The roles of delay and retroactive interference in retrieval-induced forgetting

Magdalena Abel · Karl-Heinz T. Bäuml

Abstract When subjects study items from different categories and then repeatedly retrieve some of the items from some of the categories, retrieval practice typically improves recall of the practiced items but impairs recall of related but unpracticed items, relative to control items from unpracticed categories. Here, we report the results of three experiments, in which we examined practiced and unpracticed items' delay-induced forgetting (Exp. 1) and their susceptibility to retroactive interference (Exps. 2 and 3). Control items showed the expected memory impairment after longer delay between practice and test and in the presence of retroactive interference. In contrast, both the practiced and the related unpracticed items showed hardly any forgetting under these conditions. The findings are consistent with the results from recent testing-effect studies, which have reported reduced delay-induced forgetting and reduced susceptibility to interference for retrieval-practiced items, and generalize the results to related unpracticed items. The findings are discussed with respect to the inhibitory and noninhibitory accounts of retrieval-induced forgetting, as well as the possible role of selective segregation processes, which may be induced by retrieval practice.

Keywords Retrieval practice · Retrieval-induced forgetting · Testing effect · Delay-induced forgetting · Retroactive interference

Introduction

Selectively retrieving part of a previously studied episode can result in forgetting of the remaining, nonretrieved information. Such retrieval-induced forgetting (RIF) has repeatedly been

M. Abel · K.-H. T. Bäuml (✉)
Department of Experimental Psychology, Regensburg University,
93040 Regensburg, Germany
e-mail: karl-heinz.baeuml@psychologie.uni-regensburg.de

demonstrated using the retrieval-practice paradigm (M. C. Anderson, Bjork, & Bjork, 1994). In this paradigm, subjects often study a categorized item list (e.g., PROFESSION–teacher, PROFESSION–electrician, VEGETABLE–pepper, VEGETABLE–tomato), and, after study, repeatedly retrieve half of the items from half of the categories (e.g., PROFESSION–elec__?). The typical finding is that, on a later category-cued recall test, memory performance for the practiced items (i.e., PROFESSION–electrician) is enhanced, but memory performance for the unpracticed items from the practiced categories (i.e., PROFESSION–teacher) is impaired, relative to the control items from the unpracticed categories (i.e., VEGETABLE–pepper, VEGETABLE–tomato). The two effects of retrieval practice have been found over a wide range of materials and settings, including visuospatial materials (Ciranni & Shimamura, 1999), autobiographical memory (Barnier, Hung, & Conway, 2004), foreign-language acquisition scenarios (Levy, McVeigh, Marful, & Anderson, 2007), and conversations (Coman, Manier, & Hirst, 2009).

RIF is often regarded as the outcome of a cognitive control process operating during retrieval practice. While it is assumed that retrieval practice strengthens the practiced items, the proposal is that, during repeated retrieval attempts, the not-to-be-retrieved (unpracticed) items interfere and are inhibited so as to overcome the interference (e.g., M. C. Anderson et al., 1994; M. C. Anderson & Spellman, 1995). Although the results of a large number of both behavioral and neurocognitive studies have supported this account (for reviews, see M. C. Anderson, 2003; Bäuml, Pastötter, & Hanslmayr, 2010), it has also been criticized, and a noninhibitory explanation has been suggested (e.g., Camp, Pecher, & Schmidt, 2007; Jakab & Raaijmakers, 2009). According to this alternative view, retrieval practice just strengthens the retrieval-practiced items, which at test creates a high level of interference, and thus reduces recall of the related unpracticed items, which supposedly creates the RIF effect (for recent discussions of these accounts, see Storm & Levy, 2012, and Raaijmakers & Jakab, 2013).

Focusing on the beneficial effects of retrieval practice, it has repeatedly been demonstrated in the testing-effect literature that retrieval-practiced items show reduced forgetting relative to restudied items if a delay interval of several days is introduced between study and test (e.g., Karpicke & Roediger, 2007; Roediger & Karpicke, 2006). A similar finding has been reported in the retrieval-practice paradigm, with reduced forgetting of practiced items relative to control items after delay intervals of 12 or 24 h (e.g., Abel & Bäuml, 2012; Chan, 2009, Exp. 1; but see MacLeod & Macrae, 2001). Moreover, these studies have provided evidence that the reduction in delay-induced forgetting is not restricted to the retrieval-practiced items, but can generalize to the related unpracticed items. Consistently in a number of studies the RIF effect has been absent after longer delays (Abel & Bäuml, 2012; Chan, 2009; MacLeod & Macrae, 2001; Racsmány, Conway, & Demeter, 2010), although this result has not arisen in all of the previous studies (Garcia-Bajos, Migueles, & Anderson, 2009; Storm, Bjork, & Bjork, 2012).[1]

Bjork and colleagues have provided theoretical accounts that can explain why practiced and unpracticed items show reduced delay-induced forgetting. Regarding the effect on the practiced items, Kornell, Bjork, and Garcia (2011) proposed that retrieval practice strengthens the successfully retrieved items more than items that are restudied. At the core of their proposal is the view that (the stronger) successfully retrieved items show the same decrease in strength level with delay as (the weaker) restudied items. However, because of their higher initial strength, successfully retrieved items may remain above the recall threshold even after longer delay, so that, on average, retrieval-practiced items show less delay-induced forgetting than do restudied items (for details of the account, see Kornell et al., 2011). Regarding the effect on the unpracticed items, Storm et al. (2012) suggested that the originally inhibited unpracticed items may show recovery from inhibition if the recall test is delayed. According to their proposal, recovery from inhibition is intrinsic to the very idea of inhibition, and unpracticed items

that were not recallable on an early test may become recallable on a later test due to intermittent release processes. Consistent with the proposal, these authors found that unpracticed and control items suffered equal amounts of obliviscence with delay, but that unpracticed items benefited more from reminiscence than did control items.

