Learning and Motivation 49 (2015) 23–30

Contents lists available at ScienceDirect

Learning and Motivation

journal homepage: www.elsevier.com/locate/l&m

# Retrieval practice does not safeguard memories from interference-based forgetting

Almut Hupbach *

*Lehigh University, USA*

ARTICLE INFO

*Article history:*
Received 7 October 2014
Received in revised form 23 January 2015

*Keywords:*
Testing effect
Retrieval practice
Retroactive interference
Memory reconsolidation

ABSTRACT

Retrieval enhances long-term retention. However, reactivation of a memory also renders it susceptible to modifications as shown by studies on memory reconsolidation. The present study explored whether retrieval diminishes or enhances subsequent retroactive interference (RI) and intrusions. Participants learned a list of objects. Two days later, they were either asked to recall the objects, given a subtle reminder, or were not reminded of the first learning session. Then, participants learned a second list of objects or performed a distractor task. After another two days, retention of List 1 was tested. Although retrieval enhanced List 1 memory, learning a second list impaired memory in all conditions. This shows that testing did not protect memory from RI. While a subtle reminder before List 2 learning caused List 2 items to later intrude into List 1 recall, very few such intrusions were observed in the testing and the no reminder conditions. The findings are discussed in reference to the reconsolidation account and the testing effect literature, and implications for educational practice are outlined.



Retrieval practice or testing is one of the most powerful memory enhancers. Testing that follows shortly after learning benefits long-term retention more than studying the to-be-remembered material again (Roediger & Karpicke, 2006a, 2006b). This effect has been shown using a variety of materials and paradigms, such as text passages (e.g., Roediger & Karpicke, 2006a), paired associates (Allen, Mahler, & Estes, 1969), general knowledge questions (McDaniel & Fisher, 1991), and word and picture lists (e.g., McDaniel & Masson, 1985; Wheeler & Roediger, 1992; Wheeler, Ewers, & Buonanno, 2003). Testing effects have been observed in traditional lab as well as educational settings (Grimaldi & Karpicke, 2015; Larsen, Butler, & Roediger, 2008; McDaniel, Anderson, Derbish, & Morrisette, 2007). Testing not only improves long-term retention, it also enhances subsequent encoding (Pastötter, Schicker, Niedernhuber, & Bäuml, 2011), protects memories from the buildup of proactive interference (PI; Nunes & Weinstein, 2012; Wahlheim, 2014), and reduces the probability that the tested items intrude into subsequently studied lists (Szpunar, McDermott, & Roediger, 2008; Weinstein, McDermott, & Szpunar, 2011). The reduced PI and intrusion rates are assumed to reflect enhanced list discriminability or improved within-list organization. Enhanced list discriminability in turn helps participants distinguish different sets or sources of information and allows them to circumscribe the search set during retrieval to the relevant list (e.g., Congleton & Rajaram, 2012; Halamish & Bjork, 2011; Szpunar et al., 2008).

* Correspondence to: Department of Psychology, Lehigh University, 17 Memorial Drive East, Bethlehem, PA 18015, USA.
*E-mail address:* hupbach@lehigh.edu

http://dx.doi.org/10.1016/j.lmot.2015.01.004

If testing increases list discriminability, then it should also protect the tested list(s) from RI and intrusions from material that is encoded after retrieval practice. However, testing also necessarily reactivates a memory, and according to the reconsolidation account reactivation re-introduces plasticity into the memory trace, making it especially vulnerable to modifications (e.g., Dudai, 2004; Nader, Schafe, & LeDoux, 2000; for a recent review, see e.g., Hupbach, Gomez, & Nadel, 2013). Increased vulnerability to modification would suggest increased rather than reduced RI and intrusions. The few studies addressing this issue have yielded mixed results, with some suggesting that retrieval practice diminishes RI (Halamish & Bjork, 2011; Potts & Shanks, 2012), and others showing that retrieval practice can exacerbate the potential negative effects of post-retrieval learning (e.g., Chan & LaPaglia, 2013; Chan, Thomas, & Bulevich, 2009; Walker, Brakefield, Hobson, & Stickgold, 2003).

Chan and colleagues (Chan & Langley, 2011; Chan et al., 2009; Thomas, Bulevich, & Chan, 2010) assessed the effects of testing on suggestibility in a misinformation paradigm. After watching a television episode, participants answered cued-recall questions about it (retrieval practice) or performed an unrelated distractor task. Then, all participants read a narrative, which summarized the video but also contained some misleading information. A final cued-recall test revealed that participants in the retrieval practice condition recalled more misleading details and fewer correct details than participants in the distractor condition; that is, retrieval increased the misinformation effect (retrieval-enhanced suggestibility, RES). Chan et al. (2009) discuss two mechanisms that can explain this finding. First, since testing can potentiate subsequent new learning (e.g., Izawa, 1967; Tulving & Watkins, 1974), initial testing might have improved encoding of the misinformation. Indeed, when a modified final test was used, which encouraged the recall of both the correct information and the misinformation, participants in the retrieval practice condition recalled more misinformation than participants in the distractor condition (Chan et al., 2009). Second, retrieval might have rendered the memory more susceptible to interference by misinformation, an explanation that is in line with the reconsolidation account. Indeed, Chan and LaPaglia (2013) found reduced recognition of the correct information when retrieval preceded the presentation of misinformation (cf. Walker et al., 2003 for a similar effect in procedural memory).

