Postretrieval new learning does not reliably induce human memory updating via reconsolidation

Tom E. Hardwicke[a,1], Mahdi Taqi[a], and David R. Shanks[a]

[a]Department of Experimental Psychology, University College London, London WC1H 0AP, United Kingdom

Edited by James L. McClelland, Stanford University, Stanford, CA, and approved April 4, 2016 (received for review January 27, 2016)

Reconsolidation theory proposes that retrieval can destabilize an existing memory trace, opening a time-dependent window during which that trace is amenable to modification. Support for the theory is largely drawn from nonhuman animal studies that use invasive pharmacological or electroconvulsive interventions to disrupt a putative postretrieval restabilization ("reconsolidation") process. In human reconsolidation studies, however, it is often claimed that postretrieval new learning can be used as a means of "updating" or "rewriting" existing memory traces. This proposal warrants close scrutiny because the ability to modify information stored in the memory system has profound theoretical, clinical, and ethical implications. The present study aimed to replicate and extend a prominent 3-day motor-sequence learning study [Walker MP, Brakefield T, Hobson JA, Stickgold R (2003) *Nature* 425(6958): 616–620] that is widely cited as a convincing demonstration of human reconsolidation. However, in four direct replication attempts ($n = 64$), we did not observe the critical impairment effect that has previously been taken to indicate disruption of an existing motor memory trace. In three additional conceptual replications ($n = 48$), we explored the broader validity of reconsolidation-updating theory by using a declarative recall task and sequences similar to phone numbers or computer passwords. Rather than inducing vulnerability to interference, memory retrieval appeared to aid the preservation of existing sequence knowledge relative to a no-retrieval control group. These findings suggest that memory retrieval followed by new learning does not reliably induce human memory updating via reconsolidation.

reconsolidation | sequence learning | memory updating | forgetting | replication

Reconsolidation theory proposes that retrieval of existing memory traces causes them to destabilize, triggering a transient molecular restabilization ("reconsolidation") process during which they are open to modification (1, 2). If reconsolidation enables memory modification in humans, it could have profound theoretical (3), clinical (4), and ethical (5) implications. For example, the ability to erase "pathological" memory traces that contribute to posttraumatic stress disorder, addiction, and phobias, offers the potential of permanent relief from these conditions (4).

Proponents of reconsolidation theory suggest that there is broad empirical support across a range of species, tasks, and memory types (2, 4, 6), but several authors have expressed skepticism about the extent to which existing studies rule out alternative explanations (7–9). Extending reconsolidation investigations to human participants has proved particularly challenging. Support for the theory is largely based on nonhuman animal studies in which invasive interventions, such as electroconvulsive shock or pharmacological treatment, are delivered following retrieval of an established memory trace (2, 6). By contrast, ethical constraints have led to the use of new learning as a postretrieval intervention in many investigations with human participants (e.g., refs. 10–12; for review, see ref. 13). In both cases, the observation of substantial trace-dependent performance impairments on a subsequent test is taken as evidence that the intervention has disrupted the reconsolidation of the memory trace, resulting in its modification or destruction. Although physiological interventions are intended to directly disrupt the putative molecular substrates of reconsolidation, considerable ambiguity surrounds the envisioned mechanism by which a behavioral intervention might influence these same processes. Nevertheless, there are prevalent claims about the functional role of reconsolidation as a memory "updating" mechanism (14–16) whereby existing memory traces are selectively "rewritten" by postretrieval new learning (12, 17).

It is worth noting that the reconsolidation controversy is only the latest chapter in an enduring historical debate about the locus of interference and forgetting effects (18). On the one hand, amnesia for previously recallable information has been attributed to storage deficits: the permanent physical modification of memory traces by postencoding and postretrieval interventions [e.g., "consolidation" (19); "unlearning" (20); "destructive updating" (21); "reconsolidation" (2)]. On the other hand, amnesia has been attributed to mechanisms operating during trace retrieval that temporarily modulate trace-dependent performance without necessarily influencing the underlying memory trace [e.g., "response competition" (22); "cue-dependent forgetting" (23); "state-dependent retrieval" (24); "context-dependent forgetting" (25)]. These retrieval deficit accounts can explain experimentally induced amnesia without invoking claims about physical trace disruption that cannot be directly observed. They also provide a more convincing account of the widespread finding that impairments of trace-dependent performance are often temporary and show high propensity for recovery under favorable retrieval conditions (26). This debate is particularly pertinent to the evaluation of reconsolidation studies because retrieval deficit explanations are often overlooked (7–9).

A particularly prominent finding reported by Walker et al. (ref. 27, group 7, hereafter referred to as the original study) is widely cited as a convincing demonstration of reconsolidation-mediated

Significance

Reconsolidation-updating theory suggests that existing memory traces can be modified, or even erased, by postretrieval new learning. Compelling empirical support for this claim could have profound theoretical, clinical, and ethical implications. However, demonstrating reconsolidation-mediated memory updating in humans has proved particularly challenging. In four direct and three conceptual replication attempts of a prominent human reconsolidation study, we did not observe any reconsolidation effects when testing either procedural or declarative recall of sequence knowledge. These findings suggest that the considerable theoretical weight attributed to the original study is unwarranted and that postretrieval new learning does not reliably induce human memory updating via reconsolidation.

