Research

Intrusions in episodic memory: reconsolidation or interference?

Angela Klingmüller,[1] Jeremy B. Caplan,[1,2] and Tobias Sommer[1]

[1]Department for Systems Neuroscience, Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; [2]Psychology Department, Neuroscience and Mental Health Institute, University of Alberta, Edmonton, AB T6G 2R3, Canada

It would be profoundly important if reconsolidation research in animals and other memory domains generalized to human episodic memory. A 3-d-list-discrimination procedure, based on free recall of objects, with a contextual reminder cue (the testing room), has been thought to demonstrate reconsolidation of human episodic memory (as noted in a previous study). Our goal was to replicate the central result, a high intrusion rate during recall of the target list, and evaluate the reconsolidation account relative to an alternative account, based on state-dependent learning and interference. First, replication was not straightforward (Experiment 1). Second, using a very unique, highly salient context (Experiment 2), the method produced a qualitative replication, but it was small in magnitude. A critical assumption of the reconsolidation account, that the target list is reactivated and destabilized during re-exposure to the study context, was not supported (Experiment 3). Although troubling for the reconsolidation account, the findings can be easily accommodated by an alternative account that does not assume additional neurobiological processes underlying the destabilization of consolidated memories, instead explaining intrusion rates simply in terms of well-established cognitive effects, such as item-to-context binding and interference during retrieval.

According to consolidation theories, newly encoded memories are initially in an active, labile state and can be disrupted. Over time, memory traces become stabilized, i.e., consolidated (McKenzie and Eichenbaum 2011). Reconsolidation theory, on the other hand, proposes that a consolidated memory re-enters a labile state when it is reactivated and is thus again susceptible to modification before being reconsolidated (Hardt et al. 2010; Schiller and Phelps 2011). Evidence for reconsolidation theory has derived mainly from animal fear conditioning studies where manipulations such as injecting a protein synthesis inhibitor into the amygdala after reactivating a consolidated fear memory can lead to a selective amnesia (Sara 2008; Schiller and Phelps 2011).

It would be important to know if reconsolidation is such a general phenomenon that it could even apply to human episodic memory. In human procedural and fear memory studies, the post-retrieval learning of interfering material has been proposed as a noninvasive method to modify the reactivated memory before its reconsolidation (Walker et al. 2003; Schiller et al. 2010). To study the reconsolidation hypothesis in human episodic memory, a similar procedure was designed, also based on learning interfering information after reactivation of a previously learned memory (Hupbach et al. 2007). This procedure was used in a subsequent series of studies and was adapted by others to further characterize reconsolidation in episodic memory (Hupbach et al. 2008, 2009, 2011; Wichert et al. 2011, 2013a; Jones et al. 2012, 2015; Potts and Shanks 2012; Dongaonkar et al. 2013; Gershman et al. 2013; Wichert et al. 2013b; Hupbach 2015).

In this "3-d-list-interference paradigm," on Day 1, participants learn a list of objects (List 1, Table 1) in one testing room, A, (Fig. 1). On Day 2, participants are given a "reminder cue" intended to reactivate List 1-memory. The full reminder cue involves being asked in the same room (A) by the same experimenter to think about the general procedure of Day 1. Immediately following the reminder, a list of different objects is learned (List 2).

On Day 3, participants are asked, again in room A, to recall the List-1 objects. The control group learns List 2 on Day 2 in a different testing room (room B) with a different experimenter and is not reminded about Day 1.

The central result is that participants in the Full Reminder group not only recall List-1 objects on Day 3, but also intrude objects from List-2. This increased List-2 intrusion rate in the Full Reminder group was taken as evidence in favor of the reconsolidation hypothesis. In particular, it was argued that List 1-memory upon being reactivated on Day 2 was then susceptible to be "updated" by insertion of the List-2 items prior to being reconsolidated. It was later shown that the same spatial context, room A, on Day 2 is critical as reminder to produce List-2 intrusions on Day 3, whereas the identity of the experimenter and reminder question had no influence (Hupbach et al. 2008). In some of the adaptations of the design, a decrease in List 1-memory in the Full Reminder group was observed instead of an increased List 2-intrusion rate that was then interpreted as reflecting reconsolidation (Wichert et al. 2011, 2013a,b).

The results of these studies are widely accepted as evidence that reconsolidation occurs in human episodic memory (Schiller and Phelps 2011; Schwabe et al. 2014), However, plausible alternative accounts have been proposed to explain the pattern of results. These accounts rely only on well-established cognitive processes, like interference during memory retrieval, and not on the assumption of novel neurobiological processes, like destabilization and reconsolidation (Riccio et al. 2006; Sederberg et al. 2011). The most straightforward interference-based account is a special case of classic state-dependent learning which refers to the observation that memory is superior when the study and test contexts are similar (DeVietti and Larson 1971; Unsworth

Corresponding author: tsommer@uke.de

Article is online at http://www.learnmem.org/cgi/doi/10.1101/lm.045047.117.

Table 1. List-1 and -2 objects

List-1	List-2
Balloon	Apple
Toy car	Battery
Flower	Book
Clothes hanger	Paper clip
Crayon	CD
Feather	Egg cup
Globe	Elephant
Light bulb	Euro bill
Cable	Clinical thermometer
Plush toy	Thread
Glue	Set square
Spoon	Waist belt
Magnifying glass	Cream
Playing card	Cellular phone
Whistle	Comb
Key	Chain
Sponge	Candle
Sock	Nutcracker
Sunglasses	Playmobil puppet
Star	Puzzle piece
Flashlight	Zipper
Calculator	Shovel
Cup	Swimmies
Teabag	Rock
Tennis ball	Straw
Clothespin	Tissue
Dice	Toy pot
Toothbrush	Watch
Folding rule	Envelope
Rubber duck	Hot-water bottle

et al. 2012). State-dependent learning was proposed to explain the blocking of reconsolidation by electroconvulsive shocks and protein synthesis inhibitors in animals (Riccio et al. 2002, 2006; Gisquet-Verrier and Riccio 2012; Gisquet-Verrier et al. 2015), but the increased List-2 intrusion rate observed in the 3-d-list-interference paradigm can be explained by the same logic. List-1 and -2 objects are associated with the same context (room A) only in the Full Reminder group. On Day 3, Room A is an effective contextual cue for List-1, but also for List-2 objects. Thus, both List-1 and List-2 objects will be retrieved on Day 3, explaining why List-2 items intrude when the participant attempts to recall List 1. When the room is changed on Day 2, List-2 objects are associated only with room B. On Day 3, Room A will only cue retrieval of List-1 objects, resulting in few List-2 intrusions. Moreover, because the Full Reminder group retrieves List-2 objects along with List-1 objects on Day 3, this increases the effective pool of response candidates. This increased competition for retrieval reduces the probability of List-1 objects being retrieved, for the Full Reminder group compared with the Control group. Thus, the interference-based account predicts both more List 2-intrusions and fewer List-1 recalls in the Full Reminder group.