Retrieval-practiced items have not only been demonstrated to show reduced delay-induced forgetting, but to show reduced susceptibility to interference as well, at least in the testing-effect literature. Employing a retroactive interference task, Halamish and Bjork (2011, Exp. 3) reported that retrieval practice on originally studied target material reduces the targets' susceptibility to interference when further, non-target material is studied between practice and the final recall test (for related results, see Potts & Shanks, 2012). To account for the finding, the authors suggested that retrieval practice may help to distinguish tested information from interfering information, and thus reduce the practiced items' susceptibility to retroactive interference (Halamish & Bjork, 2011, p. 810). Both the findings and the suggested account are in line with studies on the role of retrieval practice for proactive interference, which have shown that the testing of previously studied nontarget material can improve recall of subsequently studied target material and have indicated a role of segregation processes in this form of recall improvement (e.g., Bäuml & Kliegl, 2013; Szpunar, McDermott, & Roediger, 2008; Tulving & Watkins, 1974).

Whether the finding that retrieval practice insulates practiced items against retroactive interference is restricted to testing-effect paradigms or generalizes to the retrieval-practice paradigm has not been examined to date. Although it may appear likely a priori that the reduced susceptibility to interference would generalize to the practiced (relative to the control) items in the retrieval-practice paradigm, it is less clear how interference might affect the unpracticed items. On the basis of Halamish and Bjork's (2011) suggestion that retrieval practice can help to distinguish retrieval-practiced information from interfering information, one might speculate that unpracticed items would show reduced susceptibility to interference, as well. Indeed, if retrieval triggered segregation processes that were not restricted to the retrieval-practiced material, but generalized to those items that were studied as members of the same list (and category) as the practiced items but were not themselves practiced, then not only could the practiced items, but also the related unpracticed items, show reduced susceptibility to retroactive interference.

This study reports the results of three experiments designed to examine practiced and unpracticed items' delay-induced forgetting and the items' susceptibility to interference in the retrieval-practice paradigm. In Experiment 1, we examined the effect of delay between practice and test on recall of the practiced, unpracticed, and control items; subjects studied a semantically categorized item list, practiced some of the items

---

[1] Although no, or at least reduced, RIF has typically been observed for (wake) delay intervals of 12 h and more, it has been demonstrated that RIF can be maintained if regular nocturnal sleep follows immediately upon retrieval practice (e.g., Abel & Bäuml, 2012). These results indicate that it may be critical for the results of RIF experiments whether retrieval practice is followed by a wake or sleep delay. We accounted for this finding in our experiments by conducting the sessions before noon or during early afternoon (see the Method section of Exp. 1). Some other studies have also reported that RIF persists with delay (Migueles & Garcia-Bajos, 2007; Saunders, Fernandes, & Kosnes, 2009; Storm, Bjork, Bjork, & Nestojko, 2006). However, these studies used a repeated-test design in which subjects were given an initial test shortly after retrieval practice and were then tested on the same items again 24 h later. With such a design, however, the RIF effect found in the delayed test may have been caused by the recall difference in the initial test, rather than by any lasting retrieval-practice effects (see Abel & Bäuml, 2012, or Storm et al. 2012).

from some of the categories, and then were tested on the material after a short, 3-min, or a longer, 24-h, delay interval. Using similar material, in Experiment 2 we examined the effect of retroactive interference on recall of the practiced, unpracticed, and control items; subjects again studied a semantically categorized item list, practiced some of the items from some of the categories, and then encoded further items from the originally studied categories before they were tested on the initially studied items. Experiment 3 was largely identical to Experiment 2, but it followed prior work (e.g., Ciranni & Shimamura, 1999; Spitzer & Bäuml, 2009) and employed episodic instead of semantic categories in order to examine items' susceptibility to interference. Subjects studied items presented in different font colors in order to establish episodic (color) categories, practiced some of the items from some of the color categories, and then were tested on the initially studied items; crucially, the subjects had encoded further items between practice and test that were presented in the same font colors as the originally studied items.

On the basis of some of the prior work on the effects of delay between practice and test for practiced and unpracticed items (Abel & Bäuml, 2012; Chan, 2009) and the theoretical framework introduced by Bjork and colleagues (Kornell et al., 2011; Storm et al., 2012), we expected to find reduced delay-induced forgetting for both practiced and unpracticed items in Experiment 1. On the basis of prior work on the effects of retroactive interference for practiced items in the testing-effect paradigm and Halamish and Bjork's (2011) segregation account of the findings, we expected in both Experiments 2 and 3 to find reduced susceptibility to interference for the practiced items. If not only the practiced items, but also the related unpracticed items, were subject to such segregation processes, then not only the practiced, but also the unpracticed, items might show a reduced effect of retroactive interference.

Experiment 1

Method

Subjects

A group of 32 students enrolled at Regensburg University took part in the experiment in return for either partial course credit or a compensatory amount of money ($M = 23.0$ years, range 20–28 years). All of the subjects spoke German as their native language.

Material

The materials consisted of two sets of items, each containing 36 concrete nouns. Each item set was used equally often across experimental conditions. Within each set, the items belonged to six different semantic categories, with each category comprising six exemplars (Van Overschelde, Rawson, & Dunlosky, 2004). Within categories, all items had unique initial letters.

