Prediction Errors Disrupt Hippocampal Representations and Update Episodic Memories

*Alyssa H. Sinclair[1,2] Grace M. Manalili[2], Iva K. Brunec[3,4],

R. Alison Adcock[1], & Morgan D. Barense[2,5]

[1] Duke University, *Department of Psychology & Neuroscience*, Durham, NC, 27705, USA

[2] University of Toronto, *Department of Psychology*, Toronto, ON, M5S 3G3, Canada

[3] Temple University, *Department of Psychology*, Philadelphia, PA, 19122, USA

[4] University of Pennsylvania, *Department of Psychology*, Philadelphia, PA, 19104, USA

[5] Baycrest Hospital, *Rotman Research Institute*, Toronto, ON, M64 1W1, Canada

* Corresponding Author: allie.sinclair@duke.edu

Significance

1   Our brains draw on memories to predict the future; when our predictions are incorrect,

2   we must modify our memories to improve future predictions. Past studies have demonstrated that

3   the hippocampus signals *prediction error*, or surprise, but have not linked this neural signal to

4   memory updating. Here, we uncover this missing connection: We show that prediction errors

5   change the role of the hippocampus, reversing the relationship between hippocampal activation

6   and memory outcomes. We examine the mechanisms of this shift in neural processing, showing

7   that prediction errors disrupt the temporal continuity of hippocampal patterns. We propose that

8   prediction errors disrupt sustained representations and destabilize memories, enabling updating.

9   Our findings bear implications for improving education, understanding eyewitness memory

10   distortion, and treating pathological memories.

Abstract

1

2   The brain supports adaptive behavior by generating predictions, learning from errors, and

3   updating memories to incorporate new information. *Prediction error,* or surprise, triggers

4   learning when reality contradicts expectations. Prior studies have shown that the hippocampus

5   signals prediction errors, but have never linked this neural signal to memory updating. Here, we

6   uncover new mechanisms that reveal this missing link. In a human fMRI study, we elicited

7   mnemonic prediction errors by interrupting familiar narrative videos immediately before the

8   expected endings. We found that prediction error reversed the relationship between univariate

9   hippocampal activation and memory: greater hippocampal activation predicted memory

10   preservation after expected endings, but memory updating after surprising endings. In contrast to

11   previous studies, we showed that univariate activation was insufficient for understanding

12   hippocampal prediction error signals. We explained this surprising finding by tracking both the

13   evolution of hippocampal activation patterns and connectivity between the hippocampus and

14   neuromodulatory regions. We found that hippocampal activation patterns stabilized as each

15   narrative episode unfolded, consistent with sustaining episodic representations. Prediction errors

16   disrupted these sustained representations, and the degree of disruption predicted memory

17   updating. The relationship between hippocampal activation and subsequent memory depended

18   on concurrent basal forebrain activation, providing new evidence about how cholinergic

19   modulation may regulate attention and memory. We conclude that prediction errors create

20   conditions that favor memory updating, prompting the hippocampus to abandon ongoing

21   predictions and make memories malleable.

Introduction

In daily life, we continuously draw on past experiences to predict the future. Expectation and surprise shape learning across many situations, such as when we discover misinformation in the news, receive feedback on an exam, or make decisions based on past outcomes. When our predictions are incorrect, we must update our mnemonic models of the world to support adaptive behavior. *Prediction error* is a measure of the discrepancy between expectation and reality; this surprise signal is both evident in brain activity and related to learning success (1–6). The brain dynamically reconstructs memories during recall, recreating and revising past experiences based on current information (7). The intuitive idea that surprise governs learning has long shaped our understanding of memory, reward learning, perception, action, and social behavior (2, 8–14). Yet, the neural mechanisms that allow prediction error to update memories remain unknown.

Past research has implicated the hippocampus in each of the mnemonic functions required for learning from prediction errors: retrieving memories to make predictions, identifying discrepancies between past and present, and encoding new information (2, 15–20). Functional MRI (fMRI) studies have shown that hippocampal activation increases after predictions are violated; this surprise response has been termed *mismatch detection* (18, 19, 21–23), or *mnemonic prediction error* (20). These past studies have shown that the hippocampus *detects* prediction errors. Several theoretical frameworks have hypothesized that this hippocampal prediction error signal could update memories (17, 20, 24–27), but this crucial link for understanding how we learn from error has not yet been demonstrated.