The essence of the difference between the reconsolidation and interference account for the observed pattern of results in the 3-d-list-interference paradigm is that the reconsolidation account assumes that on Day 3, only an updated version of the List-1 memory trace exists (containing both List-1 and List-2 items). In contrast, the interference account assumes the original List-1 memory trace still exists, in addition to memory traces encoded during Day 2. Intrusions on Day 3 are due to confusion between those different memory traces (interference) that occurs during the recall process. The interference account seems more parsimonious because it does not require additional neurobiological and cognitive processes (undoing of consolidation-related neurophysiological and mnestic changes). It is based only on well-described mnestic phenomena, such as item–context association

and memory cueing by context reinstatement (Unsworth et al. 2012). The interference account is also more explanatory because it assumes that the same principles explain intrusions regardless of time scale, that is, in the 3-d-list-interference as well as in older single-session experiments on associative interference that observed a similar effect of changing the learning context (Nagge 1935; Bilodeau and Schlosberg 1951; Greenspoon and Ranyard 1957; Dallett and Wilcox 1968). Parsimony, therefore, favors the interference account unless results are provided that could not be readily explained by known phenomena.

The first goal of the current study was to replicate previous findings using the 3-d-list-interference paradigm. Simple replication is important not only due to the relevance and implications of the reconsolidation account, but also because of the failure to replicate two widely cited studies on reconsolidation of human procedural and fear memories—also using the interference approach to modify the reactivated memory before its reconsolidation (Soeter and Kindt 2011; Hardwicke et al. 2016). The second aim was to further specify the boundary conditions in which increased List-2 intrusions occur, and evaluate whether these are consistent with the reconsolidation and the interference account.

General methods

A total of 180 volunteers, mostly students, participated in the study (mean age 25.6 ± 4.5 yr). They were recruited via a job exchange website at the University of Hamburg, and were compensated with €30. Experimental procedures were approved by the Internal Review Board of the German Psychological Association.

Twenty participants were randomly assigned to each experimental condition, and the mean age of participants did not differ between conditions. We decided to increase the sample size from
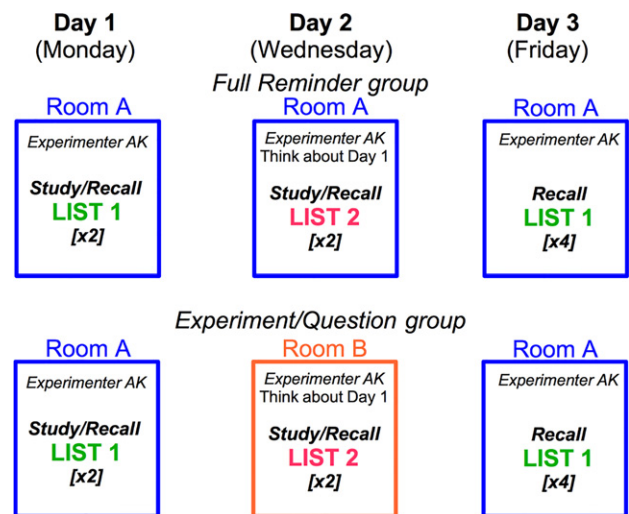

Figure 1. Schematic depiction of the procedure used in Experiments 1 and 2. Experimenter AK refers to the first author, who tested all participants in all the experiments reported here. The Full Reminder group is modeled on the same-named group in Hupbach et al. (2007); on Day 2, participants are tested by the same experimenter, in the same room as Day 1 and are asked to describe the procedures used on Day 1, but are interrupted if they start to recall any list objects. The Experimenter/Question group is modeled on the same-named group in Hupbach et al. (2008); on Day 2, they are tested by the same experimenter and are asked to describe the procedures on Day 1, but are tested in a different room. Both components of the Full Reminder cue did not result in List-2 intrusions on Day 3 in Hupbach et al. (2008) similar to the No reminder group.

Table 2. Memory performance

	Day 1 List-1 recall [mean (SD)]	Day 2 List-2 recall [mean (SD)]	Day 3	
			List-1 recall [mean (SD)]	List-1 intrusions [mean (SD)]
Experiment 1				
Full reminder replication	90.5% (5.4)	89.4% (4.8)	60.6% (19.5)	10.3% (14.4)
Full reminder	70.0% (11.8)	70.8% (11.5)	34.9% (21.3)	7.4% (8.9)
Experimenter/question	69.7% (9.7)	74.0% (9.6)	40.7% (14.4)	7.1% (8.0)
Experiment 2				
Basement full reminder	73.8% (12.8)	74.3% (10.0)	35.1% (15.5)	13.8% (13.0)
Basement experimenter/question	74.2% (7.5)	72.5% (9.2)	40.6% (12.5)	4.7% (6.8)
Experiment 3				
Full reminder–LPS	73.7% (9.6)		56.1% (14.5)	
No reminder	72.8% (12.8)		55.9% (15.0)	
Basement full reminder–LPS	71.5% (11.8)		50.8% (12.3)	
Basement no reminder	72.0% (10.5)		48.8% (10.7)	

12 to 20 participants per condition compared with previous studies using the 3-d-list-interference paradigm based on the following power calculations. We calculated the power of finding a significant ($\alpha = 0.05$) difference in List 2-intrusion rate between the Full Reminder and the Experimenter/Question groups with $N = 12$ per condition based on previous publications (Hupbach et al. 2007, 2008) using G*Power as follows (Faul et al. 2009). We extracted the standard errors from the corresponding figures (the numbers were not reported) and computed the standard deviations (14.2 for the Full Reminder group, Hupbach et al. 2007, 8.9. for the Experimenter/Question group, 2008). Based on these values and the published mean intrusion rates, an effect size of Cohen's $d = 1.502$ and a power $(1 - \beta) = 0.97$ were obtained, assuming a directed, one-tailed hypothesis (more intrusions in the Full Reminder group). Similarly, we obtained an effect size of Cohen's $d = 0.92$ for the difference in List-1 memory between the Full Reminder and the Experiment/Question groups (based on from the figures extracted standard deviations of 5.79 and 3.85, respectively) and a power $(1 - \beta) = 0.70$. Although the power with 12 participants was therefore sufficiently high to detect a difference between List-2 intrusion rate, we increased the sample size to $N = 20$ per condition because this increased the power of detecting a difference in List-1 memory substantially, to $(1 - \beta) = 0.89$. Moreover, this seems more similar to comparable free-recall experiments and it reduces the risk of sampling error, increasing the generalizability of findings.

Memory performance on all 3 d, for all experimental groups, is listed in Table 2. When ANOVAs were computed Greenhouse–Geisser corrections for violations of sphericity were used where necessary.