In contrast to Chan and colleagues' findings, a study by Potts and Shanks (2012) suggests that testing protects memories from the negative influences of post-retrieval encoding of related material. Potts and Shanks asked participants to learn English–Swahili word pairs (List 1, A–B). One day later, one group of participants took a cued recall test of List 1 (testing condition) immediately before learning English–Finnish word pairs with the same English cues as were used in List 1 (List 2, A–C). Additionally, several control groups were implemented: one group was tested on List 1 without learning a second list, one group learned List 2 without prior retrieval practice, and one group did not participate in this session at all. On the third day, all participants took a final cued-recall test of List 1. Although retrieval practice per se did not enhance List 1 memory (i.e., no testing effect in the groups that did not learn List 2), it protected memory from RI (see Halamish & Bjork, 2011 for a similar result in a one-session study). Crucial for assessing the reconsolidation account is the comparison between the groups that learned List 2 either after List 1 recall or without prior List 1 recall. Contrary to the predictions derived from the reconsolidation account, final List 1 recall was enhanced when retrieval of List 1 preceded learning of List 2.[1] While this clearly shows that testing counteracts RI, it would be premature to conclude that testing prevented the disruption of memory reconsolidation, because (a) retrieval practice without List 2 learning led to minimal forgetting between Day 2 and 3, while retrieval practice followed by List 2 learning led to significant memory decline, and (b) a reactivation condition that is independent from retrieval practice is missing. One could argue that repeating the cue words in List 2 likely reactivated memory for the original associations. It has been shown that the strength of reactivation (Detre, Natarajan, Gershman, & Norman, 2013) and the specific reminder structure (Forcato, Argibay, Pedreira, & Maldonado, 2009) determine whether or not a memory will be affected by post-reactivation procedures.

The current study re-evaluates the question of how testing affects RI and intrusions. It uses a reconsolidation paradigm (Hupbach, Gomez, Hardt, & Nadel, 2007; Hupbach, Hardt, Gomez, & Nadel, 2008; Hupbach, Gomez, & Nadel, 2009; Hupbach, Gomez, & Nadel, 2011) to assess how testing in comparison to other reactivation procedures affects declarative memory. This paradigm will allow for a direct evaluation of the hypotheses that testing makes declarative memories vulnerable to interference, or that testing protects memories from the potential negative effects of subsequently learned material, as suggested by the list-separation hypothesis (e.g., Congleton & Rajaram, 2012; Halamish & Bjork, 2011; Szpunar et al., 2008). This question has important practical implications. For instance, when students test their memory while preparing for an exam, will such testing increase or reduce interference and intrusions from information that is learned afterwards?

## Experiment

The aim of the present study is to assess how retrieval practice affects RI and intrusions in a reconsolidation paradigm. The paradigm consists of three sessions, separated by 48 h delays. In Session 1, participants learn a list of everyday objects. In Session 2, participants are either reminded of the first learning episode or not, and then learn a second list of objects. In prior studies, only subtle reminders were used (e.g., Hupbach et al., 2007): Participants in the reminder group returned to the same room, and they worked with the same experimenter as in Session 1. Additionally, they were asked to describe the general experimental procedure of Session 1. Participants in the no-reminder condition were seen in a different room

[1] Intrusion levels are not reported by Potts and Shanks (2012) but can be assumed to be generally low given that List 1 and List 2 contained quite different materials, i.e., Swahili vs. Finnish words.

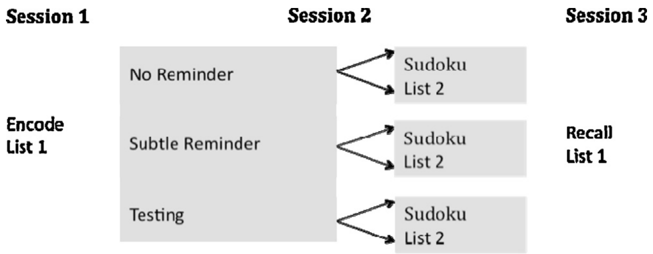


**Fig. 1.** Overview of the experimental design and the different reminder and post-reminder conditions.

and worked with a different experimenter, and were not asked to describe the procedure. In Session 3, participants were asked to recall List 1. Hupbach et al. (2007) found that participants in the reminder and no reminder groups did not differ in the number of List 1 objects recalled, but that participants in the reminder group intruded a significant number of List 2 objects into recall of List 1. This was rarely observed in the no-reminder group. Hupbach et al. interpreted this intrusion effect as a memory updating mechanism that incorporates new information into reactivated memories through reconsolidation processes. Specifically, the reminder "re-opened" List 1 memory such that new items from List 2 could be incorporated into List 1 memory (Hupbach et al., 2007, 2008, 2009, 2011).

