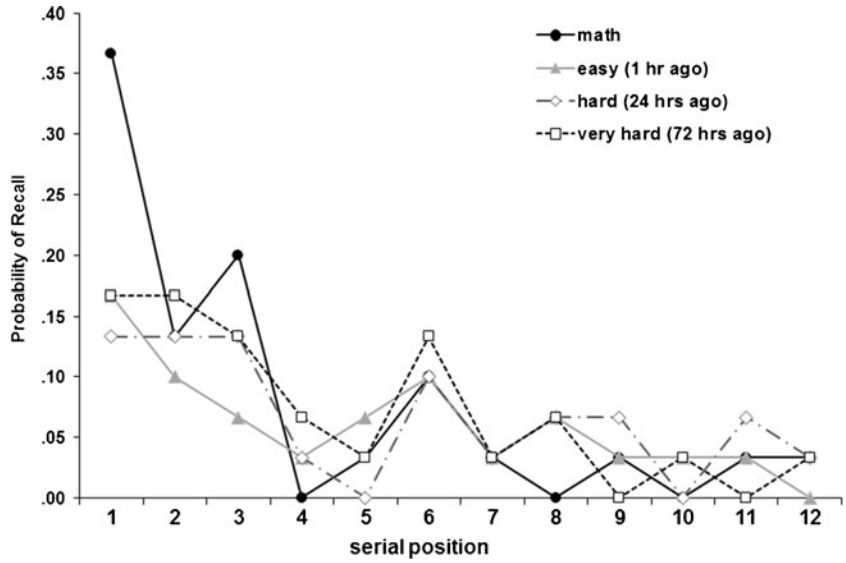

Fig. 7 First-item recall probabilities from List 2 as a function of experimental condition and serial position in Experiment 3

major change from the Experiment 3 procedure was that between the study of L2 and L3, participants engaged in free recall of L1 as opposed to cued recall. They were given a blank sheet of paper and were asked to retrieve the L1 words that had been studied in the previous session. All other procedural details were similar to those of Experiment 3.

Results and discussion

List 1 recall The results are displayed in Fig. 8 (top panel). A one-way ANOVA on the proportions of L1 words recalled by each group (easy retrieval vs. hard retrieval vs. very hard retrieval) revealed a significant effect of group, $F(2, 87) = 58.44$, $MSE = .024$, $p < .001$. The easy-retrieval group had higher recall than did the hard-retrieval group, $t(58) = 2.38$, $p < .05$, which in turn had higher recall than did the very-hard-retrieval group, $t(58) = 2.39$, $p < .05$.

List 2 recall Fig. 8 (bottom left panel) summarizes the L2 recall findings. A one-way ANOVA on the proportions of L2 recall by group (math vs. easy retrieval vs. hard retrieval vs. very hard retrieval) revealed significant differences, $F(3, 116) = 3.13$, $MSE = .021$, $p < .05$. When the math group was excluded from the analyses, the effect of group was no longer evident, $F < 1$. As in the previous experiments, the math group had higher List 2 recall than did the easy-retrieval group [$t(58) = 3.31$, $p < .01$], the hard-retrieval group [$t(58) = 2.15$, $p < .05$], and the very-hard-retrieval group [$t(58) = 2.15$, $p < .05$].

List 3 intrusions The results concerning L3 items erroneously intruding during L2 recall also replicated those of the previous experiments. We found a significant effect of group, $F(3, 116) = 3.15$, $MSE = .002$, $p < .05$ (see Fig. 8, bottom right panel): There were more L3 intrusions in the math group than in the easy-retrieval group [$t(58) = 2.84$, $p < .01$] or the hard-retrieval group [$t(58) = 2.16$, $p < .05$], and marginally more than in the very-hard-retrieval group [$t(58) = 1.66$, $p < .10$]. If the math group was excluded from the analyses, the three retrieval conditions did not differ from each other, $F < 1$.

First-response probabilities from List 2 The probabilities of first recall from L2 are shown in Fig. 9. The math group tended to initiate recall from the first item on L2 more often than from any other items on the list [$F(11, 319) = 4.78$, $p < .001$, for the effect of serial position in the math group]. This effect was substantially reduced in the retrieval conditions [$F(11, 957) = 1.72$, $p = .06$, for the effect of serial position in the three retrieval groups combined]. The only significant differences between the four groups emerged in the recall of the item from the first serial position, $F(3, 116) = 3.64$, $p < .05$. When the math group was excluded from the analyses, the three retrieval conditions did not differ from each other, $F < 1$, and hence we collapsed the data across the difficulty factor. Consistent with the prior experiments, the math group had much better access to the first item from L2 than did the retrieval groups, $t(118) = 3.31$, $p < .001$.

