Journal of Memory and Language 106 (2019) 18–28

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

List-method directed forgetting after prolonged retention interval: Further challenges to contemporary accounts

Magdalena Abel*, Karl-Heinz T. Bäuml

Department of Experimental Psychology, Regensburg University, Germany

ARTICLE INFO

Keywords:
List-method directed forgetting
Delay
Mental context change
Selective rehearsal

ABSTRACT

Numerous studies on list-method directed forgetting (LMDF) have shown that people can voluntarily forget information when cued to do so. But the cognitive mechanism(s) behind this form of forgetting are still subject to debate. The present study focused on two explanations of LMDF: selective rehearsal and mental context change. Experiment 1 addressed the context-change account by comparing the persistence of LMDF with that of context-dependent forgetting. Results showed that LMDF, but not context-dependent forgetting, was lasting, which is inconsistent with the context-change account. Experiments 2 and 3 addressed the selective-rehearsal account by examining whether persistence of LMDF depends on the status of (intentionally vs. incidentally) encoded items and the type of distractor activity (demanding vs. undemanding) between study and test. Results showed that LMDF was lasting for both intentionally and incidentally studied items but was absent after an undemanding distractor task, which disagrees with the selective-rehearsal account. The present findings challenge both the context-change and the selective-rehearsal account as well as a dual-mechanisms view, which assumes a role of both types of mechanisms in LMDF.

Introduction

Can people voluntarily forget stored memory contents that are no longer relevant? Several lines of research indicate that this is indeed the case and that humans can, to a certain degree, exert control over the contents of their minds (for a review, see Anderson & Hanslmayr, 2014). One task developed to examine voluntary forgetting of outdated information in the lab is list-method directed forgetting (LMDF; Bjork, 1970). In this task, subjects study two lists of items and, after study of the first list, are asked to either remember the list for a later test or try to forget the list, pretending that it would not be relevant for the later test and can be forgotten. After study of the second list, recall of first-list items is tested, irrespective of whether a remember cue or a forget cue was provided for the list. The typical finding is that subjects who received the forget cue show reduced first-list recall compared to subjects who received the remember cue. This directed forgetting of first-list items is a robust finding that arises over a wide range of study materials and experimental settings (for reviews, see Bäuml, Pastötter, & Hanslmayr, 2010; MacLeod, 1998; Sahakyan, Delaney, Foster, & Abushanab, 2013).

Theoretical accounts of LMDF

Several accounts have been proposed to explain directed forgetting of first-list items. The oldest account is selective rehearsal (Bjork, 1970). This account assumes that, when subjects are cued to remember list 1, they try to maintain the list in memory by engaging in mental rehearsal during list-2 study. In contrast, when subjects are cued to forget the first list, there is no such rehearsal, which may reduce later recall of list-1 items. The retrieval-inhibition account assumes that the forget cue activates an inhibitory control process that impairs access to list 1 and thereby reduces recall of the first-list items (Geiselman, Bjork, & Fishman, 1983). Finally, the context-change account assumes that subjects deliberately change their mental context in response to a forget cue. As a consequence, the mental context at test no longer matches the context during study of list 2, which again impairs recall of the first-list items (Sahakyan & Kelley, 2002).

Although the selective-rehearsal account is consistent with the basic finding of list-1 forgetting, it is challenged by a number of results in the LMDF literature. For instance, different rehearsal activities should become evident on recognition tests, but studies show that directed forgetting effects are usually absent in item recognition (e.g., Basden, Basden, & Gargano, 1993; Geiselman et al., 1983). Similarly, directed forgetting has not only been found to arise for intentionally studied, but

* Corresponding author at: Department of Experimental Psychology, Regensburg University, 93040 Regensburg, Germany.
E-mail address: magdalena.abel@ur.de (M. Abel).

https://doi.org/10.1016/j.jml.2019.02.002
Received 9 February 2018; Received in revised form 25 January 2019
Available online 19 February 2019

also for incidentally encoded material, which, according to the selective-rehearsal hypothesis, should not be subject to strategic rehearsal processes and should not show directed forgetting (Geiselman et al., 1983; Sahakyan & Delaney, 2010). Therefore, the selective-rehearsal account is typically dismissed in current explanations of LMDF (e.g., Sahakyan et al., 2013; but see Delaney, Nghiem, & Waldum, 2009).

In contrast to selective rehearsal, both the retrieval-inhibition account and the context-change account can explain a wide range of LMDF findings that go beyond the basic finding of list-1 forgetting. Single findings have even been interpreted as specific evidence for the retrieval-inhibition account (e.g., Conway, Harries, Noyes, Racsmany, & Frankish, 2000; Hanslmayr et al., 2012) or the context-change account (e.g., Lehman & Malmberg, 2011; Sahakyan, Waldum, Benjamin, & Bickett, 2009), but it is not always clear whether they can really provide such specific support (see also Sahakyan et al., 2013). Despite this theoretical ambiguity, recently the context-change account has often been preferred over the inhibition account, mostly because of the possible link of the account to the general literature on context effects in memory, which makes the account conceptually richer and leads to a couple of novel predictions on LMDF (for an example, see Sahakyan & Goodmon, 2010).

One line of evidence in favor of the context-change account comes from studies reporting parallel effects between LMDF as induced by a forget cue and context-dependent forgetting as induced by an imagination task. In imagination tasks, participants, for instance, are asked after study of list 1 to mentally walk through the house of their parents or describe what they would like to do if they were invisible (e.g., Jonker, Seli, & MacLeod, 2013; Pastötter & Bäuml, 2007; Sahakyan & Kelley, 2002). Contrasting the effects of such imagination tasks to the effect of a forget cue, the two resulting forms of forgetting were found to be accompanied by similar serial-position curves for list-1 items (Sahakyan & Foster, 2009) and to emerge in the presence of list-2 encoding but not in its absence (Pastötter & Bäuml, 2007). Similarly, both forms of forgetting were found to be reduced if, immediately before the test, the study context of list 1 was reinstated by context reinstatement procedures (Sahakyan & Kelley, 2002) or, at test, a subset of the list 1 items was provided as part-list cues to facilitate recall of the other items (Bäuml & Samenieh, 2012). Such parallels support the context-change account, because the account assumes that, at least over a wide range of settings, LMDF and context-dependent forgetting should be empirically equivalent. Indeed, until recently, no study has been reported pointing to behavioral dissociations between the two forms of forgetting and identifying factors that influence the one form of forgetting (e.g., context-dependent forgetting), but not the other (e.g., directed forgetting).

Using prolonged retention intervals to examine the context-change account of LMDF

Examining LMDF and context-dependent forgetting after a prolonged retention interval, Abel and Bäuml (2017) recently added a new twist to the theoretical debate on LMDF. In two experiments, these researchers found context-dependent forgetting as induced by imagination tasks to be present after short delays (30 s or 3 min), but to be eliminated after longer delays of 20 min or even 24 h, which conceptually replicated prior work on the transiency of context-dependent forgetting (Divis & Benjamin, 2014; see also General Discussion). In contrast, they found directed forgetting to be not only present after the two short delays, but to persist in similar size after the two prolonged delays. The finding of relatively transient context-dependent forgetting is consistent with the context-change account. Indeed, following prior work on mental context drift, which assumes that a gradual mismatch between the mental context present at encoding and at test builds up with delay (e.g., Estes, 1955; McGeoch, 1932), experimentally induced context changes, like those triggered by imagination tasks, should become less relevant with delay and no longer reduce recall relative to a no-context change condition (see also Divis & Benjamin, 2014; Sahakyan & Kelley, 2002). In contrast, the finding of lasting LMDF challenges the context-change account of LMDF by indicating that LMDF and context-dependent forgetting differ with delay and may thus be mediated by different mechanisms.

To conclude from Abel and Bäuml's (2017) report that the context-change account, or some other noninhibitory account of LMDF, necessarily provides an inadequate explanation of LMDF might be premature, however, for at least two reasons. The first reason is empirical in nature. To examine the persistence of LMDF, Abel and Bäuml (2017) used a procedure that differed in multiple aspects from more typical LMDF tasks. In particular, they used two study cycles instead of one during learning to avoid floor effects after longer delay; they employed initial-letter cued recall instead of free recall at test; and they asked subjects to recall a subset of the studied items only to get relatively pure measures of the forgetting effects (for a discussion, see Abel & Bäuml, 2017). Arguably, these deviations from more typical tasks may have affected the results. For instance, additional study cycles may add further context features to the items and initial-letter cued recall may rely less on context information than a free-recall task, both of which could attenuate context-dependent forgetting and thus be at the heart of the reported difference between LMDF and context-dependent forgetting. It is therefore important to examine whether the findings reported by Abel and Bäuml (2017) are restricted to the employed experimental procedure, or rather generalize to more typical LMDF tasks. Finding such generalization would challenge the context-change account.