The crucial question of the present study is how retrieval/testing in comparison to subtle reminder procedures will affect List 1 recall and intrusions from List 2 into List 1. Based on the idea that testing promotes list separation and enhances list discriminability (e.g., Congleton & Rajaram, 2012; Szpunar et al., 2008), testing List 1 memory before learning the second list of objects should reduce intrusions from List 2 into List 1. Additionally, based on Potts and Shanks (2012), we should see reduced RI in the form of enhanced memory in the testing condition in comparison to a subtle reminder or no reminder condition. However, if testing reactivates memories making them especially vulnerable to RI and intrusions, we should find impaired List 1 recall and increased intrusions from List 2 into List 1 in the testing condition, especially in comparison to a no reminder condition, and potentially also in comparison to a subtle reminder condition.

## Methods

### *Design and participants*

All participants took part in three separate sessions. Participants in the different groups were treated identically in Sessions 1 and 3. In Session 2, two independent variables were manipulated between subjects, resulting in a 3 (reminder) × 2 (post-reminder task) factorial design. Participants were either asked to recall List 1 (testing condition), to describe the experimental procedure of Session 1 (subtle reminder condition), or were not reminded of the previous session (no reminder). Additionally, participants either learned a second list of objects (interference condition) or performed an unrelated distractor task (solving Sudoku puzzles; control condition). Seventy-four undergraduate students (46 females, 28 males) from the University of Arizona and Lehigh University participated in the experiment. They received course credit or financial compensation for participation. Subjects were randomly assigned to the six conditions ($N = 12$ for the no-reminder and subtle reminder conditions, and $N = 13$ for the testing conditions).

### *Materials*

List 1 and List 2 materials each consisted of 20 unrelated common objects that were used in previous reconsolidation studies (Hupbach et al., 2007, 2008, 2009). List 1: balloon, bow, calculator, toy car, crayon, cup, dice, feather, flashlight, flower, glue, key, sock, sponge, spoon, sunglasses, teabag, tennis ball, toothbrush, whistle. List 2: apple, band-aid, battery, book, cassette tape, cellular phone, comb, dollar bill, elephant, envelope, paper clip, puzzle piece, rock, shovel, straw, thread, tissue, toy pot, watch, zipper.

### *Procedure*

Fig. 1 depicts the experimental design. The three sessions took place on Monday, Wednesday, and Friday of the same week, i.e., sessions were separated by 48 h. Participants were informed at the beginning of Session 1 that they would have to memorize different lists of objects on the different days. Participants were tested one at a time. The procedure follows closely the one used in Hupbach et al. (2007).

*In Session 1*, the experimenter pulled out one object at a time from a bag in random order and placed it in a distinctive basket. Participants were asked to name each object and remember it for a later memory test. After all 20 objects (List 1) had been put into the basket, the experimenter hid the basket and asked the participants to remember as many objects as possible. This procedure was repeated until the participants remembered at least 17 of the 20 objects or until a maximum of four learning trials was reached.

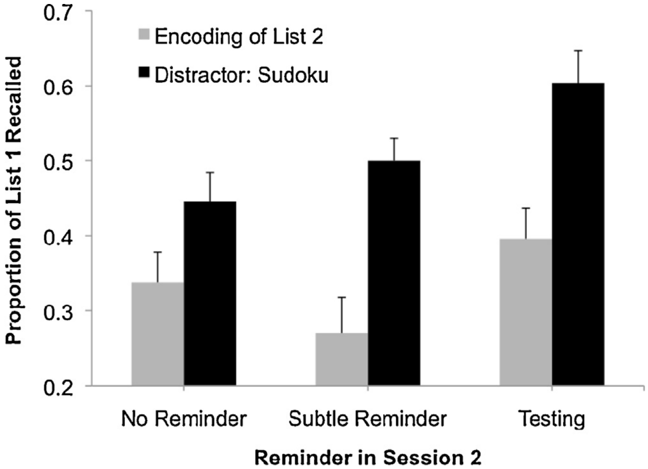


**Fig. 2.** Mean proportion of objects recalled from List 1. In Session 2, participants were either asked to describe the List 1 learning episode (subtle reminder) or they were asked to recall List 1 (testing condition). Another group of participants was not reminded of the first learning episode (no reminder). Additionally, participants either learned a second list of objects in Session 2 or solved Sudoku puzzles. Error bars represent standard errors of means.