Design

The experiment had a 3×2 design. The factors item type (practiced items, unpracticed items, control items) and delay (3 min, 24 h) were manipulated within subjects. In both delay conditions, the retrieval-practice paradigm was applied: After initial study of one of the two item sets, partial retrieval practice created practiced items, unpracticed items, and control items. The two delay conditions differed in whether the final memory test for the previously studied items was administered after a short (3-min) delay or after a longer (24-h) delay. The sequence of conditions was balanced across subjects. Sessions were conducted before noon or during the early afternoon to prevent subjects from going to sleep within a few hours after encoding, thus hopefully ruling out effects of sleep-associated memory consolidation (e.g., Abel & Bäuml, 2012; Diekelmann & Born, 2010). Indeed, the subjects reported that they had regularly gone to bed at night and had not taken any naps during the day.

Procedure

Study phase In each of the two delay conditions, the subjects studied 36 items belonging to six different categories. The items were presented together with their category label centrally on the computer screen for 3 s each. They were displayed individually and in a pseudorandomized order, with no two items of the same category following each other.

Retrieval-practice phase Immediately after study, subjects in both conditions were asked to recall half of the words from four of the six categories in two successive retrieval cycles. The words' category labels and unique word stems were provided as retrieval cues. The subjects had 8 s to recall each item and to write down their answers on a sheet of paper.

Delay manipulation Either a short or a long delay interval was placed between retrieval practice and test. In the short-delay condition, subjects solved simple arithmetic problems for 3 min before taking the final memory test. In the long-delay condition, they engaged in the same task for 3 min, but then were allowed to leave the laboratory and returned after 24 h to complete the final test.

Test phase Before testing started, subjects were asked to try to remember as many of the previously presented items as possible. The words' category labels and unique first letters were provided as retrieval cues and were presented
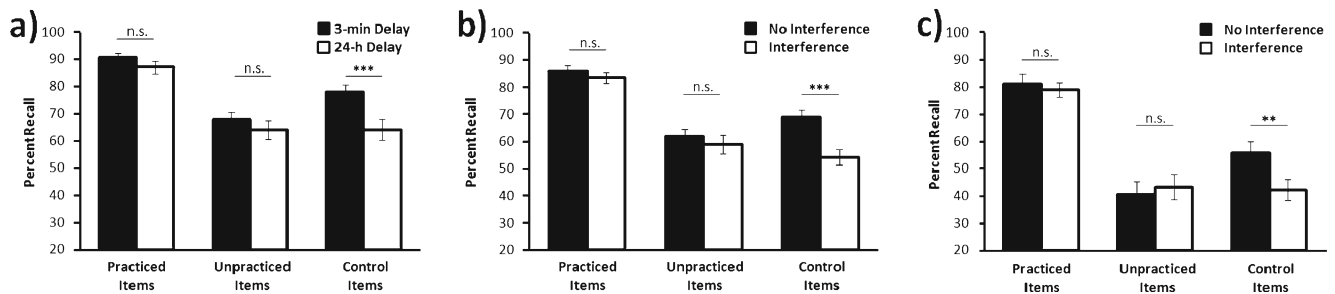
Fig. 1 (**a**) Results of Experiment 1: Mean recall performance for the three item types (practiced, unpracticed, control) as a function of delay (3 min, 24 h). The item materials were semantically categorized. (**b**) Results of Experiment 2: Mean recall performance for the three item types (practiced, unpracticed, control) as a function of interference level (no interference, interference). The item materials again were semantically categorized. (**c**) Results of Experiment 3: Mean recall performance for the three item types (practiced, unpracticed, control) as a function of interference level (no interference, interference). The item materials were episodically categorized. Error bars represent standard errors. $^{**}p \le .01$; $^{***}p \le .001$; n.s. = nonsignificant

successively in a blocked, randomized manner for 8 s each, positioned centrally on the computer screen. The sequence of categories was random, but all of the items of a category were tested successively. The subjects had 8 s to recall each item and were asked to write down their answers on a sheet of paper, and then the next retrieval cue appeared on the screen.

When the first delay condition was completed, subjects were informed that the studied material would no longer be needed. They were given a break and, after the break, were asked to memorize new item material for the second experimental condition. When this condition was completed, they were debriefed and thanked for their participation.

Results

Retrieval success

The mean recall success rates in the retrieval-practice phase were 95.1 % ($SD = 6.2$) in the short-delay and 93.4 % ($SD = 6.7$) in the long-delay condition. The difference was not significant, $t(31) = 1.02$, $p = .316$.

Delay-induced forgetting

Figure 1a shows recall of the practiced, unpracticed, and control items after the two delay intervals. A 3×2 analysis of variance (ANOVA) with the factors item type (practiced items, unpracticed items, control items) and delay (3 min, 24 h) revealed significant main effects of item type, $F(2, 62) = 43.48$, $MSE = 212.44$, $p < .001$, $\eta^2 = .58$, and delay, $F(1, 31) = 13.46$, $MSE = 180.68$, $p = .001$, $\eta^2 = .30$. The main effect of item type reflects the pattern of better recall for practiced than for control items, and of better recall for control than for unpracticed items (see below for details); the main effect of delay reflects a general decrease in recall in the long-delay condition. In addition, a significant interaction between the

two factors, $F(2, 62) = 4.11$, $MSE = 140.75$, $p = .021$, $\eta^2 = .12$, suggests that delay affected the three item types differently.