What mechanisms could link hippocampal prediction errors to memory updating? A leading hypothesis is that prediction errors shift the focus of attention and adjust cognitive processing (20, 28–32). After episodes that align with expectations, we should continue

1   generating predictions and shift attention *internally*, sustaining and reinforcing existing

2   memories. However, after prediction errors, we should reset our expectations and shift attention

3   *externally*, preparing to encode new information and update memories. Consistent with this idea,

4   mnemonic prediction errors have been shown to enhance the hippocampal *input* pathway that

5   supports encoding, but suppress the *output* pathway that supports retrieval (20). We propose that

6   surprising events may also change intrinsic hippocampal processing, changing the effect of

7   hippocampal activation on memory outcomes.

8   A separate body of research has demonstrated how memories can be updated. Animal and

9   human research on *reconsolidation* has shown that recalling a memory can temporarily

10   destabilize the memory trace in the brain, creating a window of opportunity for memory updating

11   (33–36). Prediction error is crucial for destabilizing and updating memories (1, 3, 37).

12   Reconsolidation paradigms have elicited prediction errors by reactivating memories with

13   incomplete reminders, which are imperfect replications of the original encoding experience (e.g.,

14   presenting a conditioned stimulus without the expected outcome) (1, 3, 38, 39). Yet, past

15   reconsolidation studies have not measured the implied neural prediction error signals (in the

16   hippocampus or elsewhere) or linked them to memory updating.

17   Neuromodulation may be a critical factor that regulates hippocampal processing and

18   enables memory updating. Currently, there is mixed evidence supporting two hypotheses:

19   acetylcholine and/or dopamine could act upon the hippocampus to regulate processing after

20   surprising events (24–27, 29, 31, 40). Several models have proposed that acetylcholine from the

21   medial septum (within the basal forebrain) regulates the balance between input and output

22   pathways in the hippocampus (27–29, 41–44), thus allowing stored memories to be compared

23   with perceptual input (31, 44, 45). After prediction errors, acetylcholine release could change

1   hippocampal processing and enhance encoding or memory updating (26, 29, 40, 43, 45). On the

2   other hand, dopamine released from the ventral tegmental area (VTA), if transmitted to the

3   hippocampus, could also modulate hippocampal plasticity after prediction errors. Past studies

4   have shown that the hippocampus and VTA are co-activated after surprising events (46, 47).

5   Other work has shown that co-activation of hippocampus and VTA predicts memory encoding

6   and integration (48–51). Similarly, reconsolidation studies have suggested that both dopamine

7   and acetylcholine may contribute to destabilizing memories, especially hippocampally-

8   dependent memories (37, 52, 53). Overall, extant evidence suggests that basal forebrain and

9   VTA neuromodulation are both candidate mechanisms for regulating hippocampal processing

10   and memory updating.

11   In the present study, we adapted methods from reconsolidation paradigms (which are

12   optimized to update memories) to identify the missing link between hippocampal prediction error

13   signals and memory updating. Using an fMRI task with human participants, we examined trial-

14   wise hippocampal responses to prediction errors during narrative videos. During the Day 1

15   encoding session, participants viewed 70 full-length videos that featured narrative events with

16   salient endings (e.g., a baseball batter hitting a home run) (Figure 1A). During the Day 2

17   reactivation session, participants watched the videos again (Figure 1B). We elicited mnemonic

18   prediction errors by interrupting half of the videos immediately before the expected narrative

19   ending (e.g., the video ends while the baseball batter is mid-swing). These surprising

20   interruptions were comparable to the prediction errors employed in reconsolidation studies (1).

21   Half of the videos were presented in Full-length form (identical to the encoding session), and

22   half were presented in Interrupted form (eliciting prediction error). Reconsolidation group

23   participants (n = 24) completed the Day 2 session while undergoing an fMRI scan, whereas

1   Immediate control group participants (n = 24) completed the study in a behavioral testing room

2   and were not scanned. Our primary fMRI analyses examined the period immediately following

3   the offset of Full and Interrupted videos (Post-Video fixation period) during the Day 2 session in

4   the Reconsolidation group. Importantly, this design compares neural responses to surprising and

5   expected video endings while controlling for visual and auditory input.