Forty-eight hours after Session 1, *Session 2* was administered. Participants in the subtle-reminder condition were asked, “Do you remember this basket and what we did with it?” Participants were encouraged to describe the procedure but were stopped if they started to recall specific items. Participants in the testing condition were not asked to describe the general procedure but were instead asked to recall the objects from List 1, and the experimenter noted the remembered objects. For the subtle reminder condition, Session 2 took place in the same room and with the same experimenter as Session 1. For participants in the testing condition and the no-reminder group, participants worked with a different experimenter in a different room. Half of the participants in all conditions learned a second list of 20 objects (List 2). A slightly different presentation method than in Session 1 was used, such that the procedure itself would not act as a reminder (cf. Hupbach et al., 2007). All objects were placed in front of the participants, who were asked to name each object. After all objects had been named, participants were given another 30 s to study them. Then, the experimenter removed the objects and asked the participants to recall as many of the objects as possible. The study procedure was repeated until participants recalled at least 17 objects, or for a maximum of four learning trials. The other half of the participants in each condition solved Sudoku puzzles for 10 min (this matched the average time it took the other participants to learn List 2).

*In Session 3*, participants were instructed to recall as many objects as possible from List 1 only, and the experimenter noted the remembered objects. In all conditions, the experimenter from Session 1 administered the test, and the test took place in the same room as Session 1 encoding.

## Results

*Learning List 1 and List 2*: The number of trials (1–4) it took participants to recall a minimum of 17 objects was recorded. Participants who recalled less than 17 objects in the final (fourth) trial received a score of 5. One-way ANOVAs with Reminder (No Reminder, Subtle Reminder, Testing) as the between-subjects variables were carried out separately for List 1 and List 2 encoding. For List 1, the ANOVA revealed a marginally significant effect [$F(2, 71) = 3.02$, $MSE = 1.06$, $p = .06$, $\eta_P^2 = .08$]. Post hoc tests (Tukey) showed that the Testing group ($M = 3.96$, $SD = .87$) needed significantly more trials to learn List 1 than the Subtle Reminder group ($M = 3.25$, $SD = 1.07$, $p = .05$). No other comparisons were significant ($p \geq .33$). More importantly, the number of trials to learn List 2 did not differ between groups [$F(2, 71) = 2.06$, $MSE = 1.49$, $p = .14$]. This finding shows that there were no group differences in exposure to List 2, which could have contributed to the interfering effects of List 2 onto List 1 memory.

*Recall List 1*: The mean proportions of objects recalled from List 1 in the different experimental groups are displayed in Fig. 2.

The number of objects recalled from List 1 was analyzed with a 3 × 2 factorial ANOVA with Reminder (No Reminder, Subtle Reminder, Testing) and Post-Reminder Task (Sudoku vs. List 2) as between-subjects variables. Both factors significantly affected recall performance, Task: $F(1, 68) = 29.96$, $MSE = 8.15$, $p < .01$, $\eta_P^2 = .31$; Reminder: $F(2, 68) = 5.16$, $MSE = 8.15$, $p < .01$, $\eta_P^2 = .13$. The interaction was not significant ($p = .30$). The Task effect shows that learning a second list in Session 2 significantly diminished recall of List 1 ($M = .33$, $SD = .16$) in comparison to solving Sudoku puzzles ($M = .52$, $SD = .15$). The Reminder effect was further analyzed with planned contrasts comparing the Testing condition against the Subtle Reminder and No-Reminder conditions, in order to assess the testing effect. These contrasts showed that recall was significantly higher in the Testing condition ($M = .50$, $SD = .18$) than in the other two conditions (Subtle Reminder: $M = .39$, $SD = .18$, $p < .01$; No Reminder: $M = .39$, $SD = .14$, $p = .01$).

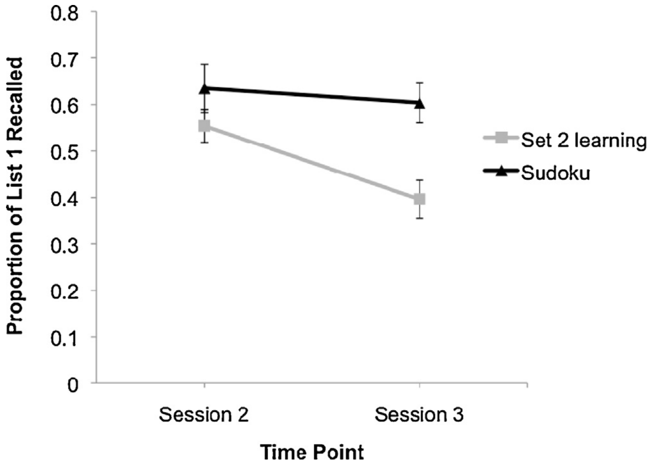


**Fig. 3.** Mean proportion of objects recalled in the testing condition in Sessions 2 and 3 in relation to the task performed after recall in Session 2. Error bars represent standard errors of means.