General discussion

Across four experiments and several dependent measures, we observed a consistent pattern of results: Retrieving L1 between encoding of L2 and L3 led to forgetting of L2 at the time of final test, relative to solving math between the lists. The magnitude of forgetting was unaffected by the degree of L1 retrieval difficulty, regardless of whether difficulty was manipulated by the number of letter cues or by varying the

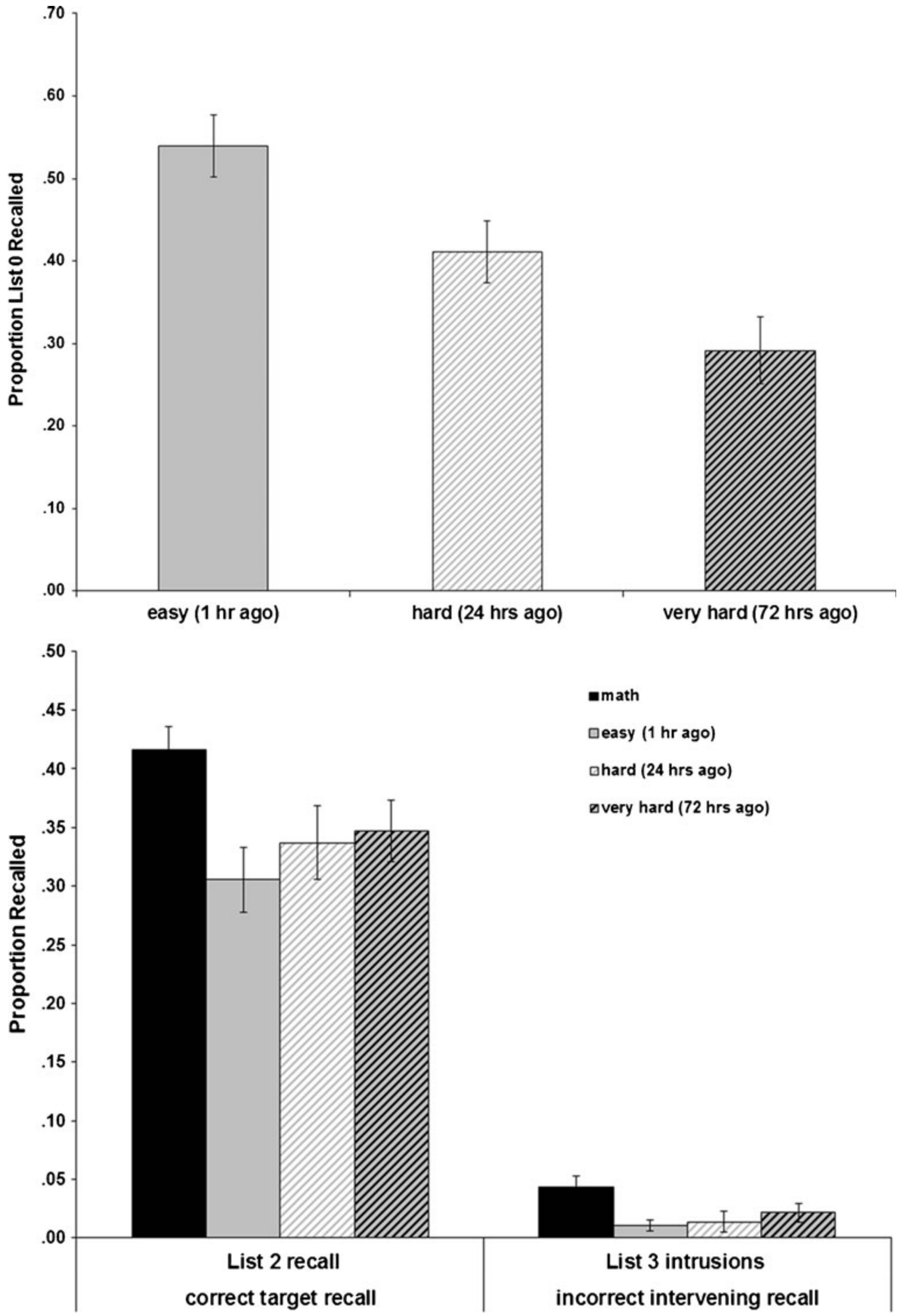

Fig. 8 Experiment 4: (Top) Proportions of List 1 recalled as a function of retrieval difficulty. (Bottom) Proportions of List 2 recalled (left panel) and proportions of List 3 intrusions (right panel) as a function of experimental condition. The errors bars represent ± *SE*s of the means

delay between L1 and L2. The results were also unaffected by whether we used confusable items across the lists (e.g., Exp. 1) or nonconfusable items (e.g., Exps. 2–4). The retrieval of L1 not only had negative consequences for memory, by impairing L2 recall, but it also had a positive effect, by isolating L2 from interference, as evidenced by reduced L3 intrusions. Other results in the literature have also demonstrated enhanced list differentiation as a result of retrieval trials (e.g., Chan & McDermott, 2007; Pastötter, Schicker, Niedernhuber, & Bäuml, 2011; Szpunar, McDermott, & Roediger, 2008), although these studies typically involved immediate testing after each study list, as opposed to retrieval of a prior list, as in the present experiments. Retrieval of the prior list caused participants to lose access to the beginning of the L2 context, which was more accessible in the math groups than in the retrieval groups, as evidenced by

