SOCIAL SCIENCES: Psychological and Cognitive Sciences

Research Report

Title: Impairing Existing Declarative Memory in Humans by Disrupting Reconsolidation

Authors: Jason C.K. Chan\textsuperscript{a,b,1}, Jessica A. LaPaglia\textsuperscript{a}

See final version here: http://www.pnas.org/content/110/23/9309.short

\textsuperscript{a}Department of Psychology,  
\textsuperscript{b}Neuroscience Interdisciplinary Program,  
Iowa State University,  
Ames, IA 50014  
USA

\textsuperscript{1}Corresponding Author:  
Jason C.K. Chan  
Department of Psychology  
Iowa State University  
W112 Lagomarcino Hall  
Ames, IA 50014  
USA  
(515) 294-6946  
ckchan@iastate.edu
Abstract

Over the past decade, a large body of research has shown that memory traces can become labile upon retrieval and must be restabilized. Critically, interrupting this reconsolidation process can abolish a previously stable memory. Although a large number of studies have demonstrated this reconsolidation associated amnesia in non-human animals, the evidence for its occurrence in humans is far less compelling, especially with regard to declarative memory. In fact, reactivating a declarative memory often makes it more robust and less susceptible to subsequent disruptions. Here we show that existing declarative memories can be selectively impaired using a noninvasive retrieval-relearning technique. In six experiments, we show that this reconsolidation associated amnesia can be achieved 48 hr after formation of the original memory, but only if relearning occurred soon after retrieval. Furthermore, the amnesic effect persists for at least 24 hr, cannot be attributed solely to source confusion, and is attainable only when relearning targets specific existing memories for impairment. These results demonstrate that human declarative memory can be selectively rewritten during reconsolidation.

(Word Count: 170)
The entrenched view that memory becomes permanent upon consolidation has faced considerable scrutiny based on recent works demonstrating that retrieval can destabilize existing memories, and that the reactivated memories need to be reconsolidated (1, 2). Over the past decade, a growing body of evidence has revealed the chemical and molecular nature of reconsolidation and its behavioral consequences (3). Critically, when a consolidated memory (e.g., a conditioned fear response) is retrieved, it becomes labile and requires protein synthesis for restabilization, and later retrieval of that memory can be severely impaired if an amnesic treatment is administered during the reconsolidation process.

Despite the proliferation of research on reconsolidation, few studies have involved human subjects, perhaps because most pharmacological consolidation blockers are unsuitable for human use (3, 4). When reconsolidation associated memory impairments are demonstrated in humans, the effects have been limited to fear conditioning (5-8), motor sequence learning (9), and drug induced craving (10). To date, no study has shown reconsolidation associated impairment in declarative memory. Indeed, even when oral administration of propranolol (a systemic pharmacological consolidation blocker approved for human use) reduced the emotional response associated with a fear-inducing experience (6, 7), it left the declarative recollection intact*. Although several studies have examined whether existing declarative memories can be impaired by interference upon reactivation, none has shown memory impairments similar to those regularly exhibited in fear conditioning. Moreover, these studies used a reminder to trigger reactivation of the original memory while prohibiting subjects from actually retrieving that memory (e.g., by briefly mentioning the learning episode without asking about what was learned); thus, it is unclear whether retrieval of the original memory (and thus reconsolidation)
actually occurred (11-16). When retrieval of a learned response is ascertained through a memory test, the results almost unequivocally show that retrieval actually makes the original memory less, not more, susceptible to interference (17-20). In sum, scant evidence exists to support the idea that declarative memory undergoes reconsolidation upon retrieval.

Distinct neural systems subserve the formation and retrieval of fear, motor, and declarative memory (21). Whereas memories acquired via fear conditioning or motor sequence learning can be relatively localized neurologically (22), the encoding and retrieval of declarative memories rely on a more distributed network (23-27). This complexity might be one reason why declarative memory is particularly resistant to treatments designed to disrupt reconsolidation. Here we show that human declarative memory can be selectively impaired in a behavioral paradigm without the administration of harmful pharmacological agents; these results demonstrate that declarative memory, too, is susceptible to reactivation-induced lability.