6   Lastly, participants completed a memory test in the form of a structured interview (Figure

7   1C). On each trial, participants were cued with the name of the video and recalled the narrative.

8   The experimenter then probed for further details with pre-determined questions (e.g., "Can you

9   describe the baseball batter's ethnicity, age range, or clothing?"). Our critical measure of

10 memory updating was *false memories*, because the presence of a false memory indicates that the

11 original memory was changed in some way. Although it can be adaptive to update real-world

12 memories by incorporating relevant new information, we expected that our laboratory paradigm

13 would induce false memories because participants would integrate interfering details across

14 similar episodes (1, 7). Because we were interested in false memories as a measure of memory

15 updating, we instructed participants not to guess and permitted them to skip details they could

16 not recall. In the Reconsolidation group, participants completed the memory test on Day 3, 24

17 hours after the reactivation session. In the Immediate control group, participants completed the

18 memory test on Day 2, immediately after the reactivation session (Figure 1D). Reconsolidation

19 theory states that updating memories requires a delay, because re-stabilizing a memory trace

20 involves hours of protein synthesis (36, 54). Therefore, the Immediate control group should not

21 exhibit any memory effects that require protein synthesis-dependent reconsolidation.

22   Our approach allowed us to test several questions set up by the prior literature. First, to

23 link hippocampal prediction error signals to memory updating, we used a reconsolidation-

1   inspired paradigm to create and then distort naturalistic episodic memories. Second, to link

2   memory updating to hippocampal processing, we used multivariate analyses to track how

3   episodic representations were sustained or disrupted over time. Third, to test hypotheses about

4   neuromodulatory mechanisms, we related hippocampal processing and memory updating to

5   activation in the basal forebrain and VTA.

A Encoding Phase: Example Stimulus Video
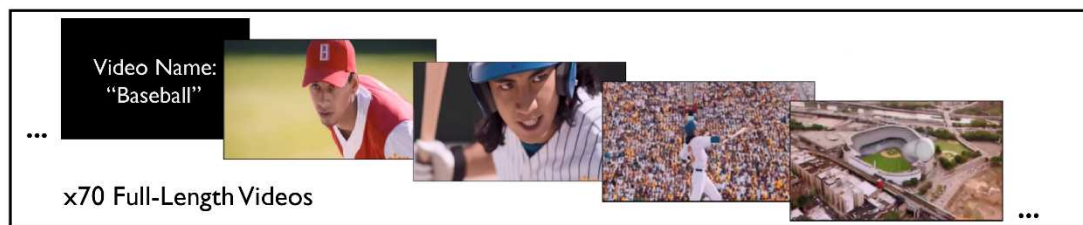
B Reactivation Phase: Example Trial

C Test Phase: Example Memory Test

Experimenter: The next video is "Baseball." Can you describe the main event of the video?
Participant: Okay, so they're **in a stadium,** and there are **lots of people watching.** The **pitcher throws the ball** and the batter **hits it out of the park.**
Experimenter: Can you describe the baseball batter? Age range, hair color, ethnicity, or clothing?
Participant: He looked **East Asian,** in his **mid-40s.** He was **wearing a green uniform.**
Experimenter: Do you remember hair color?
Participant: No, I don't remember.