Some of our central results are nonsignificant group differences. Because it is not possible to accept the Null hypothesis (H_0) using classical statistics, we evaluated the nonsignificant differences using Bayesian statistics. We computed Bayesian t-tests (Cauchy prior, width 1.0) to quantify the evidence in favor of the H_0 with the mean List-1 recall or mean List-2 intrusions across the four recall trials on Day 3 as the dependent variable (JASP Team, 2016, Version 0.8.0.0). According to convention, Bayes factors BF_{01} ranging from 1 to 3 provide anecdotal, from 3 to 10, moderate, and above 10, strong evidence in favor of the H_0.

Experiment 1

The first experiment was an attempt to replicate the previously observed pattern of results.

Methods

In Experiment 1 three groups were included, the "Full Reminder Replication group," the "Full Reminder group," and the "Experimenter Question group," in order to replicate the increased List-2 intrusion rate only in the latter group. These groups were in addition statistically compared with the corresponding groups of Hupbach et al. (2007, 2008).

The "Full Reminder Replication group" was an exact replication of the "Full Reminder group" in Experiment 1 of Hupbach et al. (2007). The materials and procedure of Hupbach et al. (2007, 2008) were replicated as closely as possible (A. Hupbach, pers. comm.). Day 1 (Monday) took place in a behavioral testing room at the institute (room A) that had some distinguishing features (relatively small room with two doors, a single window within in a pitched roof area, and a prominent sink). The experimenter (AK, first author of the study) pulled out one of the objects from List 1 (Table 1, first 20 items) from a bag, asked the participant to name it, and then placed it in a blue plastic basket. After all 20 objects were studied, the participant tried to recall as many objects as possible, in any order. This procedure was repeated until the learning criterion of ≥ 17 recalled items but to a maximum of four study–test cycles.

On Day 2 (Wednesday of the same week), in the same room as on Day 1, the same experimenter reminded participants of the Day-1 procedure by showing participants the blue basket and asking them to briefly describe the general procedure, but stopping them from recalling specific items. Participants then learned a different set of 20 objects (List 2, Table 1) using a different procedure: all 20 objects were placed on the table in an arbitrary layout in front of the participants, who named each object and were then given 30 sec to study them. The objects were removed, and participants recalled as many objects as possible. This procedure was repeated until the same learning criterion of ≥ 17 recalled items was reached, to a maximum of four study–test cycles.

On Day 3 (Friday of the same week), in the same room as on Days 1 and 2, participants were asked to recall as many objects from List 1. When participants indicated they did not remember more items, there was a 30-sec break filled with an unrelated conversation, and then the next recall round started, for a total of four rounds of free recall.

The participants in our "Full Reminder Replication group" remembered substantially more objects than the "Full Reminder group" in Hupbach et al. (2007), which we revisit in the discussion. To obtain comparable List-1 memory performance on Day 3, we made the following modifications: (i) The length of both lists was increased to 30 objects (Table 1). (ii) The maximum number of learning trials was reduced to 2, with a criterion of ≥ 25

items. These modifications are retained throughout the rest of the experiments we report. With these modifications, we added the "Full Reminder group," where the procedure was otherwise identical to the Full Reminder Replication group of Experiment 1 and the Full Reminder group of Hupbach et al. (2007).

As mentioned Hupbach et al. (2008) found that using Room A again on Day 2 was necessary and sufficient to increase List-2 intrusions, whereas the experimenter and the reminder question had no effect. Therefore, we added a control group, which we call the "Experimenter/Question" group, modeled on the Experimenter/Question group of Hupbach et al. (experiment 2, 2008). This group learned List 2 in another behavioral laboratory (room B), but with the same experimenter (AK, first author of the study) and reminder question. This laboratory was clearly distinct from the first one (large room in another floor of the building, higher level ceiling, a stretcher leaning on the wall, a large window). Otherwise, the procedure for the Full Reminder and the Experimenter/Question groups were identical.

Results

Performance on day 1 and 2: acquisition of list 1 and 2

On day 1, our Full Reminder Replication group reached criterion in 2.45 ± 0.51 (M ± SD) trials, significantly fewer than the 3.11 ± 1.26 trials needed by the participants from Experiment 1 of Hupbach et al. (2007), $t_{(54)} = 2.23$, $P < 0.015$. Our Full Reminder Replication group also took fewer trials to reach criterion for List 2 on Day 2 (1.75 ± 0.79 and 3.79 ± 1.38 trials, respectively, $t_{(54)} = 6.076$, $P = 0.0000006$).

After the modification of the experimental design (30 objects each in List, only two learning trials), trial-number ranged only from 1 to 2 so this factor has little sensitivity. Therefore, we focus our analyses on the percentage of recalled objects on the last recall trial. Participants in our Full Reminder and Experimenter/Question groups recalled equivalent proportions of List-1 objects on Day 1 and List-2 objects on Day 2 ($t_{(38)} = 0.09$, $P = 0.92$; $t_{(38)} = 1.01$, $P = 0.35$, Table 2). Thus, the groups were matched on List-1 memory on Day 1 and List-2 memory on Day 2.

Performance on day 3

Percentage of objects correctly recalled from List 1 and List 2-intrusions (averaged over all four recall trials) are displayed in Figure 2, alongside the corresponding results of Hupbach et al. (2007, 2008) to enable direct comparison. We conducted direct statistical comparisons to their results based on the published means and standard deviations that we extracted from the figures.

List–1 recall. First, we compared our Full Reminder Replication group with the Full Reminder group of Hupbach et al. (2007), which revealed that our sample recalled significantly more objects ($t_{(30)} = 3.37$, $P = 0.002$).

Next, we tested whether the modification of the experimental design (30 objects and only 2 learning trials) in our Full Reminder group affected List 1-memory and compared it with the Full Reminder group of Hupbach et al. (2007), which resulted in a nonsignificant difference ($t_{(30)} = 0.18$, $P = 0.52$). Thus, the increase in list length and decrease in trials to criterion had the desired effect of equating final-recall accuracy for List-1 items. A replication of central result of Hupbach et al. would be a List-2 intrusion rate comparable to theirs, which we test in the next section.

However, a reduction in List-1 memory in the Full Reminder group has also been argued to reflect reconsolidation in some of the studies that adapted the 3-d-list-interference paradigm.