Results

Experiment 1: Targeting specific declarative memories for impairment in humans

In the present study, we used a retrieval-relearning procedure to disrupt reconsolidation of the original memory (see fig. 1). Participants watched a video about a fictional terrorist attack during the original learning phase. Our key manipulation was whether subjects recalled specific details from the original learning episode (e.g., a terrorist used a hypodermic needle on a flight attendant) before they encountered new information (i.e., misinformation) that replaced the original information (the terrorist used a stun gun). Performance during the reactivation phase indicates moderate retention, see fig. 2. This is important because very strong memories are resistant to post-reactivation amnesic treatments (28, 29). Moreover, when recall performance during the reactivation phase was examined based on item type (i.e., whether an item would be
re-presented, omitted, or misinformed during the subsequent relearning phase), no difference emerged across any experiments, $F_s < 1.86, P_s > .17$. Therefore, any difference in performance across item type during the final test could not be attributed to baserate differences.

If retrieval triggers reconsolidation of the original memory, then the new information presented during the relearning phase should update the original memory and render it inaccessible. To assess whether amnesia of the original memory had occurred, we compared participants’ ability to recognize the original item depending on whether the item was replaced by misinformation (misinformed items) or not (neutral items). Impairment of the original memory is indicated by poorer recognition performance of the misinformed items relative to the neutral items. All statistical analyses were performed two-tailed with an alpha level of 0.05, and the dependent variable was hit rate – false alarm rate unless noted otherwise.

Much research effort has been devoted to uncovering whether existing declarative memories can be erased by learning new materials (without the reactivation component), and the results have been underwhelming (30). Consistent with these findings, results in Experiment 1 ($n = 146$) showed that relearning produced no impairment of the original memory in the absence of reactivation ($t = 0.002$, see the first white bar in fig. 3). In contrast, when relearning occurred after the original memory was reactivated, it markedly reduced recognition performance (i.e., $M = 0.55$ for neutral items and $M = 0.36$ for misinformed items), $t(69) = 3.74, P < 0.05, d = 0.57$ (see the first gray bar in fig. 3). Thus, the findings in Experiment 1 suggest that our retrieval-relearning procedure disrupted the reconsolidation process; moreover, they show that under some circumstances (namely, when retrieval precedes relearning – a manipulation rarely included in the psychological research of eyewitness memory), misinformation can indeed impair subsequent retrieval of the original memory.
Experiments 2 and 3: Reconsolidation impairment in declarative memory is time-dependent

In Experiments 2 ($n = 64$) and 3 ($n = 64$), we examined whether the amnesic effect produced by this retrieval-relearning manipulation is time dependent. In Experiment 2, a 48 hr (instead of 20 min) delay separated reactivation and relearning (fig. 1). If the memory impairment seen in Experiment 1 was based on reconsolidation disruption, it should be eliminated here because relearning occurred long after closure of the reconsolidation window ($5, 10$). Indeed, the relearning procedure now produced no memory impairment regardless of whether it occurred after reactivation, $t = 0.96$, or not, $t = 0.30$ (fig. 2). In Experiment 3, a 48 hr delay separated original learning and reactivation (fig. 1), and the relearning phase occurred immediately after reactivation. If the memory impairment in Experiment 1 was due to reconsolidation disruption, it should resurface in Experiment 3, even though the original memory has had sufficient time (48 hr) to fully consolidate following initial encoding. Consistent with our prediction, the relearning procedure once again produced substantial amnesia for the original item ($M = 0.40$ for neutral items and $M = 0.15$ for misinformed items), $t(31) = 3.50, P < 0.05, d = 0.82$ (fig. 3), but only when it occurred after retrieval ($t = 0.25$ without retrieval).