Planned comparisons were calculated in order to compare memory performance after the short and long delays, separately for the three item types. Whereas the recall of control items decreased significantly across the 24-h delay (78.1 % vs. 64.1 %), $t(31) = 4.24$, $p < .001$, $d = 0.75$, neither practiced nor unpracticed items showed reliable delay-induced forgetting [practiced items: 90.6 % vs. 87.0 %, $t(31) = 1.54$, $p = .133$; unpracticed items: 67.7 % vs. 64.1 %, $t(31) = 1.04$, $p = .307$]. Consistently, RIF (i.e., impaired memory performance for unpracticed vs. control items) was present in the short-delay condition (67.7 % vs. 78.1 %), $t(31) = 3.79$, $p = .001$, $d = 0.67$, but was absent in the long-delay condition (64.1 % vs. 64.1 %), $t(31) < 1.0$. Retrieval-induced enhancement (i.e., better memory performance for practiced vs. control items) was present in both delay conditions, $ts(31) \ge 4.77$, $ps < .001$, $ds \ge 0.88$; the enhancement, however, was greater after the long than after the short delay interval, $F(1, 31) = 5.24$, $MSE = 165.77$, $p = .029$, $\eta^2 = .15$.[2]

Discussion

By showing the typical beneficial effect for practiced items and the typical detrimental effect for unpracticed items, the

---

[2] In this experiment, delay was implemented as a within-subjects factor, and subjects successively completed the two delay conditions, with the sequence of conditions being counterbalanced across subjects. To analyze whether the sequence of conditions affected the observed pattern of results, we ran a $3 \times 2 \times 2$ ANOVA with the within-subjects factors item type and delay, and the between-subjects factor Sequence of Conditions (short-delay condition first, long-delay condition first). We found no significant main effect of sequence of conditions, $F(1, 30) = 1.65$, $p = .208$, no significant two-way interactions of sequence of conditions and the other two factors, $Fs < 1.0$, and no significant three-way interaction, $F(2, 60) = 1.02$, $p = .367$, indicating that the sequence of conditions did not affect the results. Corresponding analyses were conducted for Experiments 2 and 3, with analogous results.

results of the short-delay condition replicated the standard finding of the two effects of partial retrieval practice (e.g., M. C. Anderson et al., 1994; M. C. Anderson & Spellman, 1995). Going beyond this finding, the results for the long-delay condition still showed the beneficial effect of retrieval practice for the practiced items, but they no longer showed any detrimental effect of retrieval practice for the unpracticed items. In fact, whereas the control items showed forgetting from the short to the long delay interval, the unpracticed items did not, which made the RIF effect disappear. Like the unpracticed items, the practiced items also did not show reliable delay-induced forgetting, thus leading to greater retrieval-induced enhancement after the longer delay.

The present results mimic findings from recent work using the retrieval-practice paradigm, which also indicated that both practiced and related unpracticed items can show reduced, or even eliminated, delay-induced forgetting (e.g., Abel & Bäuml, 2012; Chan, 2009). Theoretically, the results are in line with the accounts provided by Bjork and colleagues, according to which the high level of strengthening of the practiced items reduces these items' delay-induced forgetting (Kornell et al. 2011), and recovery from inhibition can reduce delay-induced forgetting among unpracticed items (Storm et al., 2012).

Experiment 2

Experiment 2 extended Experiment 1 by examining whether practiced and unpracticed items might not only show reduced delay-induced forgetting, but also show reduced susceptibility to interference. In Experiment 2, we addressed the issue by examining the influence of retroactive interference on the recall of practiced, unpracticed, and control items, again employing the retrieval-practice paradigm. Subjects studied a semantically categorized item list and then repeatedly retrieved some of the items from some of the categories. Subjects were then asked to recall the studied material after subsequent study of another categorized (nontarget) list, or in the absence of such a list.[3]

---

[3] If forgetting occurs over time mostly because of interfering activities that occur over that time (e.g., McGeoch, 1932), one would expect the results of Experiment 2 to replicate those of Experiment 1, because both experiments would concern the effects of retroactive interference. However, arguably, delay-induced forgetting is due not only to increasing interference, but can be due to a number of other factors, including, for instance, changes in the internal and external contexts between study and test (e.g., J. R. Anderson, 2000). If so, there is no a priori guarantee that the results of Experiment 1 would generalize to the conditions of Experiment 2 (for experimental dissociations between the effects of interference and context change, see, e.g., Bäuml & Samenieh, 2012).

Method

Subjects

A new sample of 32 students took part in the experiment ($M =$ 22.3 years, range 19–28 years).

Material

Two new item sets were compiled that consisted of 12 exemplars of six semantic categories each (Van Overschelde et al., 2004). The 12 exemplars were divided into six target and six nontarget items. The target items were used to conduct the standard retrieval-practice paradigm in two consecutive conditions; the nontarget items were used for additional study that could induce retroactive interference after the retrieval-practice phase, in one of the two conditions. Within categories, all items had unique initial letters. The item sets were counterbalanced across conditions.

Design

The experiment had a 3 × 2 design. The two factors item type (practiced items, unpracticed items, control items) and Interference (interference, no interference) were both manipulated within subjects. Subjects consecutively completed two experimental conditions; in both conditions, the retrieval-practice paradigm was employed, and partial retrieval practice created practiced items, unpracticed items, and control items. The two conditions differed in whether or not retroactive interference was induced by presenting the nontarget items before the final memory test. The sequence of conditions was balanced across subjects.

Procedure

Study and retrieval-practice phases The study phase and retrieval-practice phase were identical to those of Experiment 1. Subjects initially studied the categorized item materials and then, in the intermediate practice phase, recalled half of the words from four of the six semantic categories, thus creating the three item types.