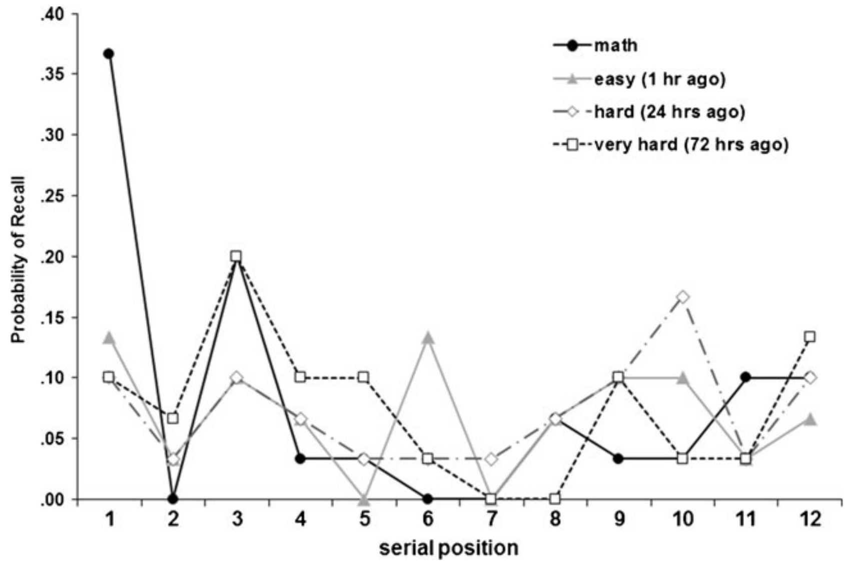

Fig. 9 First-item recall probabilities from List 2 as a function of experimental condition and serial position in Experiment 4

the first-recall functions. Other manipulations that are thought to disrupt internal context, such as directed forgetting instructions (e.g., Sahakyan & Kelley, 2002), also show reduced first-response probabilities in the forget cue condition (Lehman & Malmberg, 2009; Spillers & Unsworth, 2011).

Overall, the results are more broadly consistent with a context-change interpretation of prior-list retrieval than with any of the alternative interpretations, based on item interference, item inhibition, or retrieval difficulty. According to the context-change interpretation proposed by Jang and Huber (2008), context drifts from one list to the next list in a relatively gradual fashion in the math condition (see, e.g., Estes, 1955; Mensink & Raaijmakers, 1988). At the time of the final test, participants can either rely on the recent L3 context to access items from L2, or they can attempt to reinstate the L2 context directly. Because the contexts of L2 and L3 are similar, either list context can be a good retrieval cue for L2 items. The downside is that the contextual similarity between the two lists could make list discrimination more difficult, and hence we observed more L3 errors during L2 recall in the math group. In contrast, the act of retrieving L1 between the presentations of the other two lists changed the context between L2 and L3. If participants relied on the L3 context during the test to retrieve L2 items, it would be a poor retrieval cue for L2 items. Likewise, directly reinstating the L2 context might have been more difficult because of the intervening L1 retrieval trial, and thus recall of L2 items suffered. On the upside, the lists were better differentiated in the retrieval conditions than in the math condition, as evidenced by reduced L3 intrusions during L2 recall. Overall, reduced L2 recall and reduced intrusions from L3 demonstrate the two effects that are typically associated with external context change—the forgetting effect and the reduced interference effect (for a review, see Smith & Vela, 2001). Indeed, when the study and the test environments mismatch, recall suffers relative to when the environments are the same (e.g., Godden & Baddeley, 1975). Also, when the target list and the interfering list are learned in two different environments as opposed to the same environment, interference is substantially reduced (e.g., Bilodeau & Schlosberg, 1951; Dallett & Wilcox, 1968; Greenspoon & Ranyard, 1957; Sahakyan, 2010). Interestingly, even repetitions of the same list, when carried out in two different environments, improve recall substantially, as compared to when they occur in the same environment (e.g., Smith, Glenberg, & Bjork, 1978).