The second reason for why caution is warranted with regard to conclusions from Abel and Bäuml's (2017) prior work arises on the basis of theoretical considerations. If the finding of persistent directed forgetting but eliminated context-dependent forgetting after delay generalized to more typical LMDF tasks, then this would argue against the context-change account as a full account of LMDF, but it would not necessarily argue against a noninhibitory explanation of LMDF. Indeed, such results would well agree with a two-factor account, which claims that (transient) context-dependent forgetting mediates LMDF for short delay, but selective rehearsal mediates persistent LMDF for prolonged delay. In fact, even though selective rehearsal has previously been rejected as an explanation of (short-delay) directed forgetting (see Bäuml et al., 2010; Sahakyan et al., 2013), following MacLeod, Dodd, Sheard, Wilson, and Bibi (2003), it could play a critical role when retention intervals are prolonged. In such case, participants in the remember condition may rehearse list 1 and list 2 items during a longer retention interval, whereas participants in the forget condition may selectively rehearse the list 2 items, anticipating that only those items will be tested later. Because selective rehearsal of list 2 during the retention interval should operate in response to the forget cue, but not in response to an imagination task, such two-factor account would also provide an explanation for why LMDF is persistent, but context-dependent forgetting is transient.

The present study

There were two major goals with the present study. The first goal was to examine whether LMDF is still found to be lasting but context-dependent forgetting to be transient when using a more typical LMDF task to examine the persistence of the forgetting. In the first step, we therefore examined if Abel and Bäuml's (2017) finding of persistent directed forgetting, but eliminated context-dependent forgetting, generalizes to a more typical LMDF procedure. To address the issue, Experiment 1 employed both a short 30-s and a prolonged 20-min retention interval between study and test. In particular, it applied just one study cycle for each list and used a final free recall test to assess memory for (all) list items. Arguably, such procedure may increase the role of context at test and thus provide a much stronger test of whether context-dependent forgetting is transient or lasting. Because the prior finding of persistent directed forgetting, but transient context-dependent forgetting, is inconsistent with the context-change account of

LMDF, a generalization of Abel and Bäuml's (2017) finding to the present task would challenge the account.

The second goal of the present study was to provide a more detailed examination of the proposal that selective rehearsal mediates persistent directed forgetting. In the second step, we therefore examined in two further experiments the potential role of selective rehearsal for the maintenance of directed forgetting across delay. Experiment 2 followed Geiselman et al.'s (1983) hallmark study, and applied intermixed learn and judge words for study that were tested after a 30-s or a 20-min retention interval. If selective rehearsal is what maintains directed forgetting across delay, the effect should mainly persist for intentionally memorized learn words, but not, or to a lower extent, for incidentally encoded judge words that one would not expect to be subject to selective rehearsal. Experiment 3 manipulated the type of distractor activity that subjects were asked to engage in during a 20-min delay. On the basis of the selective-rehearsal hypothesis, a less demanding distractor task should leave more room to engage in rehearsal activities than a more demanding distractor task, and thus lead to a more pronounced directed forgetting effect.

Together, Experiments 1–3 will indicate whether (i) LMDF persists across prolonged retention interval also with a more typical LMDF procedure and is different from context-dependent forgetting in this respect, and (ii) the status of (intentionally vs. incidentally) encoded items and the type of distractor activity (demanding vs. undemanding) influence the persistence of LMDF. On the basis of these results, the study will provide critical information on whether LMDF is mediated by either mental context change alone or a combination of mental context change and selective rehearsal.

Experiment 1

Experiment 1 employed a typical LMDF procedure to examine whether LMDF and context-dependent forgetting are transient or lasting. Subjects studied two word lists and received either a remember cue or a forget cue after the first list, or they were asked to engage in an imagination task between lists. Recall of first-list items was tested after a retention interval of 30 s or a retention interval of 20 min. In contrast to the previous study by Abel and Bäuml (2017), Experiment 1 applied only one study cycle for each list, and recall was assessed by means of a free recall test for all first-list items, a procedure that may increase the role of contextual factors at test. In addition, we included two different imagination tasks in this study to see if findings on context-dependent forgetting are task-specific or generalize across different tasks. The context-change account predicts that directed forgetting as a proposed context-change effect should mimic context-dependent forgetting as induced by imagination tasks, and that both forms of forgetting should be transient and not persist across the prolonged retention interval.

Method

Participants. We determined sample sizes on the basis of prior work on LMDF and delay from our lab (Abel & Bäuml, 2013, 2017). 256 students at Regensburg University were recruited for the experiment (32 participants per condition). Mean age was 22.2 years (range 18–40 years). All subjects were fluent in German.

Material. Item material consisted of two lists of items. The sequence of lists was counterbalanced across participants. Each list comprised 16 unrelated German nouns. All item materials that were applied in the present experiments as well as all data are available on the Open Science Framework (https://osf.io/d5ky6/). Moreover, all experiments reported in this manuscript were implemented using the software E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). The software was run on standard desktop computers with the operating system Windows 7 (Microsoft, Redmond, WA). All data were analyzed using IBM SPSS Statistics for Windows, Version 25.0 (IBM Corp., Armonk, NY).

Design. The experiment applied a 4 × 2 between-subjects design with the factors of INSTRUCTION (remember cue, forget cue, two different imagination tasks) and DELAY (short delay, long delay). After list-1 study but before list-2 study, subjects were either asked to remember the first list, to forget the first list, or to engage in an imagination task. Recall was tested after a short delay of 30 s or after a longer delay of 20 min filled with distractor tasks.

Procedure. The procedure was largely identical to the procedure applied by Abel and Bäuml (2017), apart from the number of study cycles during encoding and the type of final test. In the beginning, subjects were asked to memorize two lists of items. List items were presented one at a time, in random order, and for 4 s each centrally on a computer screen. After list-1 study, subjects received either the forget or the remember cue for the first list, or they were asked to engage in an imagination task. In the forget condition, we simulated a software crash to make the cover story more plausible that a wrong list had been presented (see Abel & Bäuml, 2013; Barnier et al., 2007). Subjects were asked to forget the first list and to focus on the list coming up next instead. Simulating the software crash and presenting the cover story took roughly 60 s. In contrast, in the remember condition, subjects counted backwards in steps of two for 60 s before being asked to additionally memorize list 2. In the mental context change conditions, subjects were asked to engage in one of two different imagination tasks. In the one condition, subjects were instructed to imagine walking through their childhood homes and were asked to draw a sketch of the house (e.g., Sahakyan & Kelley, 2002); in the other condition, subjects were asked to close their eyes and think back to an international vacation (e.g., Delaney, Sahakyan, Kelley, & Zimmerman, 2010). After 60 s, subjects in all conditions additionally memorized the second item list. Duration of between-list instructions was held constant between conditions and always filled an interval of roughly 60 s.

In the short delay conditions, subjects were asked to count backwards in steps of two for 30 s before completing a final memory test of both item lists. In the long delay conditions, the same test was completed after 20 min. Subjects were asked to work on unrelated cognitive tests to fill this 20-min delay interval. These tests included the connect-the-numbers test (Oswald & Roth, 1987; for roughly 6 min), the d2 test of attention (Brickenkamp & Zillmer, 1998; for roughly 6 min), and standard progressive matrices (Raven, 2000; for roughly 8 min). The final memory test used a free-recall format. Subjects were asked to write down all list items they could remember and were given 60 s for each list. List 1 was always tested first, then list 2 was tested in the same way. Before testing started, subjects in the forget condition were debriefed and were asked to try to recall as many of the first-list items as possible, irrespective of the previous instruction to forget the list.

Results

Results for list 1. Fig. 1 shows mean recall rates of list 1 items as a function of INSTRUCTION (remember cue, forget cue, two imagination tasks) and DELAY (short delay, long delay). A 4 × 2 ANOVA showed significant main effects, indicating that recall was generally affected by INSTRUCTION, , and DELAY, $F(1, 248) = 10.86$, $MSE = 401.33$, $p = .001$, $\eta^2 = 0.04$. The two main effects were accompanied by a significant interaction, however, $F(3, 248) = 3.80$, $MSE = 401.33$, $p = .011$, $\eta^2 = 0.04$, suggesting that the delay interval influenced the effect of instruction. To follow up, we separately evaluated the influence of delay on directed forgetting and context-dependent forgetting in the two imagination tasks by means of further 2 × 2 ANOVAs.