Interference manipulation Either an additional study phase (interference condition) or a distractor task of equivalent duration (no-interference condition) was performed between retrieval practice and the test. In the interference condition, subjects studied the 36 nontarget items that belonged to the same categories as the target items that had been encoded during initial study. Again, items and their category labels were presented for 3 s each and in a pseudorandomized manner. Right before and after study of the nontarget items,

subjects counted backward in steps of three for 30 s in order to control for working memory effects. In the no-interference condition, subjects solved simple arithmetic problems between retrieval practice and test for 3 min in order to rule out time as a confounding variable.

Test phase At test, the subjects were provided with the words' category labels and unique first letters as retrieval cues and were asked to recall as many of the previously studied items as possible. The procedure was again largely identical to the one that had been applied in Experiment 1; in the interference conditions, however, the initially studied target items were tested first, and the additionally studied nontargets were tested second.

Results

Retrieval success

The mean recall success rates in the retrieval-practice phase were 90.8 % (SD = 7.8) in the presence of interference and 90.9 % (SD = 8.2) in its absence. No reliable difference emerged between the conditions, $t(31) < 1.0$.

Susceptibility to interference

Figure 1b shows recall of the practiced, unpracticed, and control items in the two interference conditions. A 3 × 2 ANOVA with the factors item type (practiced items, unpracticed items, control items) and Interference (interference, no interference) revealed significant main effects of item type, $F(2, 62) = 60.37$, $MSE = 197.74$, $p < .001$, $\eta^2 = .66$, and interference, $F(1, 31) = 12.77$, $MSE = 167.93$, $p = .001$, $\eta^2 = .29$. The main effect of item type reflects the pattern of better recall for practiced than for control items, and of better recall for control than for unpracticed items (see below for details); the main effect of interference reflects a general decrease in recall caused by the additional study of the nontarget list. In addition, a significant interaction between the two factors emerged, $F(2, 62) = 4.52$, $MSE = 177.58$, $p = .015$, $\eta^2 = .13$, suggesting that interference affected the three item types differently.

Planned comparisons were calculated to further compare memory performance for the three item types across interference conditions. Whereas the control items showed significant retroactive interference (69.0 % vs. 54.2 %), $t(31) = 4.62$, $p < .001$, $d = 0.82$, neither the practiced nor the unpracticed items suffered from such interference [practiced items: 85.7 % vs. 83.6 %, $t(31) < 1.0$, $p = .354$; unpracticed items: 62.0 % vs. 58.9 %, $t(31) < 1.0$, $p = .460$]. Likewise, RIF (i.e., impaired memory performance for unpracticed vs. control items) arose in the absence of interference (62.0 % vs. 69.0 %), $t(31) = 3.04$,

$p = .005$, $d = 0.54$, but did not arise in the presence of interference (58.9 % vs. 54.2 %), $t(31) < 1.0$, $p = .367$. Retrieval-induced enhancement (i.e., better memory performance for practiced vs. control items) emerged regardless of interference condition, $ts(31) \geq 5.85$, $ps < .001$, $ds \geq 1.04$, but it was greater in the presence than in the absence of interference, $F(1, 31) = 9.30$, $MSE = 139.99$, $p = .005$, $\eta^2 = .23$.

Discussion

The results of Experiment 2 replicated the two standard effects of partial retrieval practice—that is, the beneficial effect for practiced items and the detrimental effect for unpracticed items—at least in the absence of retroactive interference (e.g., M. C. Anderson et al., 1994; M. C. Anderson & Spellman, 1995). In contrast, in the presence of retroactive interference, only the beneficial effect, but *not* the detrimental effect, of retrieval practice emerged. This result reflects the fact that the control items, but not the unpracticed items, showed a retroactive interference effect, which made the RIF effect disappear. Like the unpracticed items, the practiced items did not show reliable susceptibility to interference, which led to a greater retrieval-induced enhancement in the presence of interference.

The present results for the practiced items mimic the results of Halamish and Bjork's (2011) recent testing-effect study, which showed a reduced retroactive interference effect for retrieval-practiced items. They are also consistent with Halamish and Bjork's view that retrieval practice triggers segregation processes between the retrieval-practiced items and the subsequently encoded items, and thus reduces the interference effect for the practiced items. The finding of parallel interference effects for the practiced and unpracticed items is new, and it indicates that the suggested segregation process may not be restricted to the practiced items, but can also be effective for the categories' unpracticed items.

Experiment 3

The results of Experiment 2 provided the first demonstration that, in the retrieval-practice paradigm, both practiced and unpracticed items can show reduced susceptibility to retroactive interference. To ensure that this novel finding was not spurious, we aimed to replicate the results. Experiment 3 addressed this issue. The experiment was largely identical to Experiment 2, but it examined the influence of retroactive interference on recall of the three item types using episodic rather than semantic categories. Prior work had demonstrated that RIF is not restricted to semantically categorized lists, but is also present when new (episodically acquired) categories are established during initial study (e.g., Abel & Bäuml, 2012;

Ciranni & Shimamura, 1999; Spitzer & Bäuml, 2009). The results of the experiment thus would show whether retroactive interference plays similar roles for RIF in episodically and semantically structured lists.

Method

Subjects

A fresh sample of 28 students took part in the experiment (M = 22.8 years, range 19–29 years).

Material

Three item sets were compiled that consisted of 24 unrelated items each (taken from different semantic categories; Van Overschelde et al., 2004). Two of the item sets were randomly chosen as target items and used to conduct the standard retrieval-practice paradigm in two consecutive conditions; the third item set was always used as additional study (or nontarget) materials to induce retroactive interference after the retrieval-practice phase in one of the two conditions. The two target item sets were counterbalanced across conditions.