Experiment 4: Reconsolidation associated amnesia cannot be explained solely by source confusions

Having established the time-dependent nature of this retrieval-relearning amnesia, we sought to rule out source confusion as the basis of this effect. Two recent papers have reported that when people are reminded of the original learning episode immediately before new learning, they often misattribute details encountered during the new learning phase to the original learning phase (11, 31). Therefore, it is possible that our manipulation did not produce an amnesic effect
on the original memory per se; rather, recognition performance was reduced due to source confusions. Specifically, participants might remember the original item but were unable to discern whether the item was encountered during original learning or relearning. Such doubts could cause participants to adopt a more conservative response criterion in a recognition test, thus reducing the hit rate without affecting accessibility of the original memory. To address this possibility, we administered a source-free recognition test in Experiment 4 (n = 72). Participants were told to respond "old" if they remembered the information from either the original learning phase or the relearning phase and to respond "new" otherwise. As with former studies, the dependent variable of interest in source-free recognition is the hit rate (32). If the amnesic effect reported in Experiments 1 and 3 were based only on source confusions and not true memory impairment, then it should not occur here. Contrary to this possibility, the retrieval-relearning amnesia remained (M = 0.73 for neutral items and M = 0.60 for misinformed items), t(35) = 2.10, P < 0.05, d = 0.50 (and again relearning produced no amnesia without reactivation, t = 0.20, fig. 3). From a procedural standpoint, this experiment is highly comparable to Experiment 1, and the effect size of the reconsolidation associated amnesia was similar in these experiments (d = 0.57 for Experiment 1 and d = 0.50 for the present experiment), which suggests that using source free recognition did not diminish the magnitude of reconsolidation associated amnesia appreciably. Nonetheless, the effect size was numerically smaller in the present experiment; thus, consistent with prior research (11, 31), reactivation-induced source confusions might have played a partial role in the memory impairment observed in the previous experiments.

**Experiment 5: Specificity of memory replacement**

In Experiment 5 (n = 84), we attempted to address why previous studies have not discovered reconsolidation associated amnesia in declarative memory. More broadly, because
encoding of new information happens on an ongoing basis in “real life” situations, why would such encoding not disrupt reconsolidation of a recently retrieved memory? To this end, we examined whether specific, rather than nonspecific, interference (33) is the key to demonstrating this reactivation-relearning amnesia effect. We suspect that the post-reactivation interfering agent must compete directly with, or replace, the originally learned information for memory impairment to occur (3, 34, 35). This would explain why systemic administration of propranolol (6) and learning of a new set of materials (12, 13, 15, 17) do not always produce amnesia of the original memory. Importantly, even in Pavlovian conditioning, reconsolidation associated amnesia is found only when a direct association exists between the interfering agent and the original memory (36). To test the idea that specific interference is needed to alter a reactivated memory, we modified Experiment 1 so that the relearning narrative presented the same misinformation (e.g., a stun gun) but in the context of an unrelated story about drug trafficking (see table S1 for details), thus turning the relearning phase into a new learning phase (i.e., nonspecific interference). As expected, the relearning procedure no longer impaired recognition performance, regardless of whether it followed retrieval, $t = 0.99$, or not, $t = 0.45$. More broadly, results from this experiment clarified why humans do not demonstrate reconsolidation associated amnesia following recall in everyday life, despite constant interference from new encoding opportunities whenever one is awake.

**Experiment 6: Disrupting reconsolidation produces persistent amnesia**

In the first five experiments, the retention interval that separated relearning and the final test was 5 min. Consequently, it is unclear whether our retrieval-relearning procedure can cause long-term amnesia of the original memory. In Experiment 6 ($n = 66$), we examined the persistence of the reconsolidation associated amnesia following a 24 hr delay. Because
fluctuations in the forgetting function of human declarative memory stabilize substantially after 24 hours, the data from this experiment would likely generalize to longer retention intervals (37). In addition, we inserted a 24 hr delay between original learning and the reactivation phase. Similar to Experiment 3, this delay allowed the original memory trace to more fully consolidate before it was reactivated. Thus, Experiment 6 was conducted over three consecutive days, with the original learning occurring on the first day, the retrieval and relearning phases occurring on the second day, and the final memory assessment occurring on the third day. Critically, memory deficits based on temporary suppression typically rebounds over time (38), so if the retrieval-relearning amnesia effect is based on short-lived inhibition or interference mechanisms, it should be eliminated in this experiment. Contrary to this possibility, relearning again caused forgetting of the original memory only when it followed retrieval ($M = 0.38$ for neutral items and $M = 0.25$ for misinformed items), $t(32) = 2.04, P < 0.05, d = 0.34$ (no-retrieval, $t = -0.72$).