Concerning directed forgetting, a 2 × 2 ANOVA showed significant main effects of INSTRUCTION (remember cue, forget cue), $F(1, 124) = 36.73$, $MSE = 373.44$, $p < .001$, $\eta^2 = 0.23$, and DELAY (short delay, long delay), $F(1, 124) = 23.62$, $MSE = 373.44$, $p < .001$, $\eta^2 = 0.16$. Recall was impaired after the forget cue compared to the remember cue (31.6% vs. 52.3%) and was lower after the 20-min delay compared to the 30-s delay (33.7% vs. 50.3%). Critically, the ANOVA did not reveal a significant

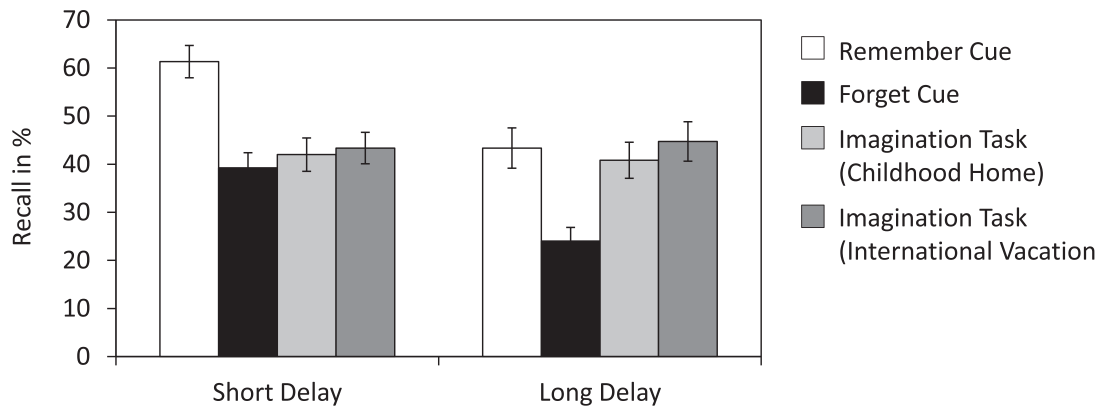

Fig. 1. Results of Experiment 1. Mean list-1 recall as a function of delay (short delay, long delay) and instruction (remember cue, forget cue, and two different imagination tasks). Error bars represent ± 1 standard error.

interaction of the two factors, $F(1, 124) < 1.00$, $MSE = 373.44$, $p = .690$, $\eta^2 = 0.001$, suggesting that the effect of the forget instruction did not depend on delay. To examine the nonsignificant interaction effect more closely, we followed Masson (2011) and applied the Bayesian information criterion (BIC) to compute posterior probabilities for the null and alternative hypotheses being correct given the observed data (D). The resulting posterior probabilities were $P_{BIC}(H_0|D) = 0.912$ and $P_{BIC}(H_1|D) = 0.088$, which can be interpreted as positive evidence in favor of the null hypothesis (see Masson, 2011; Raftery, 1995).

Concerning context-dependent forgetting, we ran two further 2 × 2 ANOVAs that contrasted the remember cue condition separately with each of the two imagination tasks. For the first imagination task (thinking back to one's childhood home), the ANOVA revealed not only significant main effects of INSTRUCTION (remember cue, imagination task), $F(1, 124) = 8.70$, $MSE = 440.14$, $p = .004$, $\eta^2 = 0.07$, and DELAY (short delay, long delay), $F(1, 124) = 6.66$, $MSE = 440.14$, $p = .011$, $\eta^2 = 0.05$, but also a significant interaction between the two factors, $F(1, 124) = 5.13$, $MSE = 440.14$, $p = .025$, $\eta^2 = 0.04$, suggesting that the mnemonic consequences of engaging in the imagination task depended on delay. Follow-up t-tests showed that context-dependent forgetting was observed after the short delay (42.0% vs. 61.3%), $t(62) = 4.00$, $p < .001$, $d = 1.00$, but not after the long delay (40.8% vs. 43.4%), $t(62) < 1.00$, $p = .653$, $d = 0.11$. For the critical nonsignificant difference in recall after the long delay, the posterior probabilities were $P_{BIC}(H_0|D) = 0.878$ and $P_{BIC}(H_1|D) = 0.122$, demonstrating positive evidence in favor of the null hypothesis (see Masson, 2011; Raftery, 1995).

For the second imagination task (thinking back to an international vacation), the ANOVA revealed parallel findings, i.e., significant main effects of INSTRUCTION (remember cue, imagination task), $F(1, 124) = 4.89$, $MSE = 450.75$, $p = .029$, $\eta^2 = 0.04$, and DELAY (short delay, long delay), $F(1, 124) = 4.89$, $MSE = 450.75$, $p = .029$, $\eta^2 = 0.04$, as well as a significant interaction effect, $F(1, 124) = 6.64$, $MSE = 450.75$, $p = .011$, $\eta^2 = 0.05$. Again, context-dependent forgetting was present after the short delay (43.4% vs. 61.3%), $t(62) = 3.83$, $p < .001$, $d = 0.96$, but not after the long delay (44.7% vs. 43.4%), $t(62) < 1.00$, $p = .816$, $d = 0.06$. Concerning the null effect after the long delay, the posterior probabilities of $P_{BIC}(H_0|D) = 0.886$ and $P_{BIC}(H_1|D) = 0.114$ again provided positive evidence in favor of the null hypothesis (Masson, 2011; Raftery, 1995).

Results for list 2. As research on LMDF has shown, the forget cue can not only impair recall of list 1 but can also improve recall of list 2 (e.g., Bjork, 1989; MacLeod, 1998). The beneficial effect of the forget cue on list-2 recall, however, arises mainly when at test list 2 is recalled first and is often absent when list 1 is recalled first (for the results of a recent meta analysis, see Pastötter, Kliegl, & Bäuml, 2012). Because the focus of the present study was on list-1 forgetting, subjects in all experiments of the present study were asked to recall list 1 items first and list 2 items second. Consistently, a 4 × 2 ANOVA showed a significant main effect of DELAY, $F(1, 248) = 26.29$, $MSE = 488.05$, $p < .001$, $\eta^2 = 0.10$, but no significant main effect of INSTRUCTION, $F(3, 248) < 1.00$, $MSE = 488.05$, $p = .670$, $\eta^2 = 0.01$ ($P_{BIC}(H_0|D) = 0.999$), and no significant interaction, $F(3, 248) = 1.74$, $MSE = 488.05$, $p = .160$, $\eta^2 = 0.02$ ($P_{BIC}(H_0|D) = 0.997$), indicating that list-2 recall did not differ between instruction conditions (see Table 1 for mean list-2 recall in the single conditions).

Discussion

The results showed both directed forgetting and context-dependent forgetting of list-1 items after a short delay of 30 s, with a similar size of effects in the two forms of forgetting, which is consistent with prior work (e.g., Pastötter & Bäuml, 2007; Sahakyan & Kelley, 2002). In contrast, after a delay of 20 min, the results showed directed forgetting, whereas context-dependent forgetting as induced by the two imagination tasks was absent, indicating that LMDF and context-dependent forgetting can differ after longer delay. This finding generalizes Abel and Bäuml's (2017) original finding to a more typical LMDF task and additionally shows that it holds across different imagination tasks. The observed dissociation of directed forgetting and context-dependent forgetting thus seems to be robust and to not much depend on procedural detail, like number of study cycles or recall format at test. The finding of no context-dependent forgetting after prolonged retention interval is consistent with the theoretical view that mental context change creates relatively transient context-dependent forgetting (e.g., Divis and Benjamin, 2004; Mensink & Raaijmakers, 1988; Sahakyan & Kelley, 2002). In contrast, the finding of lasting LMDF disagrees with such context view and therefore challenges the context-change account

Table 1
Mean list-2 recall (plus standard deviations) in Experiment 1 as a function of delay (short delay, long delay) and instruction (remember cue, forget cue, imagination tasks).

	Remember cue	Forget cue	Imagination task (childhood home)	Imagination task (international vacation)
Short delay	52.7 (22.2)	49.6 (16.9)	43.2 (23.8)	44.3 (21.8)
Long delay	32.0 (22.6)	29.9 (19.2)	32.6 (26.6)	38.7 (22.2)

of LMDF as a full explanation of LMDF.

As shown by the reported ANOVA results, the recall impairment induced by the prolonged retention interval was numerically similar in the remember and forget conditions, which at first glance may suggest similar forgetting rates in the two conditions. However, given that, after the short delay, recall levels were higher in the remember than forget conditions, proportional forgetting rates - the reduction in recall after delay relative to the initial recall level - may be the more adequate measure of forgetting rates across the delay (e.g., Wixted, 2004). Using such proportion measure, the results indeed indicate that the forgetting rate was slightly larger in the forget condition (38.9%) than the remember condition (29.2%), suggesting that the forget cue may have accelerated time-dependent forgetting. In contrast to the remember and forget conditions, there was no time-dependent forgetting at all in the two imagination conditions (for a similar result, see Abel & Bäuml, 2017). Why time-dependent forgetting after imagination tasks may be reduced, or even be absent, must remain unclear with the present experiment. However, the finding may suggest that imagination tasks entail a rather radical update of mental context, such that any further, delay-induced contextual change may be less disruptive, or not disruptive at all. Employing a larger number of delay intervals than was used in the present experiment, future work may examine time-dependent forgetting in the remember, forget, and imagination conditions in more detail.