Legend: Correct Details   False Memories
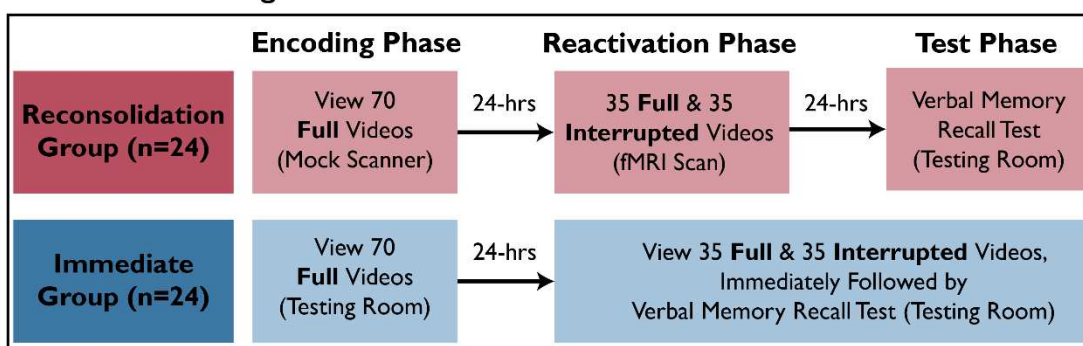
D Overview of Paradigm

1   *Figure 1*. Overview of experimental paradigm. A) During the Day 1 Encoding session, all videos
2   were presented in Full-length form. Here we show frames from a stimulus video named
3   "Baseball", depicting a home run. B) During the Day 2 Reactivation session, participants viewed
4   the videos again, but half were interrupted to elicit prediction error. Participants were cued with
5   the video name, watched the video (Full or Interrupted), and then viewed a fixation screen. The
6   "Baseball" video was interrupted when the batter was mid-swing. fMRI analyses focused on the
7   Post-Video fixation periods after each video (red highlighted box). Thus, visual and auditory
8   stimulation were matched across Full and Interrupted conditions, allowing us to compare Post-
9   Video neural activation while controlling for perceptual input. C) During the Test session,
10 participants answered structured interview questions about all 70 videos, and were instructed to
11 answer based on their memory of the Full video originally shown during encoding. Here we
12 show example text illustrating the memory test format and scoring of correct details (our
13 measure of memory preservation) and false memories (our measure of memory updating,
14 because false memories indicate that the memory has been modified). The void response ("I
15 don't remember") is not counted as a false memory. D) Overview of the experiment. All
16 participants completed Encoding, Reactivation, and Test Phases of the study. The
17 Reconsolidation group did the Test Phase 24 hours after Reactivation, whereas the Immediate
18 control group did the Test Phase immediately after Reactivation, in order to investigate whether
19 memory modification required a delay. Only the Reconsolidation group was scanned.

20   **Results**

21 **Behavioral Results**

22   We transcribed and scored memory tests for two key measures: number of unique *correct*

23 *details* (Figure 2A) and *false memories* (Figure 2B) – reflecting memory preservation and

24 updating, respectively. We also collected *confidence ratings* and scored the number of *forgotten*

25 *videos* (Supplementary Information, *Confidence and Forgetting*, Figure S1). We defined false

26 memories as distorted details that the participant recalled from the episode (e.g., "The pitcher

27 wore a green hat"). Void responses (e.g., "I don't remember") were not counted as false

28 memories, but were missed opportunities to earn points for correct details. Importantly, our

29 measures for correct details and false memories were independent; there was no limit to the

30 number of details a participant could recall about a video, and each detail was scored as correct

31 or false. We conducted linear mixed-effects regression to predict memory outcomes from the

1   fixed factors *group* (Reconsolidation and Immediate) and *reactivation type* (Full and

2   Interrupted). In all models, we included random effects to account for by-subject and by-video

3   variability (Supplementary Information, *Supplementary Methods, Linear Mixed Effects*

4   *Regression*).

5   ***Prediction Error Increased Correct Details***

6   We found that prediction errors during memory reactivation enhanced recall of correct

7   details (Figure 2A), such that participants in both groups reported more correct details for

8   Interrupted videos than Full videos ($\beta = -0.07$, 95% CI [-0.12, -0.02], $t = -2.75$, $p = .008$) (Table

9   S1A). Even though the video endings were omitted, prediction errors strengthened and preserved

10   existing memories. Participants in the Reconsolidation group recalled fewer correct details than

11   participants in the Immediate group ($\beta = 0.16$, 95% CI [0.01, 0.31], $t = 2.16$, $p < .036$), likely

12   because the Reconsolidation group completed the memory test after a 24-hour delay. There was

13   no interaction between group and reactivation type ($\beta = -0.01$, 95% CI [-0.04, 0.02], $t = -0.68$, p

14    $= .495$), indicating that the effect of prediction error enhancing correct details did not require a

15   delay.