Author contributions: T.E.H. and D.R.S. designed research; T.E.H. and M.T. performed research; T.E.H. and D.R.S. analyzed data; and T.E.H. and D.R.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: t.hardwicke.12@ucl.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1601440113/-/DCSupplemental.

memory updating in humans (e.g., refs. 2, 4, 6, 13, and 14). The results are especially compelling because the experiment conformed to the canonical 3-day reconsolidation protocol (Fig. 1) typically used in nonhuman animal studies, thus meeting several key criteria necessary for a robust investigation of reconsolidation (2, 4, 6). On day 1, participants used a computer keyboard to repeatedly tap a simple sequence of on-screen digits (e.g., 41342). Speed and accuracy improvements were observed as participants learned this initial ("Old") sequence. On day 2, participants in the Reminder group ($n = 16$) practiced the Old Sequence immediately before learning a New Sequence. The No-Reminder group did not practice the Old Sequence before new learning. The No-Intervention group practiced the Old Sequence but did not learn a New Sequence. On day 3, sequence performance was tested for all groups. The key finding was that the Reminder group's Old Sequence accuracy suffered a substantial decline (~57%) between the Reminder stage and the Test stage, although only minor decrements were observed on the speed measure (~2%). By contrast, improvements in accuracy and speed between Training and Test stages were observed in the No-Reminder and No-Intervention groups. Therefore, it would appear that the accuracy impairment in the Reminder group was contingent on the time-dependent interaction of the reminder and intervention as demonstrated in similar nonhuman animal studies (1) and widely accepted as evidence for reconsolidation (2, 4, 6). Consistent with the view that the Old Sequence memory trace had been rewritten by the new learning (12, 17), the authors suggested that reconsolidation may have "functional significance," allowing the "continued refinement and reshaping of previously learned movement skills" (ref. 27, p. 618).

However, from the perspective of the aforementioned storage–retrieval debate (18), this interpretation should be viewed with caution, especially as retrieval deficit explanations were not explored. For example, it was not clear whether the effect endured beyond the 3-day study period, or showed propensity for recovery under favorable retrieval conditions (26), effects that have been observed in several investigations of reconsolidation with nonhuman animals (e.g., refs. 28–30). In the present study, we initially sought to replicate and extend the reported reconsolidation effect (ref. 27, group 7) by examining whether it could be accounted for by retrieval deficits rather than the storage deficit mechanisms outlined under reconsolidation theory (our investigation does not address other findings, unrelated to reconsolidation, reported in the same article). We conducted a replication battery (31) consisting of both "direct replications" (32) that followed the methodology of the original study as closely as possible, and "conceptual replications" (33) that manipulated key task parameters to explore the broader validity of the reconsolidation-updating theory.

To foreshadow our findings, the complete absence of a reconsolidation effect in any of our experiments precluded any further investigation of a retrieval deficit account. Instead, we made several attempts to reproduce the effect in repeated direct replications ($n = 64$) using our own software (experiment 1), software provided by the original researchers (experiment 2), and under conditions intended to increase task difficulty (experiments 3 and 4). In our conceptual replications ($n = 48$), we used "declarative" recall conditions more consistent with the wider human reconsolidation literature (e.g., refs. 10 and 11). These experiments also involved sequence learning within a 3-day reconsolidation protocol (Fig. 1), but used sequences similar in length and structure to phone numbers (experiments 5 and 7) or computer passwords (experiment 6). A No-Reminder control group (experiment 7) enabled us to ascertain whether performance impairments were contingent on retrieval-induced vulnerability as predicted by reconsolidation theory.

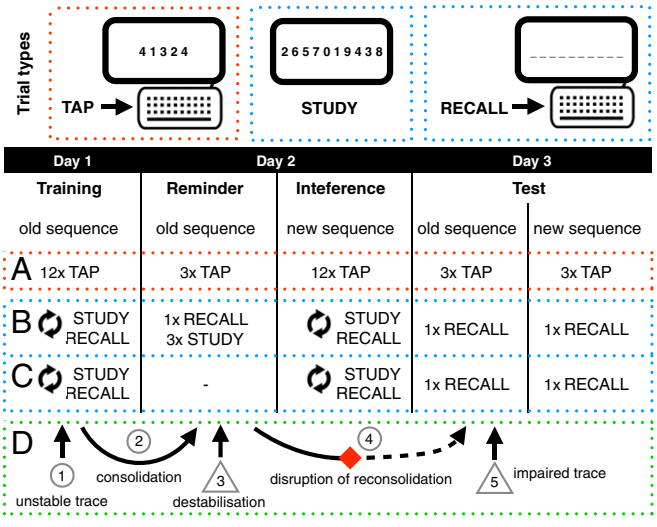

Fig. 1. Study design for Walker et al. (27) and direct replications (*A*, ··· red boundary), conceptual replications (··· blue boundary) with reminder condition (*B*) and without reminder condition (*C*), and hypothesized underlying mechanisms and events predicted by reconsolidation theory (*D*, ··· green boundary). Critical time points for calculation of the reconsolidation score (RS) are indicated by triangle symbols. See main text for details.

Results

All data (Datasets S1 and S2) and analysis scripts are publically available on the Open Science Framework (https://osf.io/gpeq4/). All experiments and measures are reported. Unequal variances in between-subject comparisons were addressed by using Welch *t* tests. Statistical significance was defined at the 0.05 level.