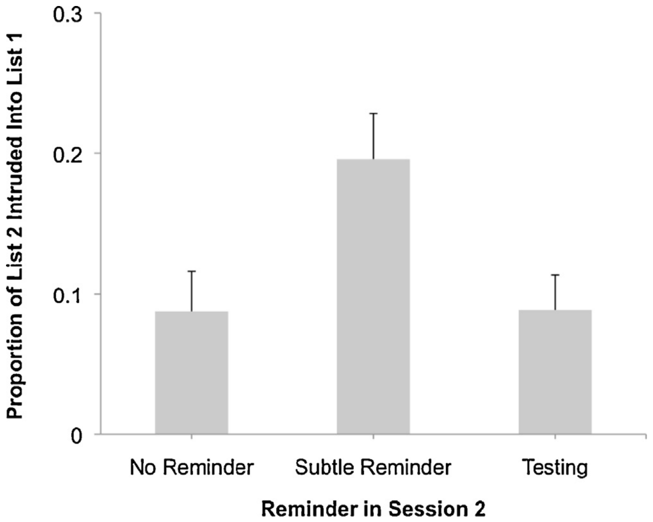


**Fig. 4.** Mean proportion of objects from List 2 that were falsely recalled as List 1 objects in the no-reminder, subtle-reminder, and testing condition. Depicted are only the three experimental groups that learned a second list in Session 2. Error bars represent standard errors of means.

In order to specifically assess how the task in Session 2 (Sudoku vs. List 2) influenced the testing effect, List 1 recall in Session 2 was compared with List 1 recall in Session 3 in the Testing condition only[2] (see Fig. 3).

A mixed factorial ANOVA with Task (Sudoku vs. List 2 learning) as the between-subjects variable and Session (2 vs. 3) as the within-subjects factor revealed significance of both main effects, Session: $F(1, 24) = 14.51$, $MSE = 3.18$, $p < .01$, $\eta_P{}^2 = .38$; Task: $F(1, 24) = 6.67$, $MSE = 16.22$, $p = .02$, $\eta_P{}^2 = .22$, and a significant interaction, $F(1, 24) = 6.58$, $MSE = 3.18$, $p = .02$, $\eta_P{}^2 = .22$. An analysis of the simple effects showed that recall performance did not differ between the two groups in Session 2, $F(1, 24) = 1.67$, $MSE = 10.15$, $p = .21$. Importantly, participants who had learned a second list after recall in Session 2 recalled significantly fewer items in Session 3 than participants who had solved Sudoku puzzles, $F(1, 24) = 12.13$, $MSE = 9.24$, $p < .0$, $\eta_P{}^2 = .34$]. At the same time, there was no significant difference (i.e., no forgetting) between recall in Sessions 2 and 3 in the group that solved Sudoku puzzles [$F < 1$], but there was a significant decline in performance in the group that learned a second list, $F(1, 24) = 20.31$, $MSE = 3.18$, $p < .01$, $\eta_P{}^2 = .46$, evidencing RI.

*Intrusions*: The analysis of intrusion rates was restricted to groups that learned a second list in Session 2. The mean proportions of objects falsely recalled from List 2 (intrusions) are displayed in Fig. 4.

The number of intrusions was analyzed with a one-way ANOVA with Reminder (No Reminder, Subtle Reminder, Testing) as the between-subjects variable. The type of reminder significantly affected intrusion rates, $F(2, 37) = 4.66$, $MSE = 4.04$, $p = .02$,

[2] The other two groups, i.e., the subtle reminder group and the no reminder group were not asked to recall List 1 in Session 2, and could therefore not be included in this analysis.

$\eta_P^2$ = .22. Post hoc comparisons (Tukey) showed that the Subtle Reminder group showed significantly more intrusions than the No Reminder ($p$ = 0.03), and the Testing group ($p$ = .03). The No Reminder group and the Testing group did not differ in their intrusion rates ($p$ = 1.0).

## Discussion

The present study assessed the competing hypotheses that testing returns memories to a vulnerable state in which they are especially susceptible to RI and intrusions as suggested by the reconsolidation account, or that testing protects memories from the potential negative effects of subsequently learned material as suggested by the list-separation hypothesis (e.g., Congleton & Rajaram, 2012; Halamish & Bjork, 2011; Szpunar et al., 2008). Before evaluating the hypotheses, it is important to note that our paradigm was generally sensitive to the testing effect: testing List 1 in Session 2 enhanced List 1 memory in comparison to the subtle reminder and the no reminder condition. Critically, this effect was independent of the post-reminder tasks. The current study therefore allows for an independent assessment of memory strengthening (testing) and memory-impairing effects (interference).