Although across experiments, data from Hupbach et al. (2007, 2008) revealed a significant difference in List-1 memory between the Full Reminder and the Experimenter/Question groups ($t_{(23)} = 2.25$, $P = 0.017$, Cohen's $d = 0.92$). We compared our corresponding groups (both with 30 objects) with one another. List 1 recall was analyzed with a 2×4 mixed ANOVA with Group (Full Reminder versus Experimenter/Question) as between-subjects factor and Trial (1–4) as a within-subjects factor. The effect of Trial reached significance ($F_{(2.6,97.7)} = 11.37$, $P = 0.000006$), indicating that participants recalled more List-1 objects with each successive trial. However, neither the main effect of Group ($F_{(1,38)} = 1.01$, $P = 0.32$) nor the interaction of Group × Trial ($F_{(2.6,97.7)} = 0.69$, $P = 0.54$) reached significance, suggesting no differences in List-1 memory between the two groups. In a Bayesian t-test with mean List-1 recall across the four trials as dependent variable $BF_{01} = 1.68$ provided anecdotal evidence for the H_0, that is, equal List-1 memory in the Full Reminder and Experimenter/Question groups. Although we could not statistically test for an interaction of the effects of Group (Full Reminder versus Experimenter/Question) and sample (Hupbach's versus ours), the significantly worse List-1 memory in the Full Reminder group in Hupbach's but not our study suggests that learning the List-2 objects in a different context affected List-1 memory in our sample less than in sample of Hupbach et al.

Intrusions from list 2. At first, we compared the List-2 intrusion rate (mean across the four recall trials) in our Full Reminder Replication (20 object, 2 learning trials) and Full Reminder (30 objects, 2 learning trials) groups to Hupbach et al.'s (2007) Full Reminder group. Our Full Reminder Replication group had significantly fewer List-2 intrusion than the Full Reminder group of Hupbach et al. (2007) ($t_{(30)} = 2.58$, $P = 0.0036$, Cohen's $d = 0.96$). Also, the List-2 intrusion rate in our Full Reminder group was significantly lower than in Hupbach et al.'s Full Reminder group ($t_{(30)} = 6.52$, $P < 0.0001$, Cohen's $d = 1.38$).

To compare our three groups to one another, the List-2 intrusion rate was analyzed with a 3×4 mixed ANOVA with Group (Full Reminder Replication versus Full Reminder versus Experimenter/Question) as a between-subjects factor and Trial (1–4) as a within-subjects factor. None of the effects reached significance (all $Fs < 0.8$). In other words, we could not replicate the greater List-2 intrusion rate in the Full Reminder compared
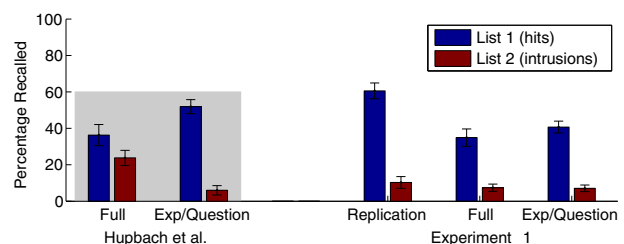

Figure 2. Percentage recalled (averaged over all four recall trials) from the target list (List 1; hits) and from the interference list (List 2; intrusions, averaged over all four recall trials), on Day 3, in Experiment 1. To enable direct comparison, the mean values from Hupbach et al. (Full: 2007; Exp/Question: 2008) are replotted here, within the gray-shaded region (*left*), alongside the means and standard errors (error bars) from the three groups in Experiment 1 (*right*). Group labels are as follows: Full—Full Reminder group (same room, experimenter, and reminder question). Exp/Question—Experimenter/Question group (different room but same experimenter and reminder question). Replication—Full Reminder Replication—attempt to replicate Hupbach et al. (2007). The Hupbach et al. and Full Reminder Replication groups studied lists of 20 objects; whereas the Full Reminder and Experimenter/Question groups of Experiment 1 studied lists of 30 objects.

with the Experimenter/Question group, which is the central result of Hupbach et al's studies (2007, 2008) and is supposed to reflect reconsolidation of List-1 memory.

Because of the critical relevance of comparing the List-2 intrusion rates for this replication attempt we computed in addition a more sensitive two-sample t-test using the mean List-2 intrusion rate in our Full Reminder and Experimenter/Question groups across the four recall trials as dependent variable ($t_{(38)} = 0.12$, $P = 0.90$, Cohen's $d = 0.039$). In a Bayesian t-test, $BF_{01} = 3.91$, favors the H_0, no difference in List-2 intrusion rate between groups.

Discussion experiment 1

The first important finding of Experiment 1 is that the current participant sample had substantially better memory performance when exactly the same number of objects was used as in the previously published studies, which might be due to differences in the populations. In particular, the samples of Hupbach et al. were undergraduate students who participated for course credit, whereas our samples were presumably older, more experienced participants (mean age 25.6 yr), who participated for monetary compensation. Such a motivational difference between participants, even of the same population, dependent on whether they participated for course credit or monetary compensation, was recently reported (Nicholls et al. 2015).

Critically, even with a modified procedure that equated List-1 memory, Experiment 1 still failed to replicate the greater List-2 intrusion rate in the Full Reminder group during final recall of List-1 on Day 3 (Hupbach et al. 2007, 2008). In fact, the relative List-2 intrusion rates in our sample were invariant to both difficulty (Full Reminder Replication versus Full Reminder) and the critical reconsolidation manipulation (Full Reminder versus Experimenter/Question). This replication failure using exactly the same experimental setting and a larger sample size seems nontrivial given the effect size and power of the previously observed effect. This suggests there may be important boundary conditions to identify the previously reliably reported greater List-2 intrusion rate in the Full Reminder groups (Hupbach et al. 2009, 2007, 2008; Gershman et al. 2013; young adults, Jones et al. 2015), which echoes a failure to replicate the elevated intrusion rate in older adults (Jones et al. 2015).

Previous studies that used modifications of the 3-d-list-interference paradigm (such as recognition instead of free-recall tests) also did not observe an increase in List-2 intrusions, but instead interpreted a decrease in List-1 memory in some of the conditions as reflecting reconsolidation (Wichert et al. 2011, 2013a,b). Hupbach et al. (2007, 2008) observed—although across experiments—lower List-1 memory in their Full Reminder compared with the Experimenter/Question group. In contrast, our participants' List-1 memory was, if at all, only very subtly and not significantly affected by learning List-2 in the same or another context.

Taken together, we failed to replicate the influence of the List-2 learning context on List-2 intrusion rate, which has been interpreted as disrupting reconsolidation. In addition, we observed only a much more subtle influence of List-2 learning context on List-1 memory, compared with Hupbach et al. (2007, 2008). It should be noted that both replication failures occurred despite our increased power due to substantially larger sample sizes. Interestingly, in the domains of procedural and fear memory, there are similar failures to replicate the modification of memories by post-retrieval learning of interfering information, previously also argued to index reconsolidation (Soeter and Kindt 2011; Hardwicke et al. 2016).

A possible explanation for our replication failures might be that the motivation and/or general experience of the volunteers

with testing contexts are boundary conditions for the previously observed increase in List-2 intrusions and decrease in List-1 memory when List 2 was learned in the same spatial context.