16   ***Prediction Error Increased False Memories***

17   We found that prediction errors selectively increased false memories after a 24-hour

18   delay in the Reconsolidation group, replicating our past behavioral results (39) (significant

19   interaction between reactivation type and group ($\beta = 0.04$, 95% CI [0.01, 0.07], $t = 2.61$, $p =$

20    $.010$) (Figure 2B, Table S1B). In other words, Interrupted videos increased false memories in the

21   Reconsolidation group ($\beta = -0.09$, $p < .001$), but not the Immediate group ($\beta = -0.01$, $p = .437$) .

22   We also found main effects of group ($\beta = -0.36$, 95% CI [-0.43, -0.29], $t = -9.68$, $p < .001$), and

1 reactivation type (β = -0.05, 95% CI [-0.08, -0.02], t = -3.31, p = .001), both driven by the effect

2 of prediction error increasing false memories in the Reconsolidation group.

3 To ensure that the effect of prediction error on memory updating was not driven by the

4 first few trials (which are presumably most surprising), we also conducted a supplementary

5 analysis that included variables for trial number and interactions (Table S2). There was no main

6 effect of trial number on false memories (β = 0.01, 95% CI [-0.03, 0.04], t = 0.48, p = .730), no

7 interaction between trial number and reactivation type (β = 0.00, 95% CI [-0.03, 0.03], t = -0.09,

8 p = .997), and no interaction among trial number, reactivation type, and group (β = 0.00, 95% CI

9 [-0.03, 0.03], t = -0.40, p = .987). Overall, these null results suggest that the effect of Interrupted

10 videos on false memories did not diminish over the course of the experiment.

11 In sum, our behavioral results showed a novel dissociation between reinforcing and

12 updating memories: Prediction errors during memory reactivation strengthened memories,

13 evident both immediately and after a delay (Figure 2A). However, memory updating (as revealed

14 by false memories) was not evident until after a delay (Figure 2B), as predicted by

15 reconsolidation theory.

16 ***Surprise Ratings and Semantic Similarity Predicted False Memories***

17 Expanding on the results reported above, we recruited an independent sample to watch

18 the videos and rate (on a 5-point Likert scale) the degree of surprise elicited by the narrative

19 interruptions (Supplementary Information, *Supplementary Methods, Online Video Ratings*). We

20 found that whereas surprise ratings were unrelated to correct details recalled by our laboratory

21 sample (Table S3), surprise was related to false memories: There was a significant interaction

22 such that more surprising videos were associated with more false memories, selectively in the

23 Reconsolidation group (β = -0.03, 95% CI [-0.06, -0.01], t = -2.07, p = .039). However, the

1   three-way interaction among surprise ratings, group, and reactivation type was not significant (β

2    $= 0.01$, 95% CI [-0.02, 0.04], $t = 0.75$, $p = .451$). Because our surprise ratings were collected

3   from a separate online sample, this measure may not be sensitive enough to detect the expected

4   three-way interaction when applied to the laboratory sample.

In the current study, we index memory updating in terms of *false* memories; however,

incorporating relevant information into memory can be an adaptive function. We hypothesized

that our paradigm would induce false memories because information would be integrated across

semantically-related episodes. To test this hypothesis, we quantified semantic similarity among

the 70 videos with a text-based analysis (Supplementary Information, *Supplementary Methods*,

Scoring of Memory Tests). All model parameters are reported in Table S4. We found that videos

that were more semantically-similar to other videos in the stimulus set produced more false

memories ($\beta = 0.11$, 95% CI [0.04, 0.19], $t = 3.09$, $p = .003$) (Table S4). There was an interaction

between semantic similarity and group predicting false memories ($\beta = -0.04$, 95% CI [-0.07, -

0.01], $t = -2.77$, $p = .006$), such that the effect of semantic similarity was stronger in the

Reconsolidation group ($\beta = 0.16$, $z = 3.90$, $p < .001$) than in the Immediate group ($\beta = 0.07$, $z =$