Direct Replications (Experiments 1–4). Consistent with the original study, we observed time-dependent improvements in accuracy and speed across the course of the Training and Interference stages, and overnight between stages (Fig. 2 and *SI Results*). The critical index of a reconsolidation effect (the percentage difference between Old Sequence performance at the Reminder stage and Test stage; from herein Reconsolidation Score or RS; Fig. 1, triangles), completely contradicted the finding of the original study (27): we observed only small fluctuations around zero for both accuracy (Fig. 3*A*) and speed (Fig. 3*B*) in all four direct replication attempts (experiments 1–4). Minor procedural differences between the replications and the original study (variability in participant age and time of testing) were ruled out as potential confounds through additional analyses (*SI Results*). Averaged across experiments, mean RS declined by <1% for accuracy, compared with ~57% in the original study, and increased by ~4% for speed. One-sample *t* tests (one-tailed) indicated that none of the RS values (Table 1) obtained in the direct replications were significantly less than zero.

As the inherent limitations of null-hypothesis significance testing constrain the degree to which one can determine the strength of evidence in favor of the null hypothesis (34), we also conducted a Bayesian analysis that enabled us to quantify the evidence in favor of the null hypothesis H_0 (RS = 0) relative to the reconsolidation hypothesis H_1 (RS < 0). Specifically, we calculated directional Bayes factors (35) using an "objective" JZS prior (Cauchy distribution with scale $r = 1$). H_1 was based on the general prediction of reconsolidation theory that trace-dependent performance should be reduced following disrupted reconsolidation of the reactivated trace (2, 4, 6). In all experiments, Bayes factors (BF_{01}) (Table 1) were larger than 1, indicating greater evidentiary support for H_0 relative to H_1.

A primary goal of replication attempts is to facilitate more precise estimates of effect-size magnitude (36). However, in light

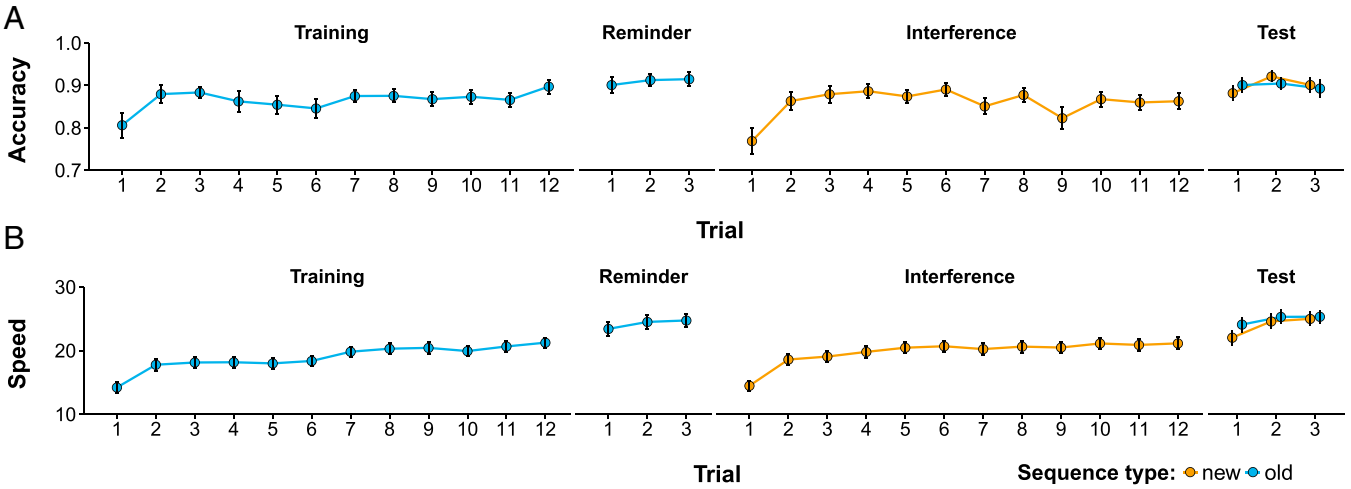

Fig. 2. Full study timeline showing mean accuracy (*A*; number of errors made relative to the number of complete sequences achieved) and mean speed (*B*; number of complete sequences achieved) by stage (Training, Reminder, Interference, and Test), trial, and sequence type, for experiments 1–4 (pooled). A full definition of these dependent variables is available in *SI Methods*. Error bars show ±SEM.

of the stark discrepancy between the finding observed in the original experiment (ref. 27, $n = 16$) and the four direct replications ($n = 64$), we focused on assessing the extent to which the collated evidence indicated that the phenomenon exists at all. That is to say, we aimed to establish whether the effect is qualitatively reproducible, as nonreplication will preclude attempts to derive greater quantitative precision in the estimation of the effect's magnitude.

Directional metaanalytic Bayes factors using t values for experiments 1–4 (Table 1) indicated greater evidentiary support for the null hypothesis (RS = 0) relative to the alternative hypothesis (RS < 0) for both accuracy ($BF_{01} = 5.743$) and speed ($BF_{01} = 36.027$). This pattern remained after incorporating an estimated t value for the original study (accuracy: $BF_{01} = 2.080$; speed: $BF_{01} = 31.317$). The complete absence of predicted outcomes across these four experiments suggests that the reconsolidation effect reported in group 7 of the original study (27) is not robust.

Conceptual Replications (Experiments 5–7).