**Disruption of reconsolidation occurs despite successful retrieval of the original memory**

We found recognition impairment due to the retrieval-relearning procedure in Experiments 1, 3, 4, and 6 (and as predicted, no impairment was observed in Experiments 2 and 5). One question is whether this amnesic effect was driven only by items that people could not recall during the reactivation phase (i.e., weaker memories) or if the effect occurred even for items that were correctly recalled during the reactivation phase (i.e., stronger memories). To address this question, we reanalyzed the results of Experiments 1, 3, 4, and 6 by examining recognition performance for only the items that were correctly recalled during reactivation. Remarkably, substantial relearning-induced amnesia is observed even for these initially recallable (and therefore strong) items. Across the four experiments, the average mean recognition probability was .89 for the neutral items and .62 for the misinformed items, with
significant impairment in each of the four experiments, $t_s > 2.14, P_s < .05$, $d_s > .55^†$. These results are particularly noteworthy because initially recallable items are almost always recognized (39), but relearning during reconsolidation had produced dramatic forgetting of these originally recallable memories.

To further investigate the effects of the reactivation-relearning procedure on subsequent memory performance, we conducted three additional analyses. First, we examined whether reactivation affected the likelihood that one would demonstrate memory impairment (see Table S2). Second, we examined whether reactivation affected the magnitude of memory impairment (see Table S3). Third, we examined whether reactivation affected the response latency associated with recognition decisions (see Table S4). Details regarding all of these analyses and the logic behind them are presented in the SI.

**Discussion**

Decades of behavioral research in humans has revealed that it is trivially easy to alter people’s memory reports and to create false memories (40), but the question surrounding whether it is possible to experimentally erase existing memories in humans has been far more contentious (30). Here, we demonstrate that retrieval can destabilize a declarative memory and render it susceptible to an amnesic treatment by relearning even after a substantial delay (i.e., at least 48 hr). Consistent with the known preconditions of reconsolidation, the timing of post-retrieval amnesic treatment is critical, such that relearning impairs the original memory only when it occurs within the reconsolidation window. Note that compared to other studies that have investigated the time course of reconsolidation, our experiments used retention intervals that are at the more extreme ends of the spectrum (e.g., immediate vs. 48 hr), whereas some other studies have used more fine-grained comparisons. These studies typically show that the reconsolidation
window may last somewhere between one and six hours. We opted for a more extreme comparison because it was not clear whether the time-related parameters found with nonhuman animals (41, 42) and non-declarative memory in humans (5, 10) are applicable to declarative memory. Future research is needed to better specify the time course of reconsolidation in human declarative memory.

In addition to the time-dependent properties of the effect, we have shown that reactivated memories are vulnerable only to interference that specifically targets existing memories, but are robust to nonspecific interference produced by new learning. Moreover, this effect persists for at least 24 hr and cannot be attributed solely to source confusions. When the relearning procedure targeted specific recollections for impairment, considerable amnesia was observed across the four experiments in which we expected reconsolidation disruption. However, when relearning occurred without reactivation of the original memory, it produced no impairments in performance across all six experiments (a nonsignificant rise of 3% in performance).