Experiment 2

The goal of Experiment 2 was to test one hypothetical boundary condition. In one prior failure to replicate the higher intrusion rate, participants were children, tested at home (Hupbach et al. 2011). However, when children were tested in a room less familiar than their home, increased List-2 intrusions were again observed. The authors suggested that a familiar room is not effective as an incidental retrieval cue to reactivate and destabilize the List-1 memory on Day 2 which enables its updating with List-2 objects before reconsolidation. Our participants very likely had experience with oral exams in rooms similar to our testing rooms, so our testing rooms may have been too familiar, even unremarkable. However, it should be emphasized that the participants had never visited these rooms before and that they were clearly distinct, as described in the General methods.

Interestingly, such an effect of the saliency of the testing room would be also consistent with earlier studies on interference theory (Nagge 1935; Bilodeau and Schlosberg 1951; Greenspoon and Ranyard 1957; Dallett and Wilcox 1968). In these studies as well two lists were learned in the same or different testing rooms followed by a final memory test for the list 1. The greater amount of retroactive interference when both lists were learned in the same testing room was taken as evidence for the theory of associative interference. The main difference of those studies to the 3-d-list-interference paradigm is that those experiments took place during a single session. Those studies emphasized that two unusual (e.g., "a dingy storeroom, partially filled with old apparatus") but also distinct testing rooms are necessary in order to observe List 2-interference in a final List-1 memory test dependent on whether List 2 is learned in the same or a different room. In an interference account, more salient spatial contexts result in stronger object–context feature binding or for participants to think to use the spatial context as a retrieval cue (Unsworth et al. 2012).

Taken together, based on both the reconsolidation and the interference account, one can hypothesize that the standard testing rooms used in Experiment 1 were not salient enough for our participants to result in List-2 intrusions. In particular, room A on Day 2 may thus not have reinstated List-1 (reconsolidation account), or participants did not use Room A as a retrieval cue on Day 3 (interference account). We tested two more groups using the identical procedure as in the Full Reminder and Experimenter/Question groups of Experiment 1, but in a very unfamiliar, unusual, and salient room.

Methods

In Experiment 2, four groups were included, the "Full Reminder group" and "Experimenter/Question group" of Experiment 1, and two new groups, the "Full Reminder basement group" and the Experimenter/Question Basement group" in order to test the effect of context in a 2 (Full Reminder versus Experimenter/Question) × 2 (behavioral laboratory versus basement) factorial design.

The two new groups were tested using the identical procedure as the Full Reminder and Experimenter/Question groups of Experiment 1, but in a smaller room in the basement of the institute. The institute is located in a red brick stone building that was established early last century. The rooms in the basement are used as storage space and for the heating unit of the building. Because few people need to go into the basement and rarely for very long, little attention is paid to aesthetics. The basement has its own
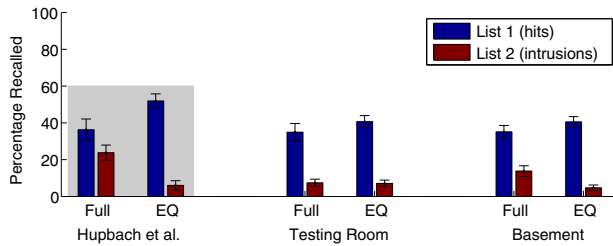
Figure 3. Percentage recalled from the target list (List 1; hits) and from the interference list (List 2; intrusions) on Day 3, in Experiment 2. To enable direct comparison, the mean values from Hupbach et al. (2007, 2008) are replotted here, within the gray-shaded region (*left*), and two groups of Experiment 1 are plotted in the *middle* bars. Error bars plot standard error of the mean across participants. Group labels are as follows: Full Reminder groups (same room, experimenter, and reminder question). EQ Experiment/Question groups (different room, but same experimenter and reminder question). For Experiment 1 (*middle* bars), rooms A and B were both testing rooms. For Experiment 2 (*right* bars), room A was the basement room and room B (EQ group only) was a testing room.

small side entrance which leads to a staircase. From the small hall downstairs, a heavy metal fire door on one side leads to a very short corridor that, in turn, leads to another heavy metal fire door. The 10-m long narrow corridor behind the second metal door is relatively dark, had not been painted for many years and has three closed doors on both sides. At the end is a small dusty room without a door, with broken brick stones in one corner and spider webs. This room was used on all days in the Full Reminder group and on Days 1 and 3 for the Experimenter/ Question group. Day 2 for the Experimenter/Question group took place in the behavioral testing room used in the previous two experiments (previously room A).

Results

The two groups tested in the basement were compared with the Full Reminder and Experimenter/Question groups of Experiment 1. Anecdotally, many participants spontaneously commented on the highly unusual testing room in the basement and commiserated with the experimenter.

Performance on days 1 and 2: acquisition of lists 1 and 2

List-1 memory on Day 1 and List-2 memory on Day 2 were analyzed with 2×2 ANOVAs with Room (behavioral laboratory versus basement) and Reminder (Full Reminder versus Experimenter/Question) as between-subjects factors. The main effect of Room produced the smallest—though not significant— P-value ($F_{(1,76)} = 2.37$, $P = 0.13$), suggesting more List-1 objects may have been recalled in the basement. All other effects were clearly nonsignificant (all Ps > 0.3 Table 2). Thus, the groups were matched on List-1 memory on Day 1 and List-2 memory on Day 2.

Performance on day 3

The mean percent of items recalled from List 1 and the mean percent of items falsely recalled from List 2 are displayed in Figure 3, along with the Results from Hupbach et al. (2007, 2008) and the Full Reminder and Experimenter/Question groups of our Experiment 1, to enable direct visual comparison.

List–1 recall. List-1 memory was analyzed with a $2 \times 2 \times 4$ mixed ANOVA with Context (behavioral laboratory versus basement) and Reminder (Full Reminder versus Experimenter/Question) as

between-subjects factors and Trial (1–4) as a within-subjects factor. The main effect of Trial ($F_{(2.2,165.9)} = 17.50$, $P<0.0001$) indicated that accuracy increased over free-recall trials, as before. The main effect of Reminder did not reach significance ($F_{(1,76)} = 2.4$, $P = 0.13$); the direction of the effect was that List-1 recall was nominally better when List-2 was learned on Day 2 in a different room (Experimenter/Question groups). All other main effects and interactions were far from significant (Ps > 0.6).

Because a potential reduction in List-1 memory in the Full Reminder group is another central result of the 3-d-list-interference paradigm we compared mean List-1 recall across the four recall trials only in the Basement groups to gain a higher sensitivity. The t-test revealed a nonsignificant difference ($t_{(38)} = 1.23$, $P = 0.23$, Cohen's $d = 0.39$). In the Bayesian t-test, $BF_{01} = 1.28$ provided anecdotal evidence for H_0, that is, equivalent List-1 memory in the Full Reminder and the Experimenter/Question groups in the Basement.