In the second component of the replication battery, we aimed to evaluate the broader validity of reconsolidation-updating theory. These experiments also involved sequence learning within a 3-day reconsolidation protocol (Fig. 1), but used sequences similar in length and structure to phone numbers (experiments 5 and 7) or computer passwords (experiment 6), and required declarative (rather than procedural) recall at the Test stage. Performance was assessed using a string-matching algorithm that provided an index of similarity between the target sequence and the sequence entered by the user (*SI Methods*). This afforded a sensitive measure of partial (or "chunked") sequence knowledge. As the pattern of performance did not vary significantly between experiment 5 and experiment 6 ("Reminder experiments"), these data were pooled for display (Fig. 4) and subsequent analyses. Participants learned either a number (experiment 5) or letter (experiment 6) sequence to a criterion on day 1 (Training stage). On day 2, these sequences were recalled (Reminder stage) before new learning (Interference stage), and on day 3, recall of the sequences was evaluated (Test stage). In the No-Reminder control group (experiment 7), there was no Reminder stage, permitting a comparison of day 3 recall in the presence or absence of the day 2 reminder.

During the Training and Interference stages, all participants successfully reached the criterion of five consecutive errorless sequence recalls (i.e., a maximum similarity score of 1.0), indicating successful learning of both the Old and New Sequences (*SI Results*). A one-way repeated-measures ANOVA indicated significant changes across stages (Training, Reminder, Test) for the Reminder experiments [$F_{(2,90)} = 8.68$, $P < 0.001$]. Follow-up paired t tests (one-tailed) showed that there was significant decline from the Training stage (1.0) to the Reminder stage [mean (M) = 0.750, SD = 0.246; $t_{(31)} = 5.742$, $P < 0.001$], and from Reminder stage to the Test stage [M = 0.638, SD = 0.315; $t_{(31)} = 2.645$, $P < 0.001$]. The No-Reminder control group (experiment 7) enabled us to ascertain whether the observed recall impairments could be causally attributed to the time-dependent interaction of memory reactivation and interference as predicted by reconsolidation theory (2, 4, 6). Despite the absence of a Reminder stage, these participants also showed a substantial performance decrement from Training (1.0) to Test (M = 0.488, SD = 0.363). A paired-samples t test (two-tailed) confirmed that this decline was significant [$t_{(15)} = 5.646$, $P < 0.001$].

These findings imply that at least some of the recall impairment observed in the Reminder experiments was not contingent on the provision of a reminder-triggered reconsolidation process. Furthermore, a between-group comparison of Test stage performance indicated poorer recall in the No-Reminder experiment (M = 0.488, SD = 0.363) than in the Reminder experiments (M = 0.638, SD = 0.315), and a two-sample t test (one-tailed) indicated no significant difference [$t_{(26.59)} = -1.409$, $P = 0.915$]. Rather than inducing a state of increased susceptibility to interference, memory reactivation resulted in numerically less recall impairment of the Old Sequence at the Test stage relative to no memory reactivation, an effect in the opposite direction to that predicted by reconsolidation theory.

Discussion

Reconsolidation-updating theory suggests that retrieval of an existing trace in the human memory system can render that trace vulnerable to modification from postretrieval new learning. In the present investigation, we attempted to replicate and extend a critical finding (27) widely considered to provide a compelling demonstration of reconsolidation-mediated memory updating in humans. In four direct-replication attempts involving procedural recall and three conceptual-replication attempts involving declarative recall, we did not observe the critical impairment effects observed in the original study and predicted by reconsolidation theory (2, 4, 6).

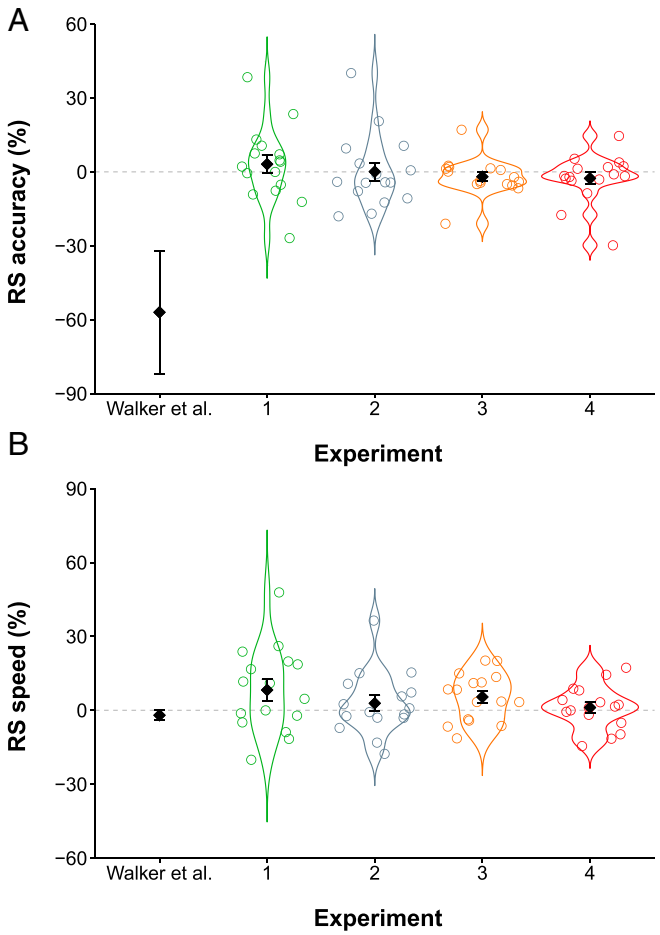

Fig. 3. Accuracy (*A*) and speed (*B*) reconsolidation scores (RSs) for Walker et al. (27); $n = 16$) and experiments 1–4 ($N = 64$). Black diamonds represent means, and error bars show SEM. Where raw data are available (experiments 1–4), individual participant scores (circles) and kernel density distributions are also depicted.