The results of the present experiments are not consistent with item interference, item inhibition, or retrieval difficulty interpretations. For example, according to the item interference/blocking account, if more L1 items are retrieved, they should block access to L2 items to a greater extent during the final test than if fewer L1 items are retrieved. Likewise, the retrieval difficulty explanation would predict that the more difficult it is to retrieve L1 items, the more forgetting of L2 we should observe. The inhibitory account would predict that in an easy retrieval condition there may be less need to invoke inhibition of L2, and therefore less forgetting of L2 items should occur on a later test; in addition, L3 intrusions during L2 recall should be the same across the math group and the retrieval groups. Our results are not consistent with any of these accounts, suggesting that the observed effects are unlikely to reflect item-specific effects but are more likely an indication of global context-change effects.

Despite the overall compatibility with a context-change interpretation, the results are inconsistent with the TCM model, which is a context-change model that assumes that successful retrieval of items updates the current state of the context. If more L1 items are retrieved, context should change more between L2 and L3, and hence we should see more L2 forgetting. Our results, however, indicated that regardless of the number of L1 items retrieved, L2 recall suffered to the same extent. These results are more compatible with an interpretation that assumes that retrieval of a prior list creates context change, but that context change is not driven by the actual items recalled (e.g., Mensink & Raaijmakers, 1988). Instead, context change may simply be driven by the process of thinking back to the previously learned information. Whether successful retrieval of previously learned information is needed to create a context change remains unclear. It could be that retrieving a single item from L1 is sufficient to create a maximum degree of context change, and any additional items that are retrieved from L1 do not make any difference—a sort of "one-shot hypothesis of context retrieval" (cf. Malmberg & Shiffrin, 2005). It could also be that the first retrieved item has a bigger impact on context change than do the subsequently retrieved items. If this is the case, our results could be compatible with a TCM model, if the impacts of retrieved items on context updating were differentially weighted. Finally, it is possible that even if no item were retrieved from L1, a context change would still be instigated by the act of thinking back in time. These are important questions that remain to be addressed by future research.

References

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

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

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1063–1087. doi:10.1037/0278-7393.20.5.1063

Anderson, M. C., & Levy, B. J. (2010). On the relationship between interference and inhibition in cognition. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 107–132). New York, NY: Psychology Press.

Bilodeau, I. M., & Schlosberg, H. (1951). Similarity in stimulating conditions as a variable in retroactive inhibition. *Journal of Experimental Psychology, 41,* 199–204.

Bjork, R. A., & Richardson-Klavehn, A. (1989). On the puzzling relationship between environmental context and human memory. In C. Izawa (Ed.), *Current issues in cognitive processes: The Tulane Flowerree Symposium on Cognition* (pp. 313–344). Hillsdale, NJ: Erlbaum.

Chan, J. C. K., & McDermott, K. B. (2007). The testing effect in recognition memory: A dual process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 431–437. doi:10.1037/0278-7393.33.2.431

Dallett, K., & Wilcox, S. G. (1968). Contextual stimuli and proactive inhibition. *Journal of Experimental Psychology, 78,* 475–480.

Delaney, P. F., Sahakyan, L., Kelley, C. M., & Zimmerman, C. A. (2010). Remembering to forget: The amnesic effect of daydreaming. *Psychological Science, 21,* 1036–1042.

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

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology, 66,* 325–331. doi:10.1111/j.2044-8295.1975.tb01468.x

Greenspoon, J., & Ranyard, R. (1957). Stimulus conditions and retroactive inhibition. *Journal of Experimental Psychology, 53,* 55–59.

Hintzman, D. L. (1988). Judgments of frequency of recognition memory in a multiple-trace memory model. *Psychological Review, 95,* 528–551. doi:10.1037/0033-295X.95.4.528

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

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

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

Lehman, M., & Malmberg, K. J. (2009). A global theory of remembering and forgetting from multiple lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 970–988.

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.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review, 39,* 353–370.

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

Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K.-H. T. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 287–297. doi:10.1037/a0021801

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93–134. doi:10.1037/0033-295X.88.2.93

Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17,* 249–255. doi:10.1111/j.1467-9280.2006.01693.x

Sahakyan, L. (2010). Environmental context change affects memory for performed actions. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 63*, 425–433. doi:10.1080/17470210903414365

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.

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

Shiffrin, R. M., Ratcliff, R., & Clark, S. E. (1990). List-strength effect: II. Theoretical mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 179–195. doi:10.1037/0278-7393.16.2.179

Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition, 6,* 342–353. doi:10.3758/BF03197465

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

Spillers, G. J., & Unsworth, N. (2011). Are the costs of directed forgetting due to failures of sampling or recovery? Exploring the dynamics of recall in list-method directed forgetting. *Memory & Cognition, 39*, 403–411. doi:10.3758/s13421-010-0038-z

Storm, B. C. (2010). Retrieval-induced forgetting and the resolution of competition. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 89–105). New York, NY: Psychology Press.

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

Copyright of Memory & Cognition is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.