List–2 intrusions. In an ANOVA with the same design as for List-1 recall, the main effect of Reminder was significant ($F_{(1,76)} = 4.9$, $P < 0.029$), with fewer List-2 intrusions for the Experimenter/ Question than the Full Reminder groups. The interaction Reminder × Context reached significance ($F_{(1,76)} = 4.2$, $P < 0.042$), indicating more intrusions by the Full Reminder than the Experimenter/Question groups in the basement (Tukey's HSD $P < 0.017$). All other main effects and interactions were not significant; the smallest of those P-values was for Trial × Reminder ($F_{(1.8,136.6)} = 1.89$, $P = 0.16$).

Contrasting only the mean List-2 intrusion rate in the basement groups confirmed the expected significant difference ($t_{(38)} = 2.77$, $P = 0.009$, Cohen's $d = 0.89$). The power to detect a significant difference ($\alpha = 0.05$) was computed: $1 - \beta = 0.87$. In a Bayesian t-test, $BF_{10} = 10.48$, providing strong evidence for the alternative hypothesis, that is, more List-2 intrusions in the Full Reminder than the Experimenter/Question group only in the basement.

Having replicated the general pattern of Hupbach et al. (2007, 2008)—a significant increase in List-2 intrusions when List-2 was learned in the same very salient context as List 1—we contrasted the intrusion rates in the Full Reminder groups between both samples. There were still significantly fewer intrusions in our sample, compared with the sample of Hupbach et al. (2007) ($t_{(30)} = 1.99$, $P = 0.027$, Cohen's $d = 0.74$).

Discussion of experiment 2

Using a highly unusual, if not bizarre, testing room, we were able to replicate, qualitatively, the previously observed increased List-2 intrusions in the 3-d-list-interference paradigm on Day 3 when List-2 was learned in the same context (Hupbach et al. 2007, 2008). The most important finding of Experiment 2 is the relevance of context saliency or uniqueness to produce intrusions in adult volunteers, but also the degree of saliency, in particular an even bizarre environment, needed to produce this still relatively low List-2 intrusion rate compared with earlier reports (Hupbach et al. 2007). This boundary condition seems therefore rather extreme in particular given the effect size and power of the originally reported effect.

From the perspective of the reconsolidation account, the results suggest that the behavioral testing rooms of the institute were not effective enough for our study population to trigger incidental reactivation and destabilization of the memory. From the perspective of interference-theory accounts, the results suggest that the behavioral testing rooms were not salient or unique enough to be a dominant component of participants' spatio-

temporal context or to be associated with the experimental list items (Unsworth et al. 2012). However, although List-1 recall of our Basement Full Reminder group was similar to the Full Reminder group of Hupbach et al. (2007), the List-2 intrusion rate was still substantially lower. It seems difficult to imagine that the remaining difference in intrusion rates between the two samples is still caused by differences in room familiarity, as the basement was rather extreme in its unusual features. Therefore, motivational differences between the two populations might again be the reason for the lower intrusion rate in current sample.

Whereas we were able to replicate the List-2 intrusions at least qualitatively, List-1 memory of our participants on Day 3 was still not significantly affected by learning List 2 in the same very salient compared with a different context. However, this effect approached significance across Experiments 1 and 2 (main effect of Reminder $P = 0.13$), suggesting only a quantitative, not qualitative, difference to the samples of Hupbach et al (2007, 2008).

Taken together, the degree of saliency that was needed to produce List-2 intrusions in our sample is—presumably together with motivation—an important boundary condition of this effect in adults. However, this result is not diagnostic of the competing accounts, as both accounts of increased List-2 intrusions rely on a stronger binding of objects to the very salient context. In particular, the interference account proposes stronger binding of List-2 objects to the salient context. More List-2 objects are then erroneously reactivated when participants intentionally reinstate the context during the final List-1 recall on Day 3. In contrast, the reconsolidation account relies on a stronger binding of List-1 to the salient context which results in a more effective, incidental cueing of List-1 memory on Day 2 and hence a stronger reactivation and destabilization. Importantly, only the reconsolidation account proposes an incidental List-1 memory reactivation on Day 2 cued by re-exposure to the salient context. Therefore, in Experiment 3, we aimed to directly test for reactivation of the List-1 memory on Day 2 in the salient context, which is a critical prerequisite of the reconsolidation account.

Experiment 3

A central tenet of reconsolidation theory is that the reactivation of consolidated memories transfers them back into a labile state. Successful reactivation of the original memory is therefore a critical prerequisite of reconsolidation to occur. In previous studies using the 3-d-list-interference paradigm, successful reactivation of List-1 memory was not observed but only hypothesized to be the cause of List-2 intrusions (Hupbach et al. 2007). In other words, the increased List-2 intrusions were taken as evidence for reconsolidation and at the same time the only observable evidence for an incidental List-1 memory reactivation—assuming the reconsolidation account is correct. This reasoning seems circular. If it were found, the detection of List-1 memory reactivation using another outcome measure would be compelling evidence in favor of the reconsolidation account. Therefore, in Experiment 3, we independently tested whether List-1 memory was indeed reactivated on Day-2 by re-exposure to the salient context.

Reconsolidation theory proposes that reactivation of memories results in their stabilization when it is not followed by learning of interfering information or disruption (Nadel et al. 2012). Such a beneficial effect of memory reactivation has also been observed in studies on the effects of successive testing, even at short timescales for which reconsolidation is not possible (Roediger and Butler 2010). Consistent with this proposal, the positive effect of reactivating a consolidated episodic memory was recently confirmed in a study on the reconsolidation hypothesis that used electroconvulsive therapy (ECT) to disrupt reconsolidation

(Kroes et al. 2014). In particular, memory reactivation was not followed by disruption in a control group which resulted in a selective improvement of those memories compared with non-reactivated memories.

In Experiment 3, we tested whether re-exposure, only to the very salient basement, but not unremarkable testing room on Day 2, and being reminded by the same experimenter to the general procedure of Day 1 by viewing the blue basket and being given the reminder question, results in a stabilized memory and a testing effect—better List-1 memory on Day 3—when it is not followed by any interfering List-2 learning. Whereas the reconsolidation account relies on reactivation of List-1 memory specifically in the unusual context, the interference account does not make specific assumptions about List-1 reactivation.

Methods

In Experiment 3, four groups were included, the "No Reminder–No List-2 group," the "Full Reminder–No List-2 group," the "No Reminder–No List-2 basement group," and the "Full Reminder–No List-2 Basement group" in order to test the effect of context in a 2 (Full Reminder versus Experimenter/Question) × 2 (behavioral laboratory versus basement) factorial design.