1.80, $p = .073$). There was also an interaction between semantic similarity and reactivation type

predicting false memories ($\beta = -0.03$, 95% CI [-0.06, -0.01], $t = -2.08$, $p = .038$), such that the

effect of semantic similarity was stronger for Interrupted videos ($\beta = 0.15$, $z = 3.44$, $p < .001$)

than for Full videos ($\beta = 0.08$, $z = 2.18$, $p = .030$). However, the three-way interaction among

reactivation type, group, and semantic similarity was not significant ($\beta = 0.02$, 95% CI [-0.02,

0.05], $t = 0.99$, $p = .323$). Overall, these results suggest that interfering details from semantically-

related videos distorted memories; this interference effect was strongest for Interrupted videos

and for participants in the Reconsolidation group. Consistent with an adaptive updating process, memories may have been updated with relevant information from other videos.
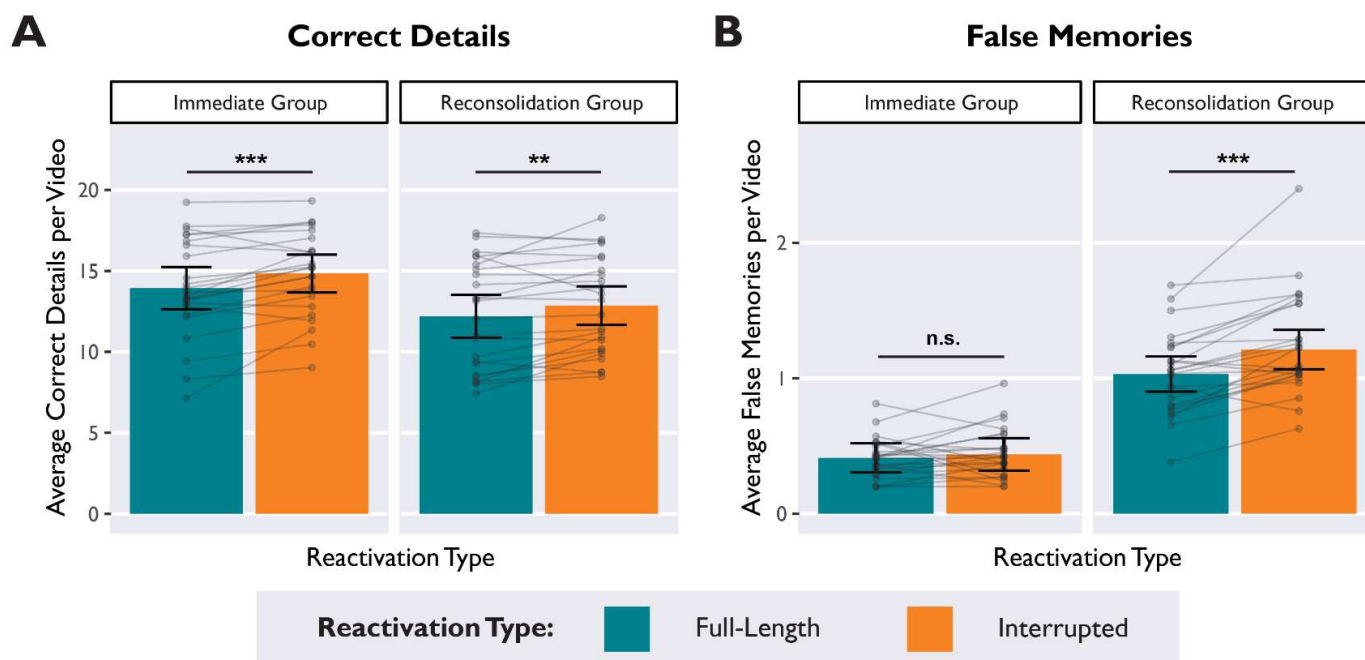

Figure 2. Prediction errors strengthened and updated memories over distinct time-courses. A) In both groups, average Correct Details were higher for videos that were Interrupted during memory reactivation, demonstrating that prediction error can strengthen memory recall both immediately and after a delay. B) In the Reconsolidation group (but not the Immediate group), average False Memories were higher for videos that were Interrupted during memory reactivation. This interaction demonstrates that prediction error enabled memory updating, but only after a delay that permitted reconsolidation. Bars depict estimated marginal means from a linear mixed effects model. Subject averages are overlaid on top to display the distribution: Dots indicate average scores by-participant, and lines connect within-subjects measures. Error bars depict 95% confidence intervals. * *p* < .05, ** *p* < .01, *** *p* < .001.