*Recall of List 1*: Learning a second list of objects in comparison to solving Sudoku puzzles in Session 2 impaired final recall of List 1 in all reminder conditions (testing, subtle reminder, no reminder). Thus, learning a new list caused RI whether a reminder was given or not, clearly showing that testing did not prevent RI. Comparing the retention of List 1 measured in Sessions 2 and 3 in the interference and control conditions (Fig. 3) provides further evidence that learning of List 2 negatively affected memory for List 1 in the testing condition. While there was little forgetting in the control group, there was substantial forgetting in the interference group. Significant RI in the testing condition, however, does not mean that testing made memories *especially* vulnerable to RI, because RI was observed in all conditions, even in the no-reminder condition.[3] Rather, it implies that RI manifests predominantly as a retrieval impairment, reducing List 1 access in Session 3. Testing List 1 in Session 2 does not attenuate this type of RI. This finding differs from Potts and Shanks' (2012) study, in which testing was reported to protect memories from RI: In their study, learning a second list did not affect List 1 memory in the testing condition but impaired memory when no test preceded List 2 learning. Potts and Shanks (2012) argue that testing strengthened List 1 memory and thus prevented RI. Although this type of strengthening did not increase recall of List 1 items when no interference was present (i.e., no general testing effect, which might be due to initial overtraining), it improved recall when interference from List 2 made List 1 retrieval more difficult (see Halamish & Bjork, 2011 for a similar argument). However, and similar to what was observed in the present study, Potts and Shanks (2012) did find evidence for RI in the testing conditions when they compared recall performances between Sessions 2 and 3 (see also Halamish & Bjork, 2011, Exp. 3). While there was no forgetting in the group that did not encode an interfering list, there was a significant decline in memory performance in the group that learned List 2, which shows testing did not fully insulate List 1 memory from interference. Hence, the degree of RI in the testing condition might be influenced by the strength of memory prior to testing, with protective effects being dependent on initial overtraining, which was the case in Potts and Shanks' study but not in the current study. Taken together, retrieval practice seems less effective in protecting memories from RI than protecting memories from PI (Szpunar et al., 2008).

Could the diminished recall of List 1 in the interference conditions reflect a process by which participants withhold List 1 items because they are uncertain whether the items belonged to List 1 or List 2 (Koriat & Goldsmith, 1996)? Based on prior studies using the same paradigm, such withholding seems rather unlikely. Specifically, when asked to recall List 2, subjects rarely intrude List 1 items (Hupbach et al., 2007). More importantly, in a recognition/source test, where subjects are forced to respond to every item and can therefore not withhold any information, subjects frequently misattribute List 2 to List 1, but rarely List 1 to List 2 (Hupbach et al., 2009). Taken together, prior studies show that List 2 items are incorporated into List 1 representations but not vice versa. Therefore, it seems unlikely that RI effects in the present study reflect a withholding of List 1 items.

*Intrusions from List 2 into List 1*: While a subtle reminder caused a substantial number of intrusions (replicating Hupbach et al., 2007), very few intrusions were observed in the testing condition and in the no reminder condition. We have previously argued that such intrusions reflect a memory reconsolidation process by which pre-existing memories are updated with new information (cf. Hupbach et al., 2007). Before further discussing the diminished updating in the testing condition, it is important to state that the low level of intrusions in this condition is *not* due to the lack of contextual reminders in Session 2. In the testing condition, encoding of List 2 took place in a different room, and memory updating in the object-learning paradigm has been shown to crucially depend on contextual reminders (Hupbach et al., 2008). To test whether intrusions in the testing condition would re-emerge if contextual reminders were present during List 2 learning, an additional 11 participants were tested in the condition that combined testing with subsequent List 2 learning, but this time, the second session took place in the same room as Sessions 1 and 3. All previously reported effects for this condition were replicated: Recall of List 1 in Session 2 ($M$ = .58, $SD$ = .11) and Session 3 ($M$ = .37, $SD$ = .11) were comparable to the testing condition that took place in a different context. Importantly, intrusions did *not* re-emerge ($M$ = .04, $SD$ = .05); on the contrary, they were even lower than in the testing condition that was carried out in a different context ($M$ = .10, $SD$ = .08; see Fig. 3). The finding

[3] Although more powerful designs might reveal larger RI in the testing condition in comparison to the no-reminder condition.

that testing diminished intrusions even when contextual reminders were present is strong evidence that testing indeed averts updating.