Four new groups were tested, two in the behavioral testing room and two in the basement, using the identical procedure as before on Day 1 and 3. However, the No Reminder–No List-2 groups did not come to the institute on Day 2 (and did not learn List 2). The Full Reminder–No List-2 groups came to the institute to the same context where they had learned List 1. They were reminded as with the Full Reminder groups of Experiments 1 and 2, not only in terms of the room, but were also by being shown the blue basket and asked the reminder question by the same experimenter. However, after the reminder question, they did not learn List-2, but performed a German fluid intelligence test (LPS 4, a subtest of the "Leistungsprüfsystem"; Horn, 1983), consisting of identifying the nonmatching digit or letter in a row by inferring an underlying rule, for 8 min.

Results

Performance on day 1: acquisition of list 1

List 1 memory was analyzed with a 2 × 2 ANOVA with Room (behavioral laboratory versus basement) and Reminder (Full Reminder–LPS versus No Reminder) as between-subjects factors that revealed no significant main effect nor interaction (all Ps > 0.6).

Performance on day 3: recall of list 1

The mean percentage of items recalled from List 1 are displayed in Figure 4. The significant main effect of Trial ($F_{(2.5, 190.4)} = 34.54$, $P = 0.0000007$) indicated again that recall increased over trials. The significant main effect of Room ($F_{(1,76)} = 5.2063$, $P = 0.025$) suggested more List-1 objects recalled on Day 3 in the behavioral laboratory. Critically, neither the main effect of Reminder ($F_{(1,76)} = 0.11$, $P = 0.74$) nor the interaction of Reminder with Room ($F(1,76) = 0.11$, $P = 0.74$, coincidentally identical to the main effect of Reminder) were significant. All interactions with Trial were nonsignificant (Ps > 0.4).

To increase the sensitivity to detect a difference, we compared mean recall rate only of the Full Reminder–No List 2 and the No Reminder groups in the basement ($t_{(38)} = 0.54$; $P = 0.59$, Cohen's $d = 0,17$). In a Bayesian t-test, $BF_{01} = 2.78$ provided anecdotal evidence for the null hypothesis, that is, equal List-1 recall in the No Reminder and Full Reminder–No List 2 groups in the basement.
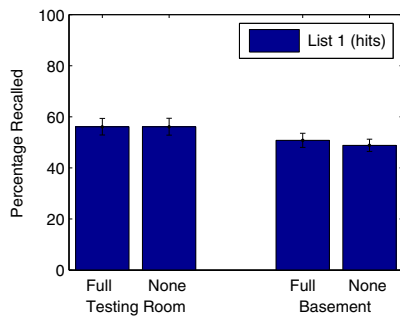
Figure 4. Percentage recalled from the target list (List 1; hits) on Day 3, in Experiment 3. Error bars plot standard error of the mean across participants. Group labels are as follows: Full–Full Reminder groups (same room, experimenter, and reminder question); these participants also performed the LPS task. None–No Reminder groups, who did not return to the institute on Day 2. Note that in this experiment, no group learned an interfering list (List 2) on Day 2. For the Testing Room groups, all sessions were conducted in the testing room. For the Basement groups, all sessions were conducted in the basement room.

Discussion of experiment 3

According to reconsolidation theory, reactivation is the critical prerequisite for the destabilization of memories, because only a reactivated memory trace becomes destabilized and modifiable before its reconsolidation. We found no evidence for an effect of the Full Reminder compared with the No Reminder on List-1 memory on Day 3—even in the salient context. In other words, we found no evidence that the Full Reminder triggered partial reactivation of the List-1 memory that would have resulted in better recall on Day 3 compared with the No Reminder, similar to the control group in a recent study that disrupted reconsolidation using ECT (Kroes et al. 2014). One might argue that List-1 memory was so weakly activated on Day 2 that it did not result in any observable consequences for the final List-1 recall on Day 3. However, a necessary precondition for reconsolidation is a strong reactivation of the first memory, strong enough also to influence final List-1 recall. Indeed, animal studies showed only when the reactivation of a memory is strong and the particular memory trace is dominant over interfering memories, does reconsolidation occur (Riccio et al. 2002; Alberini 2007; Mamiya et al. 2009; McKenzie and Eichenbaum 2011).

Taken together, the results of Experiment 3 are therefore not easily compatible with the reconsolidation account of the increase in List-2 intrusion rate in the salient context in Experiment 2. In other words, Experiment 3 tested a prediction that is central to reconsolidation theory (Nadel et al. 2012), but failed to find support for it. This suggests that the central finding, List-2 intrusions into recall of List 1, can occur without the necessary preconditions for reconsolidation. In contrast, the interference account of the List-2 intrusions in Experiment 2 is not challenged by these results because it does not assume any reactivation of List-1 memory in the Full Reminder condition. In other words, Experiment 3 resolved the interpretive ambiguity of the results of Experiment 2 in favor of the interference account.

General discussion

The most important finding of the current study is that we could not replicate (with even greater statistical power due to larger sample sizes) that the re-exposure to a rather unremarkable context in terms of a standard testing room results in higher levels of List-2 intrusions (Experiment 1). However, re-exposure to a highly unusual, salient, context did lead to an increase in List-2 intrusions

(Experiment 2). This finding would be consistent with reflecting stronger List-2 item-to-context binding in an interference account but also with a stronger reactivated, and hence more labile, List-1 memory as proposed by the reconsolidation account. Experiment 3 resolved this ambiguity, because only the reconsolidation account critically relies on a reactivation of List-1 memory on Day 2 when participants are re-exposed to the initial learning context. Contrary to a recent study blocking reconsolidation using ECT (Kroes et al. 2014), we did not observe any evidence for List-1 reactivation. This result is not easily compatible with the reconsolidation account for the List-2 intrusions observed in Experiment 2.

In conclusion, we not only identified important boundary conditions for the increase List-2 intrusion rates occurring in the 3-d-list-interference paradigm, but our results pose substantial challenges for the reconsolidation account. In contrast, the pattern of result is fully consistent with an interference account and mirrors previous findings of single-session experiments without delays that could allow for consolidation (Nagge 1935; Bilodeau and Schlosberg 1951; Greenspoon and Ranyard 1957; Dallett and Wilcox 1968). In addition, the interference account for the List-2 intrusions seems more parsimonious because it relies only on well-established cognitive processes during retrieval (DeVietti and Larson 1971; Unsworth et al. 2012) and not on proposing novel neurobiological processes (the undoing of the synaptic changes underlying complex episodic memory formation) .

References

Alberini CM. 2007. Reconsolidation: the samsara of memory consolidation. *Debates Neurosci* **1:** 17–24.