Univariate fMRI Results

The primary aim of our univariate fMRI analyses was to test the following questions: (1) Is hippocampal activation related to *reactivation type* (Full vs. Interrupted) and memory updating as indexed by subsequent *false memories?* (2) If so, does activation in the *basal forebrain* or the *VTA* moderate the relationship between hippocampal activation and memory updating?

1   We analyzed the blood oxygen level-dependent (BOLD) signal from the 24 subjects in

2   the Reconsolidation group (the Immediate group was not scanned). Our analyses focused on the

3   fixation screen presented during the Post-Video period immediately after each video offset. The

4   narrative ending of each video was either as-expected (Full) or a surprising prediction error

5   (Interrupted). Whole-brain mass univariate results are provided in the Supplementary

6   Information (*Whole-Brain Analysis*, Table S5, Figure S2). Importantly, we controlled for

7   differences in visual and auditory input by analyzing neural activation during the Post-Video

8   fixation period when visual and auditory input was identical across conditions (Figure 2B).

Some past studies have shown that prediction error signals are stronger in left

hippocampus and anterior hippocampus (18, 20, 21, 55), whereas posterior hippocampus is more

sensitive to video offsets (56). Other studies have shown that anterior and posterior hippocampus

parse continuous experience at different timescales (57, 58). On the basis of these findings, we

tested separate (non-overlapping) ROIs for left anterior, right anterior, left posterior, and right

posterior hippocampus (Supplementary Information, *Supplementary Methods, ROI Masks*).

Activation estimates from these four ROIs are included within each model to test for left/right

and anterior/posterior differences.

9   ***The Effect of Hippocampal Activation on Memory Depended on Prediction Error***

10   We used single-trial modelling to relate post-video hippocampal activation to subsequent

11   false memories. For our univariate analyses, we modelled a 2s impulse during the Post-Video

12   period (fixation screen), convolved with the canonical double-gamma hemodynamic response

13   function and phase-shifted 2s after video offset. This 2s shift targeted the peak Post-Video

14   hippocampal response identified in previous studies (59, 60). We isolated BOLD activation

15   during the Post-Video period on each trial and averaged parameter estimates across all voxels

1   within each hippocampal ROI (Methods, *Univariate fMRI Analyses*). Using linear mixed-effects

2   regression, we predicted trial-wise hippocampal activation from the following variables:

3   *reactivation type* (Full vs. Interrupted), *false memories* (continuous measure), *hemisphere* (left

4   vs. right), *axis* (anterior vs. posterior), and all interactions.

5   We found a significant interaction between reactivation type and subsequent false

6   memories associated with hippocampal activation (β = -0.06, 95% CI [-0.09, -0.04], t = -4.51, p

7   < .001) (Figure 3A) (Table S6). As predicted, this interaction demonstrated that the relationship

8   between hippocampal activation and subsequent memory differed between conditions. After Full

9   videos, greater hippocampal activation was associated with fewer subsequent false memories

10   (Figure 3A, blue; β = -0.07, z = -2.78, p = .006), consistent with the idea that the hippocampus

11   reinforces memory for episodes that just concluded (59–61). However, we observed the opposite

12   effect when events were surprising. After Interrupted videos, greater hippocampal activation was

13   associated *more* subsequent false memories (Figure 3A, orange; β = 0.05, z = 2.26, p = .024),

14   consistent with the idea that surprise drives memory updating. Overall, this interaction

15   demonstrated that the *same amount* of hippocampal activation predicted different memory

16   outcomes depending on whether the video was Full (fewer false memories) or Interrupted (more

17   false memories).