While the results of Experiment 1 challenge the context-change account of LMDF, they do not rule out other noninhibitory explanations of LMDF. For instance, the results are basically consistent with a noninhibitory two-factor account of LMDF, which attributes short-delay LMDF to context change but attributes long-delay LMDF to selective rehearsal. According to such proposal, selective rehearsal might play a critical role during longer delay: after a remember cue, subjects may rehearse both list-1 and list-2 items during the delay, whereas after a forget cue, they may engage in selective rehearsal of list-2 items only, which could maintain the forgetting present after short delay across longer retention interval. Such view would also explain why LMDF, but not context-dependent forgetting, is persistent, because selective rehearsal of list 2 during longer delay should occur mainly in response to a forget instruction but not in response to an imagination task. Experiment 2 and Experiment 3 examine the adequacy of this two-factor account in describing persistent LMDF in more detail.

Experiment 2

Experiment 2 examined if LMDF for intentionally studied material is different from LMDF for incidentally studied material after longer delay. Doing so, Experiment 2 followed Geiselman et al.'s (1983) hallmark study on the role of selective rehearsal after short delay and applied the same rationale to an investigation of the role of selective rehearsal after prolonged retention interval. In this classic study, Geiselman et al. employed a typical LMDF task, but subjects were instructed that, within each list, only some items, so-called learn words, should be memorized and would be tested later, whereas other intermixed items, so-called judge items, should be rated for their pleasantness and would not be tested later. Results showed that even though, as a whole, learn items were remembered better than judge items on a final (short-delay) test, intact and comparable directed forgetting arose for both item types. Because selective rehearsal should affect the intentionally studied learn items, but not the merely to-be-rated judge items, it was concluded that selective rehearsal did not mediate (short-delay) directed forgetting.

In Experiment 2, we applied the same method to examine whether selective rehearsal might become more important with delay, when subjects are asked to maintain mental representations of studied items across prolonged retention intervals and there is increased opportunity to engage in selective rehearsal. If this were the case and differential rehearsal activities for to-be-remembered and to-be-forgotten items mediated persistent directed forgetting, then the forgetting effect should persist primarily for intentionally studied learn items, but to a much lesser degree, if at all, for incidentally encoded judge items. Experiment 2 was run to test this noninhibitory explanation of long-delay directed forgetting.

Method

Participants. A new sample of 128 students with a mean age of 21.8 years (range: 18–33 years) was recruited for the experiment (32 subjects per condition). All subjects were fluent in German.

Material. New item material was compiled, again consisting of two lists of items. Each list comprised 24 unrelated German nouns, which were randomly divided into two sets of 12 items. Sequence of lists was counterbalanced across participants. Similarly, each set within each list served equally often as learn and judge items across participants.

Design. The experiment had a 2 × 2 × 2 mixed-factorial design. The first factor of INSTRUCTION was again manipulated between subjects, and, after list-1 study, one half of the participants were asked to remember the first list, whereas the other half were asked to forget the first list. The second factor of WORD TYPE was manipulated within subjects. Following Geiselman et al. (1983), each list was to equal parts composed of to-be-memorized learn items and to-be-rated judge items. The third factor of DELAY was again manipulated between subjects, with half of all subjects completing a recall test after a short delay of 30 s or a longer delay of 20 min.

Procedure. The experimental procedure was closely modeled after that reported by Geiselman et al. (1983). Subjects were initially informed that they would be studying a list of to-be-memorized words, but that these learn words would be intermixed with to-be-judged words that would not be tested later. Subjects were asked to try to memorize the learn words and to rate the judge words for their pleasantness. All 24 list items were presented for 5 s each, centrally on a computer screen. Learn and judge words were presented in alternating sequence, and each single item was presented below a word-type label to make sure that subjects knew whether to learn or to judge the item (e.g., LEARN - elephant, JUDGE - plum, etc.). We generated an alternating, but otherwise random sequence of learn and judge items and used this list plus its mirrored version to control for sequence effects, such that half of all subjects started with a learn item, whereas the other half started with a judge item. For learn trials, subjects were simply instructed to try to memorize the presented word as best as they could. For judge trials, subjects were asked to rate the presented word with regard to its pleasantness on a scale from 1–5 (1 = not pleasant at all; 5 = very pleasant); the scale was presented below the words on the screen, and subjects were instructed to press a key corresponding to the perceived pleasantness of the word.

When list-1 study was completed, subjects in the remember cue condition were simply asked to try to remember the just memorized learn words for a later test and to study an additional second list of learn words, again intermixed with judge words. In contrast, in the forget cue condition, subjects were informed that the just presented list was presented for practice only, to familiarize them with the alternating sequence of learn and judge words. Subjects were asked to try to forget the just presented learn words, because they would not be tested later. Instead, subjects were asked to focus on the second list, pretending that this would be the only list relevant for the upcoming memory test. The second list of intermixed learn and judge items was then presented in the same fashion as the first list.

After list-2 study, subjects in the short delay conditions were asked to count backwards in steps of 2 for 30 s before completing the final test. In the long delay conditions, all subjects were asked to work on the same unrelated cognitive tasks as in Experiment 1 to fill a 20-min delay interval. Afterwards, recall for both lists was tested. Again, subjects in the forget condition were debriefed prior to the test phase, to ensure that they would try to recall as many list-1 items as possible,

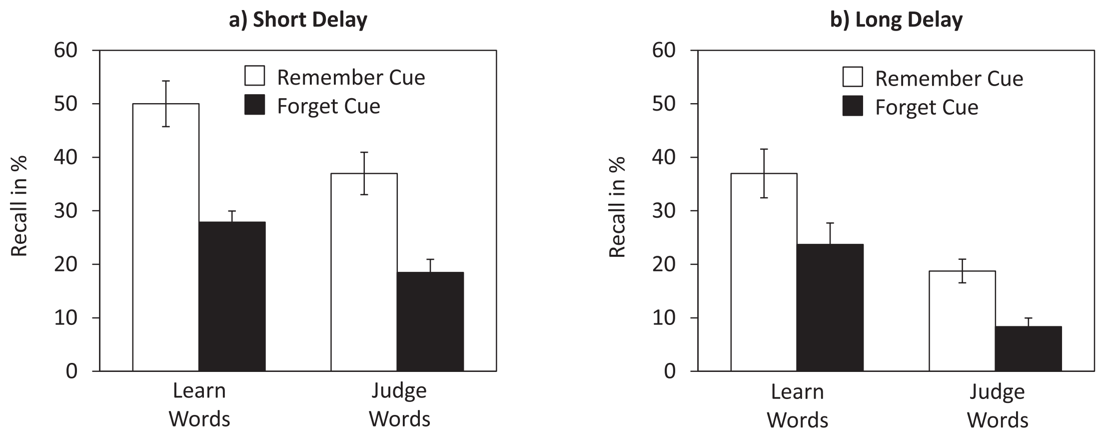

Fig. 2. Results of Experiment 2. Mean list-1 recall as a function of instruction (remember cue, forget cue) and word type (learn words, judge words). Panel (a) shows results in the short delay condition, panel (b) shows results in the long delay condition. Error bars represent ± 1 standard error.

irrespective of the previously presented instruction to forget the list. Moreover, all subjects were debriefed with regard to the presumably irrelevant judge items, and were asked to try to recall them, too. The test was a free recall test, and subjects were given 2 min per list to write down all items they could remember. List 1 was always tested first, then list 2 was tested in the same way. Following Geiselman et al., for each list, subjects received a prepared recall sheet with separate columns for learn and judge words. All subjects were asked to try to remember both types of words, and to organize them into the correct columns. When the final test was completed, subjects were thanked for their participation and debriefed.

Results

Results for list 1. In the first step, we analyzed all recalled items, irrespective of whether they were correctly classified as learn or judge items or not. Fig. 2a and b show mean list-1 recall in the short- and long-delay conditions as a function of INSTRUCTION (remember cue, forget cue) and WORD TYPE (learn, judge). A 2 × 2 × 2 ANOVA showed a significant main effect of INSTRUCTION, $F(1, 124) = 20.94$, $MSE = 0.04$, $p < .001$, $\eta^2 = 0.14$, reflecting better recall in the remember cue (32.2%) than the forget cue condition (21.4%). There was also a significant main effect of WORD TYPE, $F(1, 124) = 72.54$, $MSE = 0.03$, $p < .001$, $\eta^2 = 0.37$, with higher recall of learn items (35.8%) than judge items (17.7%), and a significant main effect of DELAY, $F(1, 124) = 20.41$, $MSE = 0.04$, $p < .001$, $\eta^2 = 0.14$, with higher recall after the short delay (33.3%) than the long delay (21.9%). Critically, the ANOVA revealed no significant interactions between any of the factors, all $Fs(1, 124) < 1.00$, all $MSEs \geqslant 0.03$, all $ps \geqslant .419$, all $\eta^2 s \leqslant 0.01$ (all $P_{BIC}(H_0|D) \geqslant 0.894$), suggesting that the effect of the forget cue compared to the remember cue on recall was not only unaffected by delay, but was also not affected by whether items were intentionally studied or incidentally encoded by means of pleasantness judgments.