This finding contradicts Chan et al.'s (2009) observation that testing increases intrusions in the form of misinformation (see also Chan & Langley, 2011; Thomas et al., 2010) and decreases later recognition of correct information (Chan & LaPaglia, 2013). One factor that might explain the different outcomes concerns the specific relationship between the original memory and the potentially interfering information. The present study focused on unspecific interference, that is, interference from material that does not share common cues with the tested material and therefore does not directly compete with it. In Chan and colleagues' studies the misinformation directly contradicted some of the target information. In fact, recognition of the original information was not impaired when retrieval of the original information was followed by encoding of an unrelated story instead of a summary of the original episode containing misinformation (Chan & LaPaglia, 2013). Specific interference in combination with enhanced encoding of the misinformation can be assumed to undermine the list separation that is triggered by testing (e.g., Szpunar et al., 2008), causing enhanced rather than reduced intrusions by misinformation and long-lasting memory impairments of the reactivated memory (Chan & LaPaglia, 2013). The specificity of interference is not the only factor determining whether new information replaces the original memory. Another factor that might modulate impairments in paradigms with specific interference concerns the detection of change (Wahlheim & Jacoby, 2013). Specifically, if subjects notice the discrepancy between the original and new information, and are able to recall the original information concurrently with encoding the new information, updating might be diminished. This is because subjects will likely discount the misleading information or remember both pieces of information (Chan & LaPaglia, 2013). Similarly, in PI paradigms, detecting that the target associated with a cue has changed (D in A–D) *and* recollecting the original target (B in A–B), lead to integration of the competing information into a unitary representation that preserves the temporal order of events (A–B–D). As a result, proactive facilitation rather than PI occurs, that is, enhanced rather than impaired memory for A–D (Wahlheim & Jacoby, 2013; see also Hintzman, 2004, 2011 for the effects of recursive reminding on memory for temporal order and frequency). Testing A–B before learning A–D leads to increased detection of change and recollection, and thus integration of competing information (Wahlheim, 2014).

Taken together, testing affected RI and intrusions differently in the current study. While testing did not prevent RI, it decreased the intrusion of List 2 items during List 1 recall in comparison to a subtle reminder condition. While RI and intrusion rates are often positively correlated, there are exceptions to this pattern. For instance, Melton and Irwin (1940) showed that extensive training on interfering material increases RI, but decreases intrusions. In the present paradigm, testing might have differentially affected processes at retrieval, i.e., RI, and processes during re-storage/reconsolidation, i.e., intrusions. Specifically, the finding that RI was observed in all conditions, i.e., regardless of memory reactivation, contradicts the idea that reactivation promotes RI.[4] Instead, learning a second list in Session 2 makes final retrieval generally more difficult. This will impair List 1 recall, unless the to-be retrieved memory is initially overlearned and then further strengthened by testing (Halamish & Bjork, 2011; Potts & Shanks, 2012), which was not the case in the present study. Prior studies using the object-learning paradigm have shown that intrusions in this paradigm, on the other hand, reflect a reconsolidation process that is triggered by memory reactivation in Session 2 (Hupbach et al., 2007, 2008, 2009, 2011). Such reactivation occurred in both the reminder and the testing conditions, but intrusions were only observed in the reminder condition. This is because the fate of a reactivated memory is not only determined by the specifics of the post-reminder treatment; it is further influenced by the strength of the reactivated memory (Wang, de Oliveira, & Nader, 2009) and by the specifics of the reactivation procedure (Forcato et al., 2009). Testing not only reactivated the memory, it also strengthened it (Roediger & Karpicke, 2006a) and increased the within-list organization (Congleton & Rajaram, 2012; Szpunar et al., 2008). These processes prevented the updating of List 1 memory with List 2 items.

## Conclusion

The current study shows that test taking strengthens memories, and while that strengthening does not always protect memories from interference-based forgetting, it prevents new information from intruding into the tested material. This is relevant for situations in which students attempt to acquire knowledge of two similar topics. Testing memory for one topic (e.g., Ancient Greece) before learning about another topic (e.g., Ancient Rome) will prevent the new facts (e.g., gods Romans worshiped such as Jupiter, Neptune) from becoming mixed up with the tested material (Greek gods such as Zeus, Poseidon).

## References

Allen, G. A., Mahler, W. A., & Estes, W. K. (1969). Effects of recall tests on long-term retention of paired associates. *Journal of Verbal Learning and Verbal Behavior*, *8*, 463–470.

Chan, J. C. K., & Langley, M. (2011). Paradoxical effects of testing: Retrieval enhances both accurate recall and suggestibility in eyewitnesses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 248–255.

Chan, J. C. K., & LaPaglia, J. A. (2013). Impairing existing declarative memory in humans by disrupting reconsolidation. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 9309–9313.

[4] Although numerically, the impairment caused by List 2 learning was twice as large in the testing condition in comparison to the no-reminder condition. Future studies with increased statistical power might detect a significant difference between the two conditions, and thus test-enhanced RI.