Although our data suggest that one can target specific existing memories for impairment following their reactivation, we believe that specific interference might be necessary, but not sufficient, to disrupt reconsolidation of declarative memory. On a processing level, it is unlikely for people to update an original memory unless they believe that the new learning phase accurately re-presents the original information. We suspect that a key to impairing the original memory is that relearning does not trigger spontaneous retrieval of the original information. Instead, reactivation of the original memory must occur prior to, but not concurrently with, relearning. Recalling the original memory while encoding the new one would likely cause people to either discount the new information or to remember both the new and original information, which would likely eliminate the updating effect. In a similar vein, research on
eyewitness suggestibility has shown that people are highly resistant to suggestions by misinformation if they detect a conflict between what was originally learned and the misinformation (43, 44). Thus, we believe that the expectation under which relearning occurs can determine whether the retrieval-relearning procedure would lead to enhancement or amnesia of the original memory.

One may wonder whether participants who received the initial test were more likely to accept the misinformation as true and thus less likely to recognize the original details during the final test. This could occur if participants, after having their memory tested during the reactivation phase, thought that the experimenter was giving them corrective feedback during the relearning phase. In other words, the retrieval-relearning procedure might have changed participants’ belief about what was correct without actually changing memory. If this were the case, we should not have observed time-dependent reconsolidation associated amnesia (in Experiments 2 and 3), nor should we have obtained the effect in the source-free recognition test (in Experiment 4), where participants were told to respond based on their memory of both the original learning and relearning phases.

The malleability of human declarative memory has been a major topic of research for decades (40, 45). Recently, the instability of memory following reactivation has been suggested to play a key role in the production and implantation of erroneous memories (46). While the unreliability of declarative memory is typically considered a disadvantage, our results show that it is possible to leverage the lability of reactivated engrams for targeted impairment, thus providing a noninvasive method to weaken the impact of unwanted memories. Aside from its noninvasive nature, this technique is particularly powerful because one can target specific recollections for impairment while sparing others. Humans are notoriously inept at suppressing
unwanted thoughts (47). In fact, attempting to block unwanted thoughts from consciousness often leads to the opposite effect (48). Remarkably, when treating post-traumatic stress disorders, therapeutic techniques that require patients to recall their traumatic memories (e.g., exposure, acceptance, and paradoxical approaches) are typically far more successful than suppression (49-51). For example, a form of acceptance therapy requires patients to recall their intrusive thoughts and reinterpret them in a safe context (e.g., imagine the thought as a band of marching soldiers emerging from the ears). Existing explanations have ascribed the efficacy of these approaches to the acceptance of the unwanted events as belonging to the past and to detaching oneself from the negative feelings associated with the experience (52). Based on the present findings, it is possible that these techniques are successful because they are, at some level, exploiting the post-retrieval updating characteristics of memory.

A note of caution is in order here. Similar to many existing experiments demonstrating reconsolidation-like effects, it is not possible to know whether our reactivation-relearning manipulation impaired memory performance by weakening the original memory (i.e., a storage deficit) or by impairing its retrieval. It is also unknown whether different retrieval environments or subject factors [e.g., differences in overall suggestibility (53), differences in executive functioning (54)] can protect one from the present forgetting effect. Thus, further research is needed to clarify the boundary conditions of our findings. But these data represent an important step towards a fuller understanding of the mechanisms responsible for the plasticity of declarative memory in humans. Knowing the limits and operating characteristics of reconsolidation blockage in declarative memory can have a profound impact on how memory is conceived theoretically. These data also further bolster the idea that reconsolidation plays a fundamental role in the formation and maintenance of memory in humans.
Methods

Subjects. All subjects were recruited from the Iowa State University community and they received either partial course credit or a payment of $15 for their participation. All participants were native English speakers.