In the second step, we considered only items for analysis that were correctly classified as learn or judge items. Classification performance and discriminability of learn and judge words was almost perfect: in the remember-cue condition, 98% of all recalled learn items and 98% of all recalled judge items were correctly classified; in the forget-cue condition, 97% of the recalled learn items and 98% of the recalled judge items were assigned to the correct label. Unsurprisingly, when considering only those items that were correctly classified, the ANOVA results reported above replicated – i.e., significant main effects of delay, instruction, and word type, all $Fs(1, 124) \geqslant 18.65$, all $ps < .001$, all $\eta^2 s \geqslant 0.13$, but no significant interaction effects, all $Fs(1, 124) < 1.00$, all $ps \geqslant .500$, all $\eta^2 s \leqslant 0.004$ – indicating that the finding of similarly pronounced directed forgetting for learn and judge words after both short and long delay did not arise because subjects confused the two word types and engaged in selective rehearsal for both of them.

When asked after the final test, 16 out of the 128 subjects reported not to have completely trusted our initial instruction that to-be-judged words would not be tested later. Arguably, these subjects may have engaged in other rehearsal activities than the rest of the sample. Excluding these 16 subjects from the analysis, another ANOVA however revealed the same results as reported above, with significant main effects of delay, instruction, and word type, all $Fs(1, 108) \geqslant 11.32$, all $ps \leqslant .001$, all $\eta^2 s \geqslant 0.10$, but no significant interaction effects, all $Fs(1, 108) < 1.00$, all $ps \geqslant .522$, all $\eta^2 s \leqslant 0.004$.

Results for list 2. Mean list-2 recall is shown separately for all conditions in Table 2. A 2 × 2 × 2 ANOVA on recalled list-2 items revealed no significant main effect of INSTRUCTION, $F(1, 124) = 2.80$, $MSE = 0.04$, $p = .097$, $\eta^2 = 0.02$, suggesting that recall did not generally differ between the remember and forget cue conditions (29.4% vs. 33.5%). The main effect of WORD TYPE was significant, $F(1, 124) = 159.01$, $MSE = 0.03$, $p < .001$, $\eta^2 = 0.56$, with higher recall of learn items (45.8%) than judge items (17.2%), and there was also a significant main effect of DELAY, $F(1, 124) = 22.07$, $MSE = 0.04$, $p < .001$, $\eta^2 = 0.15$, with higher recall after the short delay (37.2%) than the long delay (25.7%). The ANOVA also showed a significant interaction between INSTRUCTION and WORD TYPE, $F(1, 124) = 7.45$, $MSE = 0.03$, $p = .007$, $\eta^2 = 0.06$ - larger enhancement for learn words than for judge words - but no further significant two-way interactions, all $Fs(1, 124) \leqslant 1.19$, all $ps \geqslant .278$, all $\eta^2 \leqslant 0.01$, and no significant three-way interaction, $F(1, 124) = 3.49$, $MSE = 0.03$, $p = .064$, $\eta^2 = 0.03$. Like for list 1 recall, the results remained the same when only those recalled items were considered for analysis that were correctly classified as learn or judge items. These findings replicate the result of Experiment 1 that delay did not influence the effect of the forget cue. Again, because list 2 was always tested last in this experiment and because prior work shows that this test sequence can modulate list-2 findings (see Pastötter et al., 2012), the list-2 results should be interpreted with caution.

Table 2
Mean list-2 recall (plus standard deviations) in Experiment 2 as a function of delay (short delay, long delay), instruction (remember cue, forget cue), and word type (learn words, judge words).

	Remember cue		Forget cue	
	Learn word	Judge word	Learn word	Judge word
Short delay	42.5 (21.9)	25.3 (12.1)	59.6 (21.0)	21.6 (10.6)
Long delay	38.8 (27.7)	11.2 (10.5)	42.2 (26.4)	10.7 (10.4)

Discussion

Experiment 2 confirms the finding of Experiment 1 that directed forgetting of intentionally studied items can persist across delay. Most importantly, the experiment shows that a hallmark finding on LMDF after short delay previously reported by Geiselman et al. (1983) can be replicated and generalized to LMDF after prolonged retention interval. After both short and long delay between study and test, directed forgetting was not only evident for intentionally studied items, but was also present for incidentally encoded items that were judged for their pleasantness. This finding challenges the proposal that directed forgetting after longer delay is mediated by selective rehearsal. If this were the case, selective rehearsal should predominantly maintain directed forgetting for intentionally studied items, but not, or to a much lesser degree, for incidentally encoded judge items. The results of Experiment 2 thus provide first evidence that selective rehearsal may not underlie persistent directed forgetting. Experiment 3 put the selective-rehearsal hypothesis to a further test.

Experiment 3

Rehearsal activities and their effects on later recall should depend on whether there is more or less mental capacity left to engage in rehearsal during a retention interval. In both the previous experiments reported by Abel and Bäuml (2017) and the present Experiments 1 and 2, rather demanding distractor tasks in the form of unrelated cognitive tests were used to fill the 20-min delay intervals. These tightly controlled and experimenter-parsed distractor tasks were employed to gain some control over subjects' mental activity during the retention intervals, and, in particular, to keep them from constantly rehearsing the studied lists. While these tasks may not have prevented subjects from engaging in at least some rehearsal activities, such activities may become more pronounced if the distractor task was less demanding and subjects were therefore given more room for rehearsal activities. In fact, if selective rehearsal was responsible for maintaining directed forgetting across prolonged retention interval, the amount of directed forgetting should depend on the distractor activity placed between encoding and test, and should be more pronounced with a less demanding distractor task. The goal of Experiment 3 was to test this prediction, by comparing directed forgetting across two different types of distractor activities – a relatively demanding and a relatively undemanding distractor task.

Method

Participants. A new sample of 128 students with a mean age of 22.2 years (range: 19–32 years) was recruited for the experiment (32 subjects per condition). All subjects were fluent in German.

Material. We used the same item material as in Experiment 1, with two lists of 16 unrelated German nouns each. Sequence of lists was again counterbalanced across participants.

Design. The experiment had a 2 × 2 between-subjects design. The first factor of INSTRUCTION was again manipulated by asking one half of the participants to remember the first list for a later test, whereas the other half were asked to try to forget the first list and to focus on the second list instead. The second factor of DISTRACTOR ACTIVITY was also manipulated between subjects. During the 20-min retention interval, one half of the subjects were asked to engage in the same demanding cognitive tests as subjects in Experiments 1 and 2 (as well as in Abel & Bäuml, 2017), whereas the other half were asked to engage in a rather undemanding vigilance task instead. Recall was assessed after 20 min in all conditions (i.e., no short-delay condition was included in Experiment 3).

Procedure. The experimental procedure was largely identical to the procedure applied for the long-delay remember and forget cue conditions in Experiment 1, the only exception being that the type of distractor activity was manipulated during the 20-min retention interval. Whereas half of the subjects engaged in the same cognitive tasks between study and test as in Experiment 1 (i.e., the connect-the-numbers test, the d2 test of attention, and standard progressive matrices), the other half of subjects were asked to engage in a vigilance task for the same amount of time, which was designed to be rather undemanding and was modeled after the vigilance task developed by Zimmermann and Fimm (2007). During this task, subjects saw a constant stream of geometric figures of varying shapes and colors on a computer screen. They were asked to press a certain computer key when the same figure was presented twice in a row. On vigilance tasks, the critical event requiring a reaction occurs only very infrequently. In the present experiment, subjects saw a stream of 480 figures in total, but a repetition of the same figure occurred on only 30 of these trials. To ensure similar spacing of critical trials across the retention interval, unbeknownst to participants, the task was divided into three blocks of 160 trials, with each block containing 10 critical trials. Performance on the first two blocks was excellent, as is shown by very high mean rates of correct reactions on critical trials (96.1% and 96.4% correct) and very low mean numbers of errors (0.22 and 0.001). On the third block, however, performance dropped, with a mean of 69.9% correct reactions on critical trials and a mean number of 0.001 errors. When asked about their impression of the vigilance task, subjects predominantly reported that they thought it was very easy, but also pretty boring. In contrast, in the demanding distractor condition, the combination of several unrelated cognitive tasks was never perceived as boring, but as a rather diverse and challenging distractor activity.