Each item set was further divided into three clusters, each comprising eight items. Each cluster's items were presented in a different font color during study, thus creating three episodic color categories (e.g., Abel & Bäuml, 2012; Spitzer & Bäuml, 2009). In one of the conditions, the colors red, green, and blue were used; in the other condition, the colors magenta, yellow, and turquoise were applied as categories. Within categories, all items had unique word stems.

Design

The experiment had the same 3 × 2 design as Experiment 2. The two factors item type (practiced items, unpracticed items, control items) and Interference (interference, no interference) were both manipulated within subjects. The subjects completed two blocks of the retrieval-practice paradigm that differed in whether or not retroactive interference was induced by presenting the nontarget items before the final memory test. The sequence of conditions was balanced across subjects.

Procedure

The procedure was largely identical to that of Experiment 2. The only difference was that items were presented in three different font colors, for 3 s each and on two consecutive study cycles during the initial study phase, to create episodic color categories. Subjects were asked to encode the single items with respect to their font colors, thus establishing fresh (nonsemantic and not preexisting) categories. During retrieval practice, the items' word stems were presented in their respective font colors as retrieval cues, and subjects were asked to complement the cues with previously studied items from the same color category. The subjects practiced half of the items from two of the three color categories on two successive retrieval cycles. After retrieval practice, subjects either completed arithmetic problems for 3 min before taking the final test or studied the 24 nontarget items (presented in the same font colors and in the same manner as the target items during initial study). At test, items' initial letters were presented in their font colors as retrieval cues, and subjects were asked to recall as many of the initially studied items as possible.

Results and discussion

Retrieval success

The mean recall success rates in the retrieval-practice phase were 93.5 % (SD = 10.6) in the presence of interference and 93.1 % (SD = 8.7) in its absence. No difference emerged between conditions, $t(27) < 1.0$.

Susceptibility to interference

Figure 1c shows recall of the practiced, unpracticed, and control items in the two interference conditions. A 3 × 2 ANOVA with the factors item type (practiced items, unpracticed items, control items) and Interference (interference, no interference) revealed a significant main effect of item type, $F(2, 54) = 67.02$, $MSE = 344.68$, $p < .001$, $\eta^2 = .71$, but no main effect of interference, $F(1, 27) = 2.83$, $MSE = 283.85$, $p = .104$, $\eta^2 = .10$. The main effect of item type reflects the pattern of better recall for practiced than for control items, and of better recall for control than for unpracticed items (see below for details); Although we observed no main effect of interference, there was at least a trend for a general decrease in recall in the presence of interference. More importantly, a significant interaction between the two factors, $F(2, 54) = 3.46$, $MSE = 281.08$, $p = .039$, $\eta^2 = .11$, suggests that interference affected the three item types differently.

Planned comparisons were carried out to compare memory performance for the three item types across interference conditions. Whereas the control items showed significant retroactive interference (55.8 % vs. 42.2 %), $t(27) = 3.16$, $p = .004$, $d = 0.60$, neither the practiced nor the unpracticed items suffered from such interference [practiced items: 81.3 % vs. 79.0 %, $t(27) < 1.0$, $p = .525$; unpracticed items: 40.6 % vs. 43.3 %, $t(27) < 1.0$, $p = .628$]. Likewise, RIF (i.e., impaired memory performance for unpracticed vs. control items) arose in the absence of interference (55.8 % vs. 40.6 %), $t(27) = 3.43$, $p = .002$, $d = 0.65$, but did not arise in the presence of

interference (42.2 % vs. 43.3 %), $t(27) < 1.0$, $p = .842$. Retrieval-induced enhancement (i.e., better memory performance for practiced vs. control items) emerged regardless of interference condition, $ts(27) \geq 5.36$, $ps < .001$, $ds \geq 1.01$, but it was greater in the presence than in the absence of interference, $F(1, 27) = 4.86$, $MSE = 185.36$, $p = .036$, $\eta^2 = .15$.

The results of Experiment 3 replicated those of Experiment 2 using episodic rather than semantic categories. Although control items showed a reliable retroactive interference effect, no such effect was present for practiced and unpracticed items. In the same vein, a reliable RIF effect arose in the absence of retroactive interference, whereas no such effect emerged in its presence. The finding of practiced and unpracticed items' reduced susceptibility to interference that we reported in Experiment 2 thus is not restricted to semantically structured lists, but generalizes to episodically structured material.

General discussion

The results from the present series of experiments have replicated prior RIF work by showing that, after a relatively short delay between practice and test and in the absence of retroactive interference, retrieval practice improves recall of the practiced items but impairs recall of related unpracticed items, relative to the control items (M. C. Anderson et al., 1994; M. C. Anderson & Spellman, 1995). In addition, the results demonstrated that both a longer delay between practice and test and the presence of (retroactive) interference have only minor influences on the recall of practiced and unpracticed items. Indeed, both practiced and unpracticed items showed reduced delay-induced forgetting and reduced susceptibility to interference relative to the control items, which in both cases enhanced the size of the retrieval-facilitation effect for the practiced items and reduced the size of the RIF effect for the unpracticed items. While the delay finding replicates the results from some previous studies (e.g., Abel & Bäuml, 2012; Chan, 2009; see also Chan, McDermott, & Roediger, 2006), the interference finding is the first demonstration of a reduced susceptibility to interference for practiced and unpracticed items in the retrieval-practice paradigm.

The results on practiced and unpracticed items' delay-induced forgetting are consistent with theoretical accounts of retrieval-practice effects, as they have been introduced by Kornell et al. (2011) and Storm et al. (2012). Kornell et al. suggested that retrieval practice can lead to an exceptionally high level of strengthening of practiced items and that such strengthening can make the practiced items' recall rates relatively immune to delay-induced forgetting. On the basis of the inhibitory view of RIF, Storm et al. (2012) argued that

inhibited unpracticed items could be subject to intermittent release processes, so that they might not be recallable on an early test but become recallable on a later test, and thus show reduced forgetting after longer delay. The present results are in line with the two suggested accounts.