The findings of our direct replications are consistent with several recent investigations that used variations of the original paradigm (27), with either transcranial magnetic stimulation (37, 38) or new learning (39) interventions. Although the findings of these studies were interpreted as favorable evidence for reconsolidation theory, the expected Reminder–Test performance decrement was actually absent in most conditions, including close replications of the original study. When trace-dependent performance decrements were observed at short reminder lengths, they were modest, and rapidly recovered within the test session (39). It is difficult to reconcile the absence of performance impairments and the presence of recovery effects with the prediction of permanent trace modification (2). Thus, the outcomes of all three studies (37–39) are inconsistent with both the original study (27) and reconsolidation theory in general (2, 4, 6).

The findings of our conceptual replications cast further doubt on the veracity of claims that memory updating can be mediated by reconsolidation processes (12, 14–17). These experiments adhered to the canonical 3-day reconsolidation protocol and aimed to increase external validity through the use of sequences similar in structure to phone numbers or computer passwords. In addition, consistent with several studies in the human reconsolidation literature (e.g., refs. 10–12), participants completed a declarative recall task. Under these conditions, performance impairments occurred in the both the presence and absence of memory retrieval. Rather than triggering a state of heightened trace-vulnerability, retrieval actually led to numerically higher performance than in the no-retrieval control group. This finding is consistent with previous investigations of reconsolidation that found retrieval practice can afford some protection against interference (e.g., ref. 40), and a considerable body of evidence suggesting that retrieval aids rather than impairs subsequent recall (41).

Two notable aspects of human reconsolidation research are not directly addressed by the present investigation. First, there is evidence that postretrieval pharmacological interventions can attenuate emotional responding in a fear-conditioning paradigm (42). However, the reliability of these effects has also recently come under scrutiny (43). Similarly, initially promising findings based on using postretrieval extinction to disrupt reconsolidation in a fear-conditioning paradigm (17) have proved elusive in subsequent replication attempts (44, 45). It is striking that declarative recall of the conditioned stimulus–unconditioned stimulus contingency remains intact across these fear-conditioning studies, either in the presence or absence of effects on emotional responding.

Additionally, it has been suggested that "prediction error" is a necessary reconsolidation trigger (11, 46). If this were the case, it could explain the absence of reconsolidation effects in the present replications. However, it would also be surprising that a reconsolidation effect was observed in the original study because the reminder protocol required participants to practice the Old Sequence in its entirety, and thus presumably did not invoke prediction error. To justify an auxiliary theoretical assumption about prediction error, one would need to reconcile a considerable amount of contradictory evidence. For example, relative to controls, no impairment of declarative recall is observed in the aforementioned prediction error studies (11, 46), only attenuation of emotional responding (46), or ambiguous null effects on an indirect measure of trace integrity (retrieval-induced forgetting; ref. 11). Furthermore, reconsolidation-like effects have been reported when the reminder involves reinforced trials (and thus no prediction error) in both nonhuman animals (47) and humans (12), and impairment effects are absent even in studies where prediction error would be expected (44, 45). At present, therefore, it is unclear whether prediction error is either necessary or sufficient for reconsolidation effects to emerge.

Taken together, our findings cast doubt on the efficacy of new-learning interventions as a means for disrupting the reconsolidation of procedural or declarative memory in humans. The absence of reconsolidation effects in all four direct replications suggests that the considerable theoretical weight attributed to the original study (2, 4, 6, 13, 14) is unwarranted. Furthermore, the absence of retrieval-contingent impairment in the conceptual replications is inconsistent with the purported functional role of reconsolidation as an adaptive mechanism that underlies memory updating (12, 14–17). Replication will be an essential tool in future reconsolidation investigations as researchers seek to verify

Table 1. Direct replication RS statistics for accuracy and speed

Exp	DV	RS	SD	$t_{(15)}$	*P*	BF_{01}
1	Accuracy	3.13	14.95	0.84	0.79	8.97
	Speed	8.24	17.14	1.92	0.96	13.59
2	Accuracy	0.08	14.75	0.02	0.51	5.38
	Speed	2.85	12.74	0.89	0.81	9.23
3	Accuracy	−1.98	7.58	−1.04	0.16	1.90
	Speed	5.41	9.75	2.22	0.98	14.68
4	Accuracy	−2.64	9.88	−1.07	0.15	1.85
	Speed	1.12	8.70	0.52	0.69	7.52

BF_{01}, Bayes factor quantifying evidence in favor of the null hypothesis (RS = 0) relative to the reconsolidation hypothesis (RS < 0); DV, dependent variable; Exp, experiment; RS, mean reconsolidation score.