After the 20-min distractor interval, memory was again assessed by means of a free recall test. As in Experiment 1, subjects had 1 min per list to write down all list items they could remember. List 1 was always tested before list 2. Upon test completion, subjects were thanked for their participation and debriefed.

Results

Results for list 1. Fig. 3 shows mean list 1 recall as a function of INSTRUCTION (remember cue, forget cue) and DISTRACTOR ACTIVITY (demanding distractor, undemanding distractor). A 2 × 2 ANOVA showed a significant main effect of INSTRUCTION, $F(1, 124) = 10.18$, $MSE = 476.17$, $p = .002$, $\eta^2 = 0.08$, reflecting better recall in the remember cue (41.5%) than the forget cue condition (29.2%). There was no significant main effect of DISTRACTOR ACTIVITY, $F(1, 124) < 1.0$, but a significant interaction effect, $F(1, 124) = 6.93$, $MSE = 476.17$, $p = .010$, $\eta^2 = 0.05$, suggesting that type of distractor activity affected the difference in recall levels between the forget and remember cue conditions. Follow-up tests showed that intact directed forgetting was

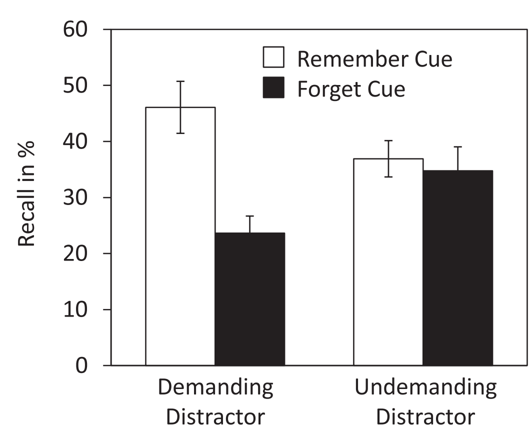

Fig. 3. Results of Experiment 3. Mean list-1 recall as a function of instruction (remember cue, forget cue) and distractor activity (demanding distractor, undemanding distractor). Error bars represent ± 1 standard error.

Table 3
Mean list-2 recall (plus standard deviations) in Experiment 3 as a function of instruction (remember cue, forget cue), and distractor activity (demanding distractor, undemanding distractor).

	Remember cue	Forget cue
Demanding distractor	40.6 (27.2)	36.1 (18.2)
Undemanding distractor	32.8 (19.1)	39.8 (22.6)

present after the demanding distractor task (23.6% vs. 46.1%), $t(62) = 4.05$, $p < .001$, $d = 1.01$, but was absent after the undemanding task (34.8% vs. 36.9%), $t(62) < 1.0$, $d = 0.10$ ($P_{BIC}(H_0|D) = 0.880$). Recall was lower after the demanding than the undemanding distractor task in the forget cue condition (23.6% vs. 34.8%), $t(62) = 2.12$, $p = .038$, $d = 0.53$, but recall did not differ statistically between distractor tasks in the remember cue condition (46.1% vs. 36.9%), $t(62) = 1.62$, $p = .110$, $d = 0.41$.

Results for list 2. Mean list-2 recall is shown separately for the single conditions in Table 3. A 2 × 2 ANOVA on recalled list-2 items showed no significant main or interaction effects, all $Fs(1, 124) \leqslant 2.18$, all $ps \geqslant .142$, indicating that list-2 recall did not differ between instruction conditions.

Discussion

The results in the condition with a demanding distractor activity between study and test replicate results of Experiments 1 and 2 by showing that directed forgetting can persist across a 20-min delay interval. The results in the condition with an undemanding distractor activity show a different picture. They indicate that easy (and potentially even boring) distractor tasks can eliminate the forgetting effect across the same delay. This finding is at odds with predictions derived from the selective-rehearsal hypothesis. If selective rehearsal mediated the maintenance of directed forgetting across longer retention intervals, if anything, the size of the forgetting effect should increase if subjects have more room to engage in selective rehearsal across delay (see also MacLeod et al., 2003). The undemanding vigilance task applied in the present experiment was definitely easy enough to free up subjects' mental capacities to enable them to engage in higher amounts (or more elaborate forms) of rehearsal than in the demanding distractor task condition. And indeed, type of distractor activity did influence recall in the present experiment, though in the opposite direction as predicted on the basis of the selective-rehearsal hypothesis.

Although Experiment 3 was conducted with the clear goal of testing the selective rehearsal hypothesis, it may also be possible to consider the results from a context-change perspective. First, the finding that, with a demanding distractor, LMDF is present across longer delay contrasts with the finding of Experiment 1 that, with the same distractor tasks, context-dependent forgetting does not survive longer delay and thus is inconsistent with the context-change account (see above). Second, on the basis of the context-change account, there may be reason to argue that directed forgetting should have been present with the undemanding distractor task used in Experiment 3. Such expectation arises if one assumes that the more demanding distractor, which consisted of multiple tasks, created additional contextual change during the delay, but the undemanding distractor, which consisted of one long period of the same activity, induced hardly any contextual change. Indeed, following the view that the effects of experimentally induced context changes should be reduced after delay when the delay includes contextual drift (e.g., Divis & Benjamin, 2014; Mensink & Raaijmakers, 1988; Sahakyan & Kelley, 2002), a reduction in LMDF might be expected mainly with the multiple demanding distractor tasks (i.e., when the delay included additional contextual change), and less with the single undemanding distractor task (i.e., when the delay included hardly any further contextual change). Clearly, this is not what the results of the experiment show. It should be stressed, however, that this reasoning is post hoc, and that it is so far unclear how exactly different types and numbers of distractor tasks affect magnitude of context change across prolonged retention intervals. One road to pursue in future empirical work may be to examine the influence of demanding versus undemanding distractor activity on the effects of imagination tasks. Potentially, if the above assumptions are correct, the results of such work may find the effects of imagination tasks to differ from those of the forget cue in *both* distractor conditions and show persistent context-dependent forgetting with the undemanding distractor only.

Aside from the question of which cognitive mechanism(s) mediate(s) LMDF, the present finding that a relatively undemanding distractor task can eliminate directed forgetting mimics other recent results by Schlichting and Bäuml (2017), who reported that wakeful resting can eliminate LMDF. Schlichting and Bäuml used rather short retention intervals of 72 s between study and test, but, similar to the present study, manipulated the distractor activity during these intervals. When a more demanding distractor activity was used (counting backwards or solving simple math equations), they found intact directed forgetting. Yet, when subjects were instead asked to relax and to either listen to music or contemplate picture material as a distractor activity, directed forgetting was eliminated. Like the results of the present Experiment 3, these findings indicate that type of distractor activity during delay can influence LMDF (for a related finding in item-method directed forgetting, see Lee, 2012).

General discussion

Using typical LMDF tasks, the present experiments provide a number of important findings on the persistence of LMDF. Experiment 1 shows a dissociation between LMDF and context-dependent forgetting after prolonged retention interval; LMDF is present after 20 min, but context-dependent forgetting is not. Experiment 2 shows that LMDF is lasting for both intentionally and incidentally encoded words, suggesting that the expectation of a future test does not influence persistence. Finally, Experiment 3 shows that the type of distractor activity between study and test can modulate LMDF after prolonged delay; while directed forgetting is present after a demanding distractor activity, the forgetting can be absent if the distractor activity is undemanding. As a whole, these results provide evidence that LMDF is not a transient phenomenon, but can last for quite a while.

Theoretical implications for accounts of LMDF

The results of Experiment 1 challenge the context-change account of LMDF as a full explanation of the forgetting. The context-change account attributes directed forgetting and context-dependent forgetting to the same context-change mechanism. On the basis of this view and the assumption that contextual drift occurs naturally with the passage of time (e.g., Estes, 1955; Mensink & Raaijmakers, 1988), both forms of forgetting should be transient and dissipate with increasing delay (e.g., Divis & Benjamin, 2014; Sahakyan & Kelley, 2002). The observation of transient context-dependent forgetting as induced by the two imagination tasks employed in Experiment 1 is consistent with this prediction, but the finding of lasting directed forgetting is not, indicating that the forgetting as induced by imagination tasks and the forgetting as induced by a forget cue are mediated by different mechanisms. Doing so, the results also reject all accounts that attribute LMDF and context-dependent forgetting to the same mechanism, including the view that both are mediated by inhibition (see Anderson, 2005).