Bilodeau IM, Schlosberg H. 1951. Similarity in stimulating conditions as a variable in retroactive inhibition. *J Exp Psychol* **41:** 199–204.

Dallett K, Wilcox SG. 1968. Contextual stimuli and proactive inhibition. *J Exp Psychol* **78:** 475–480.

DeVietti TL, Larson RC. 1971. ECS effects: evidence supporting state-dependent learning in rats. *J Comp Physiol Psychol* **74:** 407–415.

Dongaonkar B, Hupbach A, Gomez R, Nadel L. 2013. Effects of psychosocial stress on episodic memory updating. *Psychopharmacol Berl* **226:** 769–779.

Faul F, Erdfelder E, Buchner A, Lang A-G. 2009. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* **41:** 1149–1160.

Gershman SJ, Schapiro AC, Hupbach A, Norman KA. 2013. Neural context reinstatement predicts memory misattribution. *J Neurosci Off J Soc Neurosci* **33:** 8590–8595.

Gisquet-Verrier P, Riccio DC. 2012. Memory reactivation effects independent of reconsolidation. *Learn Mem Cold Spring Harb N* **19:** 401–409.

Gisquet-Verrier P, Lynch JF, Cutolo P, Toledano D, Ulmen A, Jasnow AM, Riccio DC. 2015. Integration of new information with active memory accounts for retrograde amnesia: a challenge to the consolidation/reconsolidation hypothesis? *J Neurosci Off J Soc Neurosci* **35:** 11623–11633.

Greenspoon J, Ranyard R. 1957. Stimulus conditions and retroactive inhibition. *J Exp Psychol* **53:** 55–59.

Hardt O, Einarsson EO, Nader K. 2010. A bridge over troubled water: reconsolidation as a link between cognitive and neuroscientific memory research traditions. *Annu Rev Psychol* **61:** 141–167.

Hardwicke TE, Taqi M, Shanks DR. 2016. Postretrieval new learning does not reliably induce human memory updating via reconsolidation. *Proc Natl Acad Sci* **113:** 5206–5211.

Horn W. 1983. *L-P-S Leistungsprüfsystem*, 2nd ed. Hogrefe, Göttingen.

Hupbach A. 2015. Retrieval practice does not safeguard memories from interference-based forgetting. *Learn Motiv* **49:** 23–30.

Hupbach A, Gomez R, Hardt O, Nadel L. 2007. Reconsolidation of episodic memories: a subtle reminder triggers integration of new information. *Learn Mem* **14:** 47–53.

Hupbach A, Hardt O, Gomez R, Nadel L. 2008. The dynamics of memory: context-dependent updating. *Learn Mem* **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. *Psychon Bull Rev* **18:** 787–797.

Jones BJ, Bukoski E, Nadel L, Fellous J-M. 2012. Remaking memories: reconsolidation updates positively motivated spatial memory in rats. *Learn Mem Cold Spring Harb N* **19:** 91–98.

Jones BJ, Pest SM, Vargas IM, Glisky EL, Fellous J-M. 2015. Contextual reminders fail to trigger memory reconsolidation in aged rats and aged humans. *Neurobiol Learn Mem* **120:** 7–15.

Kroes MCW, Tendolkar I, van Wingen GA, van Waarde JA, Strange BA, Fernández G. 2014. An electroconvulsive therapy procedure impairs reconsolidation of episodic memories in humans. *Nat Neurosci* **17:** 204–206.

Mamiya N, Fukushima H, Suzuki A, Matsuyama Z, Homma S, Frankland PW, Kida S. 2009. Brain region-specific gene expression activation required for reconsolidation and extinction of contextual fear memory. *J Neurosci* **29:** 402–413.

McKenzie S, Eichenbaum H. 2011. Consolidation and reconsolidation: two lives of memories? *Neuron* **71:** 224–233.

Nadel L, Hupbach A, Gomez R, Newman-Smith K. 2012. Memory formation, consolidation and transformation. *Neurosci Biobehav Rev* **36:** 1640–1645.

Nagge JW. 1935. An experimental test of the theory of associative interference. *J Exp Psychol* **18:** 663–682.

Nicholls MER, Loveless KM, Thomas NA, Loetscher T, Churches O. 2015. Some participants may be better than others: sustained attention and motivation are higher early in semester. *Q J Exp Psychol* **68:** 10–18.

Potts R, Shanks DR. 2012. Can testing immunize memories against interference? *J Exp Psychol Learn Mem Cogn* **38:** 1780–1785.

Riccio DC, Moody EW, Millin PM. 2002. Reconsolidation reconsidered. *Integr Physiol Behav Sci Off J Pavlov Soc* **37:** 245–253.

Riccio DC, Millin PM, Bogart AR. 2006. Reconsolidation: a brief history, a retrieval view, and some recent issues. *Learn Mem* **13:** 536–544.

Roediger HL, Butler AC. 2010. The critical role of retrieval practice in long-term retention. *Trends Cogn Sci.* http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20951630.

Sara SJ. 2008. Reconsolidation: historical perspective and theoretical aspects. In *Cognitive psychology of memory* (ed. Roediger HL), pp. 460–475. Elsevier, Oxford.

Schiller D, Phelps EA. 2011. Does reconsolidation occur in humans? *Front Behav Neurosci* **5:** 24.

Schiller D, Monfils MH, Raio CM, Johnson DC, Ledoux JE, Phelps EA. 2010. Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature* **463:** 49–53.

Schwabe L, Nader K, Pruessner JC. 2014. Reconsolidation of human memory: brain mechanisms and clinical relevance. *Biol Psychiatry* **76:** 274–280.

Sederberg PB, Gershman SJ, Polyn SM, Norman KA. 2011. Human memory reconsolidation can be explained using the temporal context model. *Psychon Bull Rev* **18:** 455–468.

Soeter M, Kindt M. 2011. Disrupting reconsolidation: pharmacological and behavioral manipulations. *Learn Mem* **18:** 357–366.

Unsworth N, Spillers GJ, Brewer GA. 2012. Dynamics of context-dependent recall: an examination of internal and external context change. *J Mem Lang* **66:** 1–16.

Walker MP, Brakefield T, Hobson JA, Stickgold R. 2003. Dissociable stages of human memory consolidation and reconsolidation. *Nature* **425:** 616–620.

Wichert S, Wolf OT, Schwabe L. 2011. Reactivation, interference, and reconsolidation: are recent and remote memories likewise susceptible? *Behav Neurosci* **125:** 699–704.

Wichert S, Wolf OT, Schwabe L. 2013a. Changing memories after reactivation: a one-time opportunity? *Neurobiol Learn Mem* **99:** 38–49.

Wichert S, Wolf OT, Schwabe L. 2013b. Updating of episodic memories depends on the strength of new learning after memory reactivation. *Behav Neurosci* **127:** 331–338.