Chan, J. C. K., Thomas, A. K., & Bulevich, J. B. (2009). Recalling a witnessed event increases eyewitness suggestibility: The reversed testing effect. *Psychological Science*, *20*, 66–73.

Congleton, A. R., & Rajaram, S. (2012). The origin of the interaction between learning history and delay in the testing effect: The roles of processing and retrieval organization. *Memory & Cognition*, *40*, 528–539.

Detre, G. J., Natarajan, A., Gershman, S. J., & Norman, K. A. (2013). Moderate levels of activation lead to forgetting in the think/no-think paradigm. *Neuropsychologia*, *51*, 2371–2388.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annual Review of Psychology*, *55*, 51–86.

Forcato, C., Argibay, P. F., Pedreira, M. E., & Maldonado, H. (2009). Human reconsolidation does not always occur when a memory is retrieved: The relevance of the reminder structure. *Neurobiology of Learning and Memory*, *91*, 50–57.

Grimaldi, P. J., & Karpicke, J. D. (2014). Guided retrieval practice of educational materials using automated scoring. *Journal of Educational Psychology*, *106*, 58–68.

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.

Hintzman, D. L. (2004). Judgment of frequency vs. recognition confidence: Repetition and recursive reminding. *Memory & Cognition*, *32*, 336–350.

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "coordinates of truth". *Perspectives on Psychological Science*, *6*, 253–271.

Hupbach, A., Gomez, R., Hardt, O., & Nadel, L. (2007). Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learning & Memory*, *14*, 47–53.

Hupbach, A., Hardt, O., Gomez, R., & Nadel, L. (2008). The dynamics of memory: Context-dependent updating. *Learning & Memory*, *15*, 574–579.

Hupbach, A., Gomez, R., & Nadel, L. (2009). Episodic memory reconsolidation: Updating or source confusion? *Memory*, *17*, 502–510.

Hupbach, A., Gomez, R., & Nadel, L. (2011). Episodic memory updating: The role of context familiarity. *Psychonomic Bulletin & Review*, *18*, 787–797.

Hupbach, A., Gomez, R., & Nadel, L. (2013). Episodic memory reconsolidation: An update. In C. Alberini (Ed.), *Memory reconsolidation* (pp. 233–246). Elsevier.

Izawa, C. (1967). Function of test trials in paired-associate learning. *Journal of Experimental Psychology*, *75*, 194–209.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490–517.

Larsen, D. P., Butler, A. C., & Roediger, H. L. (2008). Test-enhanced learning in medical education. *Medical Education*, *42*, 959–966.

McDaniel, M. A., & Fisher, R. P. (1991). Tests and test feedback as learning sources. *Contemporary Educational Psychology*, *16*, 192–201.

McDaniel, M. A., & Masson, M. E. J. (1985). Altering memory representations through retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 371–385.

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, *19*, 494–513.

Melton, A. W., & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *The American Journal of Psychology*, *LIII*(2), 611–641.

Nader, K., Schafe, G. E., & LeDoux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature*, *406*, 722–726.

Nunes, L. D., & Weinstein, Y. (2012). Testing improves true recall and protects against the buildup of proactive interference without increasing false recall. *Memory*, *20*, 138–154.

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.

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

Roediger, H. L., & Karpicke, J. D. (2006a). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249–255.

Roediger, H. L., & Karpicke, J. D. (2006b). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*, 181–210.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008). Testing during study insulates against the build-up of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1392–1399.

Thomas, A. K., Bulevich, J. B., & Chan, J. C. K. (2010). Testing promotes eyewitness accuracy with a warning – Implications for retrieval enhanced suggestibility. *Journal of Memory and Language*, *63*, 149–157.

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.

Wahlheim, C. N. (2014). Testing can counteract proactive interference by integrating competing information. *Memory & Cognition*, advance online publication.

Wahlheim, C. N., & Jacoby, L. L. (2013). Remembering change: The critical role of recursive reminding in proactive effects of memory. *Memory & Cognition*, *41*, 1–15.

Walker, M. P., Brakefield, T., Hobson, J. A., & Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature*, *425*, 616–620.

Wang, S. H., de Oliveira, A. L., & Nader, K. (2009). Cellular and systems mechanisms of memory strength as a constraint on auditory fear reconsolidation. *Nature Neuroscience*, *12*, 905–912.

Weinstein, Y., McDermott, K. B., & Szpunar, K. K. (2011). Testing during study insulates against proactive interference in face-name learning. *Psychonomic Bulletin & Review*, *18*, 518–523.

Wheeler, M. A., & Roediger, H. L. (1992). Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychological Science*, *3*, 240–245.

Wheeler, M. A., Ewers, M., & Buonanno, J. (2003). Different rates of forgetting following study versus test trials. *Memory*, *11*, 571–580.