The results of Experiment 1 challenge the context-change account, but they are principally consistent with an alternative two-factor explanation of LMDF. Indeed, while LMDF after shorter delay may be mediated by (transient) mental context change, during longer delay selective rehearsal may take place. After a remember cue participants may rehearse both list-1 and list-2 items during the delay, whereas after a forget cue participants may selectively rehearse the list-2 items, expecting that only these items will be tested later (see MacLeod et al., 2003). Such differential rehearsal during longer delay would lead to higher recall of rehearsed to-be-remembered than nonrehearsed to-be-forgotten items and thus to persistent directed forgetting. Because selective rehearsal would be expected to occur in response to a forget cue, but not after an imagination task, such view could explain why LMDF is lasting, but context-dependent forgetting is not.

While such a two-factor account can provide an explanation of the results of Experiment 1, it cannot explain the results of Experiments 2 and 3. Regarding Experiment 2, the account would predict lasting LMDF for intentionally studied learn items, but transient LMDF for incidentally encoded judge items, because only intentionally studied material should be subject to selective rehearsal. In contrast, the results of Experiment 2 show persistent LMDF for both item types with no difference in the size of the forgetting effect between the two item types, thus challenging the selective-rehearsal account. Regarding Experiment 3, the account would predict stronger LMDF after an undemanding than a demanding distractor task, because an undemanding activity should leave room for subjects to engage in higher amounts or more elaborate forms of selective rehearsal. Yet, the results of Experiment 3 show persistent LMDF after a demanding, but not after an undemanding distractor activity, which is inconsistent with the account.

The present experiments were designed to test the adequacy of two prominent noninhibitory accounts of LMDF, but they were not designed to test the inhibition account. Indeed, because to date no assumption on degree of persistence has been included in the inhibition account, the account is silent on whether LMDF should be transient or lasting. Although the present results thus cannot serve as an evaluation of the inhibition account, the results of Experiments 1 and 2 may suggest to impose the restriction on the account that the inhibitory effect is lasting. With such restriction, the inhibition account leads to the expectation that, in general, LMDF results observed after short delay also show up after longer delay, at least with delays as they were used in the present Experiments 1 and 2 and in Abel and Bäuml (2017). While future work may examine the validity of this expectation in more detail, the results of Experiment 3, showing no LMDF with an undemanding distractor task, may be regarded a first challenge to this expectation. Indeed, there is no obvious reason why an inhibition account would predict such finding and the first-order prediction of the account would rather be that the forgetting arises regardless of distractor activity. Because the results of Experiment 3 are also not easily reconciled with the context view of LMDF (see above), they pose a challenge to all three contemporary accounts of LMDF. This challenge may drive future - empirical and theoretical - work on LMDF.

Relation to prior work on LMDF and context-dependent forgetting

The present finding of transient context-dependent forgetting when mental context change is induced by an imagination task parallels the results of a recent study by Divis and Benjamin (2014). These authors examined how semantic generation of extralist items between the study of successively presented lists affects later recall of the first list items. On the basis of the view that semantic generation can induce mental context change (e.g., Jang & Huber, 2008; Pastötter, Schicker, Niedernhuber, & Bäuml, 2011) and the assumption that context change induces transient forgetting, recall impairment of list-1 items was expected after a short delay of 1 min but not after a prolonged delay of 15 min. The results were in line with the expectation, supporting the view that induced mental context change creates relatively transient forgetting.

The present finding of lasting LMDF is consistent with other reports of lasting LMDF in the literature, which, however, come with methodological shortcomings. For instance, both Basden and Basden (1998) and MacLeod et al. (2003) reported evidence that the forgetting of list-1 items can still be present after a 20-min retention interval. Unfortunately, none of the studies included a remember condition and all participants were asked to forget list 1 and remember list 2. Directed forgetting was calculated by subtracting recall of list-1 items from recall of list-2 items, a method that endorses possible effects of the forget cue on both list 1 and list 2 and therefore cannot answer the question to which extent the cue influenced list-1 recall after delay. In contrast, the present finding of lasting LMDF may appear inconsistent with other results from the literature. For instance, in her dissertation work, Liu (2001) applied regular word lists and reported a numerical reduction of LMDF from a 3-min to a 22-min delay, which may indicate that LMDF decreases with delay. The reduction was nonsignificant, however, so that the results may not be in conflict with the present findings. Shapiro, Lindsey, and Krishnan (2006) reported transient LMDF in an applied study, which used a notebook shopping scenario and examined the role of LMDF for advertising. Arguably, however, this study may have induced mental context change before providing a forget cue, which may be the reason for why LMDF was transient (for details, see Abel & Bäuml, 2017).

The present indication of a behavioral dissociation of directed forgetting and context-dependent forgetting is in line with the results of a few previous studies, which point to a neural dissociation of the two forms of forgetting. Bäuml, Hanslmayr, Pastötter, and Klimesch (2008) and Pastötter, Bäuml, and Hanslmayr (2008) measured electroencephalograms during list-2 encoding after subjects had received a forget cue or engaged in an imagination task. Analyzing subjects' oscillatory brain activity, they found different patterns of brain oscillations in a certain frequency band (11–13 Hz) in response to the two forgetting manipulations. While the forgetting of list-1 items in LMDF was reflected by a sustained decrease in phase synchronization between electrode sites, no evidence for such a decrease arose when the forgetting was induced by mental context change (see also Hanslmayr et al., 2012). The results from the present and these previous studies thus indicate that directed forgetting and context-dependent forgetting can be differentiated, both neurally and behaviorally.

Relation to prior computational work on LMDF

Formalizing Sahakyan and Kelley's (2002) context-change account, Lehman and Malmberg (2009) provided a computational account to explain LMDF as it occurs in a 3-list variant of the original (2-list) task. In this variant, before studying lists 1 and 2, subjects study an additional list 0, and after study of list 0 and list 1 are then cued to either remember or forget the two previous lists. This task was introduced by Lehman and Malmberg to reduce possible unwanted advantages of list 1 over list 2, which, according to the authors, may occur in the 2-list task (see Lehman & Malmberg, 2009, pp. 972–973).

Arguably, this computational account may be used to deduce predictions on how prolonged retention interval influences LMDF, to see whether the account is consistent with the present results. Application of the account to the present findings is problematic, however, because, as is indicated by Lehman and Malmberg's (2009) recall and recognition results, LMDF in the 3-list task differs fundamentally from LMDF in the 2-list task. For instance, when using the 3-list task, directed forgetting of list-1 items is present in inclusion as well as exclusion item recognition, whereas it is absent in these tests when using the 2-list task (e.g., Geiselman et al., 1983; Pastötter, Kliegl, & Bäuml, 2016); serial position curves, for both list-1 and list-2 items and in both recall and item recognition, look quite different in the 3-list than the 2-list task (e.g., Pastötter et al., 2012, 2016; Sahakyan & Foster, 2009); finally, in the 3-list task, list recency characterizes recall of lists 1 and 2 - with higher recall of list 2 than list 1 -, whereas the opposite pattern arises in the 2-list task (e.g., Geiselman et al., 1983). According to Lehman and Malmberg, the absence of list recency in the 2-list task precludes application of the account to this task and "requires a special model" (p. 974), indicating that the account cannot easily be applied to the present results.

Still, Lehman and Malmberg's (2009) account includes an explanation of list-1 forgetting, which, if it generalized from the 3-list to the 2-list task, would offer an interesting speculation on why LMDF and context-dependent forgetting differ with delay. The account rests critically on the assumption that the forget cue induces context change, but the account also suggests that context change by itself is insufficient for LMDF and that difficulty of context reinstatement at the time of test is a second critical factor to induce forgetting. On the basis of this view, forgetting of first-list items after delay would arise in the forget condition but not in the imagination conditions, *if* both context change and reinstatement difficulty were induced by the forget cue, but context change without reinstatement difficulty was induced by imagination tasks.[1] In such case, however, context change as induced by the forget cue and context change as induced by imagination tasks would not be the same, which again provided a challenge to the mental context-change view of LMDF. Future work, both empirical and computational, may address this speculation.

From list-1 forgetting to list-2 enhancement

As was the case in most LMDF work in recent years, the focus of the present experiments was on the effects of the forget cue on list-1 recall, with a reduced interest in possible effects of the forget cue on list-2 recall. This focus is reflected by the fact that, in each single experiment, subjects were asked at test to recall list 1 first and list 2 second. Although, in general, the forget cue can create two effects, impairing recall of list 1 and improving recall of list 2 (e.g., Bjork, 1989; MacLeod, 1998), asking subjects at test to recall list-1 items first often reduces, or even eliminates, any enhancement effects for list-2 items (see Pastötter et al., 2012). The results of the present experiments, particularly Experiments 1 and 3, show a similar pattern, with no beneficial effect of the forget cue on recall of list-2 items. Thus, in order to properly examine whether not only list-1 recall but also list-2 recall shows persistent effects of the forget cue, fresh experiments would be required, in which the focus is on list 2 and subjects at test are asked to recall list-2 items first.