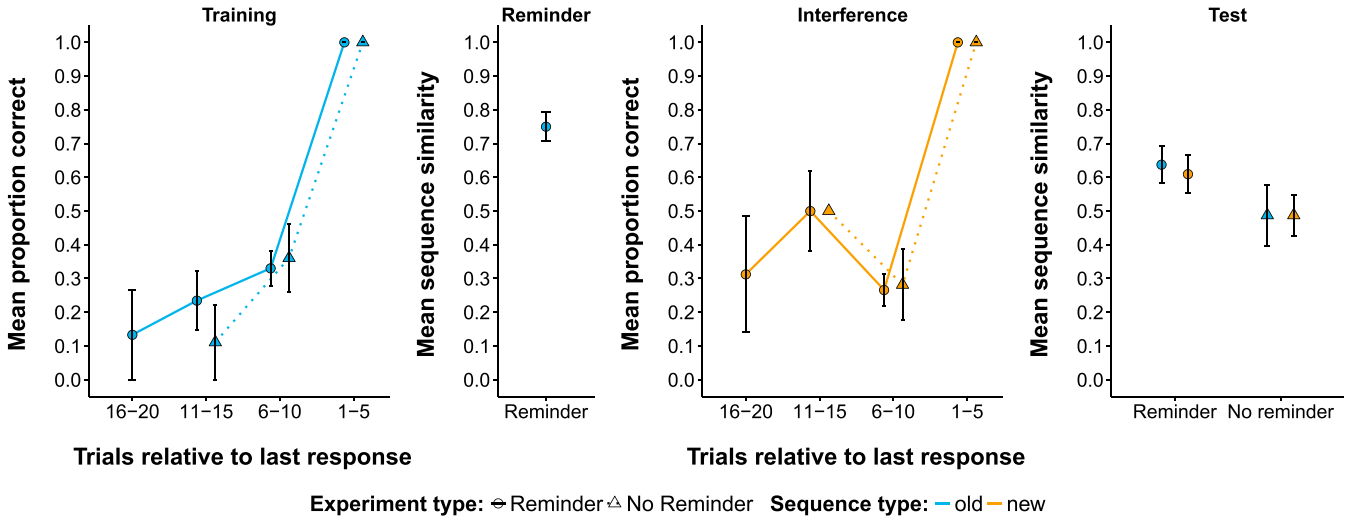

Fig. 4. Full study timeline showing performance in experiments 5 and 6 pooled (Reminder groups; $n = 32$), and experiment 7 (No-Reminder group; $n = 16$). The Training and Interference panels show mean proportion correct on $\text{RECALL}_{\text{Feedback}}$ trials across five trial bins plotted relative to participants' final response of the stage. All participants reached the performance criterion (five correct trials in a row) but required a different number of trials to do so (*SI Results*). The small number of participants who took more than 20 trials to reach criterion (Training: $n = 2$, maximum trials = 29; Interference: $n = 1$, maximum trials = 22) contribute to all relevant analyses. The Reminder and Test panels show mean sequence similarity between the target sequence and the user-entered sequence assessed on a single $\text{RECALL}_{\text{NoFeedback}}$ trial for each previously learned sequence (Old and New). Error bars show SEM.

the reliability of existing findings, identify genuine boundary conditions, and foster theoretical progress.

Methods

All experimental programs and verbatim materials are publically available on the Open Science Framework (https://osf.io/gpeq4/). Participants were recruited from the University College London (UCL) mixed-occupation subject pool and received either monetary compensation or course credits. All participants reported that they were right-handed and had no history of neurological, psychiatric, or sleep disorder. All participants provided informed consent and the study was approved by the local UCL ethics committee.

Direct Replications (Experiments 1–4).

Participants. Sixteen participants were randomly allocated to each of the four direct-replication experiments, affording a total sample size of 64 individuals (49 females; median age, 22 y; age range, 18–54 y). Two additional participants were excluded for typing an incorrect sequence at the Reminder stage, and four additional participants did not complete all three stages of the study.

Design. Participants performed a "finger-tapping" sequence learning task in three discrete sessions taking place on consecutive days (Fig. 1). Two five-digit sequences (X: 4–1–3–2–4; Y: 2–3–1–4–2) were assigned to be the Old Sequence and the New Sequence in counterbalanced order. On day 1, participants completed 12 Old Sequence trials (Training). On day 2, participants performed three Old Sequence trials (Reminder) immediately before 12 New Sequence trials (Interference). On day 3, participants completed three trials of both the Old Sequence and the New Sequence in counterbalanced order (Test). The dependent variables (see *SI Methods* for details) were the number of sequences completed during each 30-s trial ("speed") and the ratio of errors to speed ["accuracy"; 1 − (errors/speed)].

Procedure. Unless otherwise stated (*SI Methods*), the following procedures were used in all direct replications and precisely matched those reported in the original study (27). Ambiguous or missing information was clarified through contact with the senior author of the original research team. Participants were seated in front of a computer screen in a quiet room and used the four fingers of their left (nondominant) hand to respond using the four top-row numeric keys 1, 2, 3, and 4 of a standard keyboard. The task involved repeatedly tapping a five-element sequence that was displayed on the screen for 30 s (including on "test" trials), followed by 30 s of rest during which the sequence was absent. Key presses were acknowledged with white dots that accumulated on screen, but there was no feedback regarding response accuracy. A 30-s countdown timer was displayed during the rest phase to signal the approaching test phase. During the tapping phase, the screen background was green, and during the rest phase it was red. Participants were instructed to "tap out the sequence as quickly and accurately as possible." There was no within- or between-subjects timing variability in the original study because all sessions were conducted at 1:00 PM. In the present experiments, there was also no within-subject variability: participants completed sessions at precise 24-h intervals (±15 min); however, session times varied between participants (9:00 AM to 6:00 PM).

Conceptual Replications.