In contrast to the inhibitory view of RIF, the noninhibitory view of RIF claims that retrieval practice does nothing but strengthen the retrieval-practiced items, which at test is supposed to create a high level of interference and to reduce recall of the unpracticed items (e.g., Camp et al., 2007; Jakab & Raaijmakers, 2009). If retrieval practice led to an exceptionally high level of strengthening of the practiced items, this account could explain the reduced delay-induced forgetting of practiced items, very similar to how Kornell et al. (2011) explained the effect (see above). Regarding unpracticed items, the account suggests that the delay-induced decrease in the strength level of the practiced items reduces these items' (absolute) interference level. Whether such decrease would result in a reduced RIF effect after longer delay is unclear, however, because with increasing delay, the strength of unpracticed items should decrease as well, which might leave the (relative) sampling chances for the unpracticed items largely unaffected and the RIF effect unchanged.

The finding that practiced items show reduced susceptibility to interference agrees with Halamish and Bjork's (2011) suggestion that retrieval practice can help to distinguish retrieval-practiced information from subsequently encoded interfering information (see also Szpunar et al., 2008). The fact that unpracticed items also show reduced susceptibility to interference indicates that such segregation is not restricted to the retrieval-practiced material, but generalizes to those items that were studied as members of the same list and category as the practiced items but were not themselves practiced. Prior work in the testing-effect literature has examined the possible segregation effects of retrieval practice by employing categorized lists (e.g., Szpunar et al., 2008), lists of unrelated items (e.g., Pastötter, Schicker, Niedernhuber, & Bäuml, 2011), or paired associates (e.g., Halamish & Bjork, 2011). Whereas in all of these previous studies retrieval practice was given for all of the previously studied items, here we used categorized lists and subjects were asked to retrieve only some of the items from some of the categories. Therefore, the present results indicate that segregation is effective for all of the (practiced and unpracticed) items from practiced categories, but is quite ineffective for (control) items from unpracticed categories. This finding challenges the view that retrieval cycles between the study of lists merely alter subjects' internal context. Such context change should lead to specific context cues for each single list, and thus enhance list discrimination for all list items at test (e.g., Criss & Shiffrin, 2004; Howard & Kahana, 2002; Pastötter et al., 2011). In contrast, any discrimination benefits seem to differ between list items,

and to depend on which item sets were (partially) practiced and which were not.[4]

Both the inhibitory account and the noninhibitory account of RIF have problems to explain the reduced susceptibility to interference of the practiced and unpracticed items. On the basis of the inhibitory account of RIF, one may argue that unpracticed items show reduced susceptibility to interference because the encoding of the new, related information triggers release processes on the inhibited information. Although reexposure of the inhibited items after retrieval practice can in fact reduce the RIF effect (Storm, Bjork, & Bjork, 2008, 2012), such reduction has been attributed to a specific reexposure effect, which would not easily generalize to the exposure of new, related material, as it occurs in retroactive interference. On the basis of the noninhibitory account of RIF, the encoding of new, related information should increase interference for control items, practiced items, and unpracticed items. Although the relative increase in interference level may vary somewhat with item type and be slightly smaller for unpracticed items (which already suffer from a high level of interference) than for practiced and control items, such a pattern does not easily fit the observed parallel between practiced and unpracticed items' interference effects. To explain the observed parallel between practiced and unpracticed items' susceptibility to interference, both the inhibitory and noninhibitory accounts could be complemented by the assumption of retrieval-induced segregation processes.

Research on the testing effect has focused on the benefits of retrieval practice, emphasizing the strong overall recall improvements that result from retrieval practice, relative to restudy conditions (e.g., Roediger & Butler, 2011). Because the testing-effect literature typically does not distinguish between practiced and unpracticed items, research employing the retrieval-practice paradigm can supplement this literature by showing the extent to which the effects for practiced items generalize to unpracticed items. The present study serves this goal by showing that the effects of delay and interference on practiced items generalize to related unpracticed items. The present results thus not only increase our knowledge about retrieval-practice effects in the retrieval-practice paradigm, but may also help bridge the gap between retrieval-practice effects as they have been reported in the RIF literature and these effects as they have been reported in the testing-effect literature.

---

[4] The present indication for selective segregation processes shows a potential parallel to recent studies on selective (list-method) directed forgetting, in which it was shown that, in response to a "forget" cue, subjects can selectively forget previously studied irrelevant information while keeping previously studied relevant information in mind (Delaney, Nghiem, & Waldum, 2009; Kliegl, Pastötter, & Bäuml, 2013). Whether or not this parallel is mediated by similar processes may be an interesting issue for future research.

References

Abel, M., & Bäuml, K.-H. T. (2012). Retrieval-induced forgetting, delay, and sleep. *Memory, 20,* 420–428.

Anderson, J. R. (2000). *Learning and memory: An integrated approach*. New York, NY: Wiley.

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanism of forgetting. *Journal of Memory and Language, 49,* 415–445. doi:10.1016/j.jml.2003.08.006

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., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review, 102,* 68–100. doi:10.1037/0033-295X.102.1.68

Barnier, A., Hung, L., & Conway, M. (2004). Retrieval-induced forgetting of emotional and unemotional autobiographical memories. *Cognition & Emotion, 18,* 457–477. doi:10.1080/0269993034000392

Bäuml, K.-H. T., & Kliegl, O. (2013). The critical role of retrieval processes in release from proactive interference. *Journal of Memory and Language, 68,* 39–53.