Although originally regarded as the two sides of the same coin (e.g., Bjork, 1970; Geiselman et al., 1983; Sahakyan & Kelley, 2002), more recent LMDF work suggests that list-2 enhancement and list-1 forgetting are partly mediated by different mechanisms. While both the inhibition account and the context-change account explain list-1 forgetting by retrieval mechanisms, there is evidence that differential encoding contributes to list-2 enhancement. For instance, Sahakyan and Delaney (2003) suggested that a forget cue can trigger a switch to more efficient encoding strategies, thus improving list-2 encoding in general. Pastötter and Bäuml (2010) proposed that a forget cue can reset encoding of early list 2 items and thus produce a greater primacy effect for list 2. If so, not only retrieval but also encoding processes may contribute to list-2 enhancement (for such a two-factor account of list-2 enhancement, see Pastötter et al., 2012), which leads to the expectation that list-2 enhancement may be at least as persistent as list-1 forgetting. Future work may address the issue and thus complement the present results on the persistence of list-1 forgetting with results on the persistence of list-2 enhancement.

Conclusions

The results of the present study demonstrate that LMDF can persist across prolonged retention interval, whereas context-dependent forgetting does not, which indicates that LMDF is not caused by mental context change. The results also show that persistent directed forgetting is not mediated by selective rehearsal and, in contrast to such a view, is present for incidentally studied items and absent after undemanding distractor tasks. These findings challenge the context-change and selective-rehearsal accounts of LMDF as well as a dual-mechanisms view, which assumes a role of both types of mechanisms in LMDF.

[1] This speculation was suggested to us by one of the reviewers.

Authors' note

This work was supported by a grant from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG; AB 594/3-1). We thank N. Basli, A. Bauer, A. Ederer, E. Meier, A. Nickl, D. Romanowski, S. Reisinger, S. Schobel, and S. Weitlaner for their help with data collection.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jml.2019.02.002.

References

Abel, M., & Bäuml, K.-H. T. (2013). Sleep can eliminate list-method directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 946–952.
Abel, M., & Bäuml, K.-H. T. (2017). Testing the context-change account of list-method directed forgetting: The role of retention interval. *Journal of Memory and Language, 92*, 170–182.
Anderson, M. C. (2005). The role of inhibitory control in forgetting unwanted memories: A consideration of three methods. In C. MacLeod, & B. Uttl (Eds.). *Dynamic cognitive processes* (pp. 159–190). Tokyo: Springer-Verlag.
Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences, 18*, 279–292.
Bäuml, K.-H., Hanslmayr, S., Pastötter, B., & Klimesch, W. (2008). Oscillatory correlates of intentional updating in episodic memory. *NeuroImage, 41*, 596–604.
Bäuml, K.-H., Pastötter, B., & Hanslmayr, S. (2010). Binding and inhibition in episodic memory – Cognitive, emotional, and neural processes. *Neuroscience & Biobehavioral Reviews, 34*, 1047–1054.
Bäuml, K.-H. T., & Samenieh, A. (2012). Influences of part-list cuing on different forms of episodic forgetting. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 38*, 366–375.
Barnier, A. J., Conway, M. A., Mayoh, L., Speyer, J., Avizmil, O., & Harris, C. B. (2007). Directed forgetting of recently recalled autobiographical memories. *Journal of Experimental Psychology: General, 136*, 301–322.
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. 139–172). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
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, & Cognnition, 19*, 603–616.
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. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger, & F. I. M. Craik (Eds.). *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309–330). Hillsdale: Lawrence Erlbaum Associates.
Brickenkamp, R., & Zillmer, E. A. (1998). *d2 test of attention.* Göttingen, Germany: Hogrefe & Huber.
Conway, M. A., Harries, K., Noyes, J., Racsmany, M., & Frankish, C. R. (2000). The disruption and dissolution of directed forgetting: Inhibitory control of memory. *Journal of Memory & Language, 43*, 409–430.
Delaney, P. F., Nghiem, K. N., & Waldum, E. R. (2009). The selective directed forgetting effect: Can people forget only part of a text? *Quarterly Journal of Experimental Psychology, 62*, 1542–1550.
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.
Divis, K. M., & Benjamin, A. S. (2014). Retrieval speeds context fluctuation: Why semantic generation enhances later learning, but hinders prior learning. *Memory & Cognition, 42*, 1049–1062.
Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review, 62*, 145–154.
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.
Hanslmayr, S., Volberg, G., Wimber, M., Oehler, N., Staudigl, T., Hartmann, T., ... Bäuml, K.-H. T. (2012). Prefrontally driven down-regulation of neural synchrony mediates goal-directed forgetting. *The Journal of Neuroscience, 32*, 14742–14751.
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.
Jonker, T. R., Seli, P., & MacLeod, C. M. (2013). Putting retrieval-induced forgetting in context: An inhibition-free, context-based account. *Psychological Review, 120*, 852–872.
Lee, Y.-S. (2012). Cognitive load hypothesis of item-method directed forgetting. *The*

Quarterly Journal of Experimental Psychology, 65, 1110–1122.
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.
Lehman, M., & Malmberg, K. J. (2011). Overcoming the effects of intentional forgetting. *Memory & Cognition, 39*, 335–347.
Liu, X. (2001). On the dynamics of directed forgetting: Facilitation and interference in the updating of human memory. *Dissertation Abstracts International, 61*, 6159B.
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 (Vol. Ed.), *The Psychology of Learning and Motivation: Vol. 43*, (pp. 163–214). San Diego, CA: Academic Press.
Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods, 43*, 679–690.
McGeoch, J. (1932). Forgetting and the law of disuse. *Psychological Review, 39*, 352–370.
Mensink, G., & Raaijmakers, J. G. (1988). A model for interference and forgetting. *Psychological Review, 95*, 434–455.
Oswald, W. D., & Roth, E. (1987). *Der Zahlenverbindungstest (ZVT) [Connect-the-Numbers Test]*. Göttingen, Germany: Hogrefe.
Pastötter, B., & Bäuml, K.-H. (2007). The crucial role of postcue encoding in directed forgetting and context-dependent forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 977–982.
Pastötter, B., & Bäuml, K.-H. (2010). Amount of postcue encoding predicts amount of directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 54–65.
Pastötter, B., Bäuml, K.-H., & Hanslmayr, S. (2008). Oscillatory brain activity before and after an internal context change - Evidence for a reset of encoding processes. *NeuroImage, 43*, 173–181.
Pastötter, B., Kliegl, O., & Bäuml, K.-H. T. (2012). List-method directed forgetting: The forget cue improves both encoding and retrieval of postcue information. *Memory & Cognition, 40*, 861–873.
Pastötter, B., Kliegl, O., & Bäuml, K.-H. T. (2016). List-method directed forgetting: Evidence for the reset-of-encoding hypothesis employing item-recognition testing. *Memory, 24*, 63–74.
Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K.-H. T. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 37*, 287–297.
Raftery, A. E. (1995). Bayesian model selection in social research. In P. V. Marsden (Ed.). *Sociological methodology 1995* (pp. 111–196). Cambridge: Blackwell.
Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology, 41*, 1–48.
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. (2010). Item-specific encoding produces an additional benefit of directed forgetting: Evidence from intrusion errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1346–1354.
Sahakyan, L., Delaney, P. F., Foster, N. L., & Abushanab, B. (2013). List-method directed forgetting in cognitive and clinical research: A theoretical and methodological review. In B. H. Ross (Vol. Ed.), *Psychology of learning and motivation: Vol. 59*, (pp. 131–189). New York: Elsevier.
Sahakyan, L., & Foster, N. L. (2009). Intentional forgetting of actions: Comparison of list-method and item-method directed forgetting. *Journal of Memory and Language, 61*, 134–152.
Sahakyan, L., & Goodmon, L. B. (2010). Theoretical implications of extralist probes for directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 920–937.
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.
Sahakyan, L., Waldum, E., Benjamin, A. S., & Bickett, S. P. (2009). Where is the forgetting with list-method directed forgetting in recognition? *Memory & Cognition, 37*, 464–476.
Schlichting, A., & Bäuml, K.-H. T. (2017). Brief wakeful resting can eliminate directed forgetting. *Memory, 25*, 254–260.
Shapiro, S., Lindsey, C., & Krishnan, H. S. (2006). Intentional forgetting as a facilitator for recalling new product attributes. *Journal of Experimental Psychology: Applied, 12*, 251–263.
Wixted, J. T. (2004). On common ground: Jost's (1897) law of forgetting and Ribot's (1881) law of retrograde amnesia. *Psychological Review, 111*, 864–879.
Zimmermann, P., & Fimm, B. (2007). Testbatterie zur Aufmerksamkeitsprüfung (TAP) [Battery of tests for assessing attention]. Herzogenrath, Germany: Psytest.