Participants. Sixteen participants were randomly allocated to each of the three conceptual-replication experiments, affording a total sample size of 48 individuals (38 females; median age, 22 y; age range, 18–52 y). Three additional participants were excluded as they did not complete all three stages of the study.

Design. Participants performed a sequence-learning task in three discrete sessions taking place on consecutive days (Fig. 1). Two 10-item sequences with independent grammars (*SI Methods*) were assigned to be the Old Sequence and the New Sequence in counterbalanced order. For experiments 5 and 7, the sequences were numbers (X: 1–4–6–3–2–9–5–0–8–7; Y: 2–6–5–7–0–1–9–4–3–8). For experiment 6, the sequences were letters (X: l–p–k–s–f–q–j–d–x–h; Y: j–f–l–d–q–x–k–h–p–s). On day 1, an adaptive test-feedback protocol was used to ensure that all participants could recall the Old Sequence unassisted five times in a row (Training). On day 2, participants in experiments 5 and 6 recalled and restudied the Old Sequence immediately before new learning (Reminder). All participants learned the New Sequence in the same manner as Old Sequence Training (Interference). On day 3, participants were asked to recall both sequences in counterbalanced order (Test). The dependent variable was a metric of the similarity between the target (Old/New) sequence at a given stage and the sequence entered by the user ("sequence similarity"; see *SI Methods* for details).

Procedure. Participants were seated in front of a computer screen in a quiet room and responded using a standard keyboard. On STUDY trials, participants were instructed to memorize the sequence while it was displayed on screen for 5 s. No response was required. On $\text{RECALL}_{\text{Feedback}}$ trials, participants were asked to enter the sequence from memory into 10 blank placeholders (_). Correctly entered items appeared in green. Entering an item in an incorrect order caused that item to flash in red and black (4 × 0.5-s flashes over 2 s) followed by replacement with the correct item, which flashed in green and black (4 × 0.5-s flashes over 2 s), and early termination of the trial. On $\text{RECALL}_{\text{NoFeedback}}$ trials, participants also had to enter the sequence from memory; however, the trial was not interrupted if they entered items in an incorrect order and they could make corrections if they wished. All items appeared in black so there was no feedback on these trials.

The Training and Interference stages involved iterative cycles of STUDY and $\text{RECALL}_{\text{Feedback}}$ trials starting with the former. Accurately entering the whole sequence on a $\text{RECALL}_{\text{Feedback}}$ trial led to additional $\text{RECALL}_{\text{Feedback}}$ trials. Failure to complete a $\text{RECALL}_{\text{Feedback}}$ trial resulted in a STUDY trial and

the cumulative $RECALL_{Feedback}$ counter was reset. When the participant had achieved five accurate $RECALL_{Feedback}$ trials in a row, the stage was terminated.

The Reminder stage involved a single $RECALL_{NoFeedback}$ trial followed by two STUDY trials. The Test stage involved two $RECALL_{NoFeedback}$ trials where participants were asked to "Recall the OLD sequence from day one and enter it on the next screen" and, separately, "Recall the NEW sequence from day two and enter it on the next screen." Participants completed sessions at precise 24-h intervals (±15 min); however, session times varied between participants (9:00 AM to 6:00 PM).

ACKNOWLEDGMENTS. We thank the UCL Institute of Making for assistance with apparatus for experiment 4. This work was supported by Economic and Social Research Council Grant ES/J500185/1.