Materials and Procedure. In all experiments, subjects viewed a ~40 min video about a terrorist attack (the original learning phase). The video was the pilot episode of the television drama “24.” Subjects were given intentional learning instructions prior to watching the video. Reactivation of the original memory was then manipulated by having participants complete a cued recall test of the video (the reactivation condition) or play the videogame Tetris (the no-reactivation condition). During the recall test, subjects were asked 24 questions about the video (e.g., What does the terrorist use on the flight attendant?) and they had 25 sec to answer each question. No feedback was provided during this reactivation phase. After a filled delay during which subjects completed a working memory task (see fig. 1 for the exact delay implemented in each experiment), participants listened to an 8 min audio narrative that purportedly recapped the video (the relearning phase). However, among the 24 details queried during the reactivation phase, eight were presented incorrectly, and these details (misinformed items) replaced the original information (e.g., in the video, a terrorist knocked a flight attendant unconscious with a hypodermic syringe, but the narrative described the weapon as a stun gun). Eight other details that were queried during the reactivation phase were not mentioned during the narrative (neutral items), and the remaining eight details were re-presented correctly during the narrative (re-presented items). Refer to fig. S1 for data regarding the re-presented items. Participants were not informed about any inaccuracies and all information in the narrative was presented as fact. After a retention interval (which was 5 min in Experiments 1-5 and 24 hr in Experiment 6),
memory of the video was assessed in a final test. We devised a special true/false recognition test to estimate the accessibility of the original memory (55, 56). Recall tests are not suitable for our purpose because participants can withhold responses based on various metacognitive control processes (31, 57), making assessment of the true strength of a memory difficult. In the recognition test, participants encountered one statement during each trial and indicated whether the statement was true (e.g., The terrorist used a hypodermic syringe on the flight attendant) or false (e.g., The terrorist used a chloroform rag on the flight attendant). Critically, the misinformation (e.g., the stun gun) was never presented during this recognition test so that performance would not be influenced by non-memory based factors (e.g., demand characteristics). In Experiment 4, instead of this true/false recognition test, a source-free recognition test was administered. Here, participants were told to make an “old” judgment if they remembered the event detail from either the original learning or the relearning phase; otherwise, they were to make a “new” judgment.
Footnotes

* A recent study showed that emotional (but not nonemotional) declarative memory might need to reconsolidate upon retrieval (16). However, this study did not obtain overt recall responses from participants during the memory reactivation phase, thus like many other human studies, it is not possible to ascertain that retrieval was attempted. Moreover, the amnesic treatment (propranolol) was administered orally an hour before the memory reactivation phase. Therefore, it is possible that, instead of blocking reconsolidation, propranolol exerted an influence on retrieval and produced a permanent effect on the memory (4).

† To ensure relatively stable results in these conditional analyses, we included data from only participants who recalled at least three items correctly for both the neutral and misinformed items during the reactivation phase. As such, the degrees of freedom for these analyses are smaller than those of the analyses that included all items \( (df_{Exp\ 1} = 52, df_{Exp\ 3} = 10, df_{Exp\ 4} = 27, df_{Exp\ 6} = 14) \).
Author Contributions

J.C.K.C. designed research, analyzed data, and wrote the paper. J.A.L. designed research, performed research, and analyzed data. We thank Keely Burke for collecting data of Experiment 6.


Figure Legends

Figure 1. Timeline of the experimental procedure. Arrows at the bottom of the figure indicate major methodological changes for Experiments 2-6.

Figure 2. Performance during the memory reactivation phase of Experiments 1-6.

Figure 3. Reactivation is required for targeted impairment of existing declarative memory. Each bar is the difference in recognition accuracy between the misinformed (relearned) items and the neutral (not-relearned) items. The white bars indicate performance in the no-reactivation condition and the gray bars indicate performance in the reactivation condition. A negative score indicates poorer memory following relearning. As can be seen, relearning produced no forgetting of the original memory when it was not preceded by reactivation in any of the experiments. However, in experiments where reconsolidation associated amnesia was expected (as indicated by asterisks), the retrieval-relearning procedure led to substantial impairment of the original memory. Error bars display 95% confidence interval.
Original Learning

Reactivation (Memory Test)

No-Reactivation (Distractor Task)

20 min

Relearning

5 min

Recognition Memory Test

24 hr in E6

Source-free Recognition in E4

48 hr in E3

24 hr in E6

48 hr in E2

0 min in E3 & E6

Learn unrelated materials in E5