Bäuml, K.-H. T., Pastötter, B., & Hanslmayr, S. (2010). Binding and inhibition in episodic memory—Cognitive, emotional, and neural processes. *Neuroscience and Biobehavioral Reviews, 34,* 1047–1054.

Bäuml, K.-H. T., & Samenieh, A. (2012). Selective memory retrieval can impair and improve retrieval of other memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 488–494. doi:10.1037/a0025683

Camp, G., Pecher, D., & Schmidt, H. G. (2007). No retrieval-induced forgetting using item-specific independent cues: Evidence against a general inhibitory account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 950–958. doi:10.1037/0278-7393.33.5.950

Chan, J. C. K. (2009). When does retrieval induce forgetting and when does it induce facilitation? Implications for retrieval inhibition, testing effect, and text processing. *Journal of Memory and Language, 61,* 153–170. doi:10.1016/j.jml.2009.04.004

Chan, J. C. K., McDermott, K. B., & Roediger, H. L., III. (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology. General, 135,* 533–571. doi:10.1037/0096-3445.135.4.553

Ciranni, M. A., & Shimamura, A. P. (1999). Retrieval-induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1403–1414. doi:10.1037/0278-7393.25.6.1403

Coman, A., Manier, D., & Hirst, W. (2009). Forgetting the unforgettable through conversation: Socially shared retrieval-induced forgetting of September 11 memories. *Psychological Science, 20,* 627–633.

Criss, A. H., & Shiffrin, R. M. (2004). Context noise and item noise jointly determine recognition memory: A comment on Dennis and Humphreys (2001). *Psychological Review, 111,* 800–807. doi:10.1037/0033-295X.111.3.800

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.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience, 11,* 114–126. doi:10.1038/nrn2762

Garcia-Bajos, E., Migueles, M., & Anderson, M. C. (2009). Script knowledge modulates retrieval-induced forgetting for eyewitness events. *Memory, 17,* 92–103.

Halamish, V., & Bjork, R. A. (2011). When does testing enhance retention? A distribution-based interpretation of retrieval as a memory modifier. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 801–812.

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

Jakab, E., & Raaijmakers, J. G. W. (2009). The role of item strength in retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 607–617. doi:10.1037/a0015264

Karpicke, J. D., & Roediger, H. L., III. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language, 57,* 151–162. doi:10.1016/j.jml.2006.09.004

Kliegl, O., Pastötter, B., & Bäuml, K.-H. T. (2013). List-method directed forgetting can be selective: Evidence from the 3-list and the 2-list tasks. *Memory & Cognition, 41,* 452–464. doi:10.3758/s13421-012-0264-7

Kornell, N., Bjork, R. A., & Garcia, M. A. (2011). Why tests appear to prevent forgetting: A distribution-based bifurcation model. *Journal of Memory and Language, 65,* 85–97.

Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting your native language: The role of retrieval-induced forgetting during second-language acquisition. *Psychological Science, 18,* 29–34. doi:10.1111/j.1467-9280.2007.01844.x

MacLeod, M. D., & Macrae, C. N. (2001). Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science, 12,* 148–152.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review, 39,* 352–370. doi:10.1037/h0069819

Migueles, M., & Garcia-Bajos, E. (2007). Selective retrieval and induced forgetting in eyewitness memory. *Applied Cognitive Psychology, 21,* 1157–1172.

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

Potts, R., & Shanks, D. R. (2012). Can testing immunize against interference? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 1780–1785.

Raaijmakers, J. G. W., & Jakab, E. (2013). Rethinking inhibition theory: On the problematic status of the inhibitory theory for forgetting. *Journal of Memory and Language, 68,* 98–122.

Racsmány, M., Conway, M. A., & Demeter, G. (2010). Consolidation of episodic memories during sleep: Long-term effects of retrieval practice. *Psychological Science, 21,* 80–85. doi:10.1177/0956797609354074

Roediger, H. L., III, & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15,* 20–27. doi:10.1016/j.tics.2010.09.003

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

Saunders, J., Fernandes, M., & Kosnes, L. (2009). Retrieval-induced forgetting and mental imagery. *Memory & Cognition, 37,* 819–828. doi:10.3758/MC.37.6.819

Spitzer, B., & Bäuml, K.-H. (2009). Retrieval-induced forgetting in a category recognition task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 286–291.

Storm, B. C., Bjork, E. L., & Bjork, R. A. (2008). Accelerated relearning after retrieval-induced forgetting: The benefit of being forgotten. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 230–236.

Storm, B. C., Bjork, E. L., & Bjork, R. A. (2012). On the durability of retrieval-induced forgetting. *Journal of Cognitive Psychology, 24,* 617–629.

Storm, B. C., Bjork, E. L., Bjork, R. A., & Nestojko, J. F. (2006). Is retrieval success a necessary condition for retrieval-induced forgetting? *Psychonomic Bulletin & Review, 13,* 1023–1027. doi:10.3758/BF03194002

Storm, B. C., & Levy, B. J. (2012). A progress report on the inhibitory account of retrieval-induced forgetting. *Memory & Cognition, 40,* 827–843. doi:10.3758/s13421-012-0211-7

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

Tulving, E., & Watkins, M. J. (1974). On negative transfer: Effects of testing one list on the recall of another. *Journal of Verbal Learning and Verbal Behavior, 13,* 181–193. doi:10.1016/S0022-5371(74)80043-5

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language, 50,* 289–335. doi:10.1016/j.jml.2003.10.003