1. Nader K, Schafe GE, Le Doux JE (2000) Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature* 406(6797):722–726.
2. Nader K, Hardt O (2009) A single standard for memory: The case for reconsolidation. *Nat Rev Neurosci* 10(3):224–234.
3. Nadel L, Land C (2000) Memory traces revisited. *Nat Rev Neurosci* 1(3):209–212.
4. Schwabe L, Nader K, Pruessner JC (2014) Reconsolidation of human memory: Brain mechanisms and clinical relevance. *Biol Psychiatry* 76(4):274–280.
5. Hui K, Fisher CE (2015) The ethics of molecular memory modification. *J Med Ethics* 41(7):515–520.
6. Tronson NC, Taylor JR (2007) Molecular mechanisms of memory reconsolidation. *Nat Rev Neurosci* 8(4):262–275.
7. Cahill L, McGaugh JL, Weinberger NM (2001) The neurobiology of learning and memory: Some reminders to remember. *Trends Neurosci* 24(10):578–581.
8. Riccio DC, Millin PM, Bogart AR (2006) Reconsolidation: A brief history, a retrieval view, and some recent issues. *Learn Mem* 13(5):536–544.
9. Millin PM, Moody EW, Riccio DC (2001) Interpretations of retrograde amnesia: Old problems redux. *Nat Rev Neurosci* 2(1):68–70.
10. Hupbach A, Gomez R, Hardt O, Nadel L (2007) Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learn Mem* 14(1-2):47–53.
11. Forcato C, Argibay PF, Pedreira ME, Maldonado H (2009) Human reconsolidation does not always occur when a memory is retrieved: The relevance of the reminder structure. *Neurobiol Learn Mem* 91(1):50–57.
12. Chan JCK, LaPaglia JA (2013) Impairing existing declarative memory in humans by disrupting reconsolidation. *Proc Natl Acad Sci USA* 110(23):9309–9313.
13. Schiller D, Phelps EA (2011) Does reconsolidation occur in humans? *Front Behav Neurosci* 5:24.
14. Lee JLC (2009) Reconsolidation: Maintaining memory relevance. *Trends Neurosci* 32(8):413–420.
15. Dudai Y (2009) Predicting not to predict too much: How the cellular machinery of memory anticipates the uncertain future. *Philos Trans R Soc Lond B Biol Sci* 364(1521): 1255–1262.
16. Hardt O, Einarsson EÖ, Nader K (2010) A bridge over troubled water: Reconsolidation as a link between cognitive and neuroscientific memory research traditions. *Annu Rev Psychol* 61(1):141–167.
17. Schiller D, et al. (2010) Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature* 463(7277):49–53.
18. Miller RR, Matzel LD (2006) Retrieval failure versus memory loss in experimental amnesia: Definitions and processes. *Learn Mem* 13(5):491–497.
19. McGaugh JL (2000) Memory—a century of consolidation. *Science* 287(5451):248–251.
20. Melton AW, Irwin JM (1940) The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Am J Psychol* 53(2): 173–203.
21. Loftus EF (1979) The malleability of human memory. *Am Sci* 67(3):312–320.
22. McGeoch JA (1942) *The Psychology of Human Learning: An Introduction* (Longmans, New York).
23. Tulving E (1974) Cue-dependent forgetting: When we forget something we once knew, it does not necessarily mean that the memory trace has been lost; it may only be inaccessible. *Am Sci* 62(1):74–82.
24. Eich JE (1980) The cue-dependent nature of state-dependent retrieval. *Mem Cognit* 8(2):157–173.
25. Capaldi EJ, Neath I (1995) Remembering and forgetting as context discrimination. *Learn Mem* 2(3-4):107–132.
26. Bouton ME (2002) Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biol Psychiatry* 52(10):976–986.
27. Walker MP, Brakefield T, Hobson JA, Stickgold R (2003) Dissociable stages of human memory consolidation and reconsolidation. *Nature* 425(6958):616–620.
28. Power AE, Berlau DJ, McGaugh JL, Steward O (2006) Anisomycin infused into the hippocampus fails to block "reconsolidation" but impairs extinction: The role of re-exposure duration. *Learn Mem* 13(1):27–34.
29. Lattal KM, Abel T (2004) Behavioral impairments caused by injections of the protein synthesis inhibitor anisomycin after contextual retrieval reverse with time. *Proc Natl Acad Sci USA* 101(13):4667–4672.
30. Eisenberg M, Dudai Y (2004) Reconsolidation of fresh, remote, and extinguished fear memory in Medaka: Old fears don't die. *Eur J Neurosci* 20(12):3397–3403.
31. Rosenthal R (1990) Replication in behavioral research. *J Soc Behav Pers* 5(4):1–30.
32. Simons DJ (2014) The value of direct replication. *Perspect Psychol Sci* 9(1):76–80.
33. Schmidt S (2009) Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Rev Gen Psychol* 13(2):90–100.
34. Dienes Z (2011) Bayesian versus orthodox statistics: Which side are you on? *Perspect Psychol Sci* 6(3):274–290.
35. Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G (2009) Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev* 16(2):225–237.
36. Cumming G (2012) *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-analysis* (Routledge, New York).
37. Censor N, Horovitz SG, Cohen LG (2014) Interference with existing memories alters offline intrinsic functional brain connectivity. *Neuron* 81(1):69–76.
38. Censor N, Dimyan MA, Cohen LG (2010) Modification of existing human motor memories is enabled by primary cortical processing during memory reactivation. *Curr Biol* 20(17):1545–1549.
39. de Beukelaar TT, Woolley DG, Wenderoth N (2014) Gone for 60 seconds: Reactivation length determines motor memory degradation during reconsolidation. *Cortex* 59: 138–145.
40. Potts R, Shanks DR (2012) Can testing immunize memories against interference? *J Exp Psychol Learn Mem Cogn* 38(6):1780–1785.
41. Roediger HL, Butler AC (2011) The critical role of retrieval practice in long-term retention. *Trends Cogn Sci* 15(1):20–27.
42. Kindt M, Soeter M, Vervliet B (2009) Beyond extinction: Erasing human fear responses and preventing the return of fear. *Nat Neurosci* 12(3):256–258.
43. Bos MGN, Beckers T, Kindt M (2014) Noradrenergic blockade of memory reconsolidation: A failure to reduce conditioned fear responding. *Front Behav Neurosci* 8:412.
44. Golkar A, Bellander M, Olsson A, Öhman A (2012) Are fear memories erasable?—reconsolidation of learned fear with fear-relevant and fear-irrelevant stimuli. *Front Behav Neurosci* 6:80.
45. Kindt M, Soeter M (2013) Reconsolidation in a human fear conditioning study: A test of extinction as updating mechanism. *Biol Psychol* 92(1):43–50.
46. Sevenster D, Beckers T, Kindt M (2014) Prediction error demarcates the transition from retrieval, to reconsolidation, to new learning. *Learn Mem* 21(11):580–584.
47. Duvarci S, Nader K (2004) Characterization of fear memory reconsolidation. *J Neurosci* 24(42):9269–9275.
48. Lakens D (2013) Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Front Psychol* 4:863.
49. Suzuki A, et al. (2004) Memory reconsolidation and extinction have distinct temporal and biochemical signatures. *J Neurosci* 24(20):4787–4795.
50. Van der Loo M (2014) The stringdist package for approximate string matching. *R J* 6(1):111–122.