18   Neither the main effect of reactivation type on hippocampal activation (β = 0.04, 95% CI

19   [-0.01, 0.09], t = 1.64, p = .106), nor the main effect of false memories on hippocampal

20   activation (β = -0.01, 95% CI [-0.05, 0.03], t = -0.43, p = .673) was significant. These null results

21   underscore the value of examining the effect of prediction error on both hippocampal activation

22   and memory outcomes. Additionally, there were no significant main effects of hemisphere or

1   long axis position, nor any other significant interactions, indicating that effects did not differ

2   across hippocampal ROIs. Parameter estimates for all variables are provided in Table S6.

3   Lastly, although our primary research questions pertained to false memories, we also

4   tested a model that included both false memories and correct details (Table S7). We found that

5   correct details also interacted with reactivation type to predict hippocampal activation, but

6   contrasted with our results for false memories (β = 0.03, 95% CI [0.01, 0.06], t = 2.31, p = .021).

7   After Full videos, hippocampal activation was positively associated with correct details but

8   negatively associated with false memories. After Interrupted videos, these associations reversed

9   direction: hippocampal activation was positively associated with false memories but negatively

10   associated with correct details. However, the effect of hippocampal activation on correct details

11   was weaker than for false memories, and the slopes for each reactivation type did not

12   significantly differ from zero when tested independently (Full: β = 0.01, z = 0.31, p = .759;

13   Interrupted: β = -0.06, z = -1.90, p = .058). Importantly, the effect of hippocampal activation on

14   false memories remained unchanged after controlling for correct details (Full: β = -0.07, z = -

15   2.52, p = .012; Interrupted: β = 0.05, z = 2.03, p = .043), demonstrating that the effect of memory

16   updating was distinct from recall success for correct details (Table S7).

17   ***Hippocampal-Basal Forebrain Connectivity Predicted Memory Outcomes***

18   Next, we tested hypotheses about neuromodulatory mechanisms by examining activation

19   in the basal forebrain (which contains the medial septal nucleus, the primary source of

20   acetylcholine in the hippocampus) (29, 31, 43) and the VTA (which contains dopaminergic

21   neurons that project to the hippocampus) (16, 24, 62). First, we added a parameter to the model

22   described above to examine average basal forebrain activation during the Post-Video period (i.e.,

23   the 2-second segment of the Post-Video fixation period, consistent with modelling of

1   hippocampal activation). We also tested whether basal forebrain activation interacted with

2   *reactivation type*, *false memories, hemisphere,* or *axis* to predict hippocampal activation.

3   There was a significant three-way interaction among basal forebrain activation,

4   reactivation type, and false memories predicting hippocampal activation (β = -0.05, 95% CI [-

5   0.07, -0.02], t = -3.34, p < .001) (Figure 3B). This interaction demonstrated that the relationship

6   between hippocampal activation and subsequent memory (Figure 3A) was evident only when the

7   basal forebrain was also strongly activated (Figure 3B, right) (Full: β = -0.14, z = -4.3, p < .001;

8   Interrupted: β = 0.07, z = 2.4, p = .015). When basal forebrain activation was weak, hippocampal

9   activation was unrelated to memory (Figure 3B, left) (Full: β = 0.01, z = 0.25, p = .800;

10   Interrupted: β = 0.04, z = 1.24, p = .213). There was also a four-way interaction among

11   hemisphere, basal forebrain activation, reactivation type, and false memories predicting

12   hippocampal activation; this indicated that the three-way interaction described above was

13   stronger in right hippocampus than in left hippocampus (β = -0.04, 95% CI [-0.08, -0.01], t = -

14   2.32, p = .021). There was no effect of long axis position. All parameter estimates are provided

15   in Table S8.

16   Next, we examined the role of VTA activation. We modified the model described above

17   by replacing the basal forebrain activation variable (and interactions) with VTA activation

18   parameters. We found that VTA activation was positively related to hippocampal activation

19   during the Post-Video period (β = 0.16, 95% CI [0.05, 0.28], t = 2.90, p = .008). However, there

20   was no interaction among VTA activation, reactivation type, and false memories (β = 0.002,

21   95% CI [-0.05, -0.05], t = 0.06, p = .953). Thus, there was no evidence that VTA moderated the

22   effect of hippocampal activation on memory. All parameter estimates are provided in Table S9.

1   In a separate model, we also tested for a main effect of prediction error on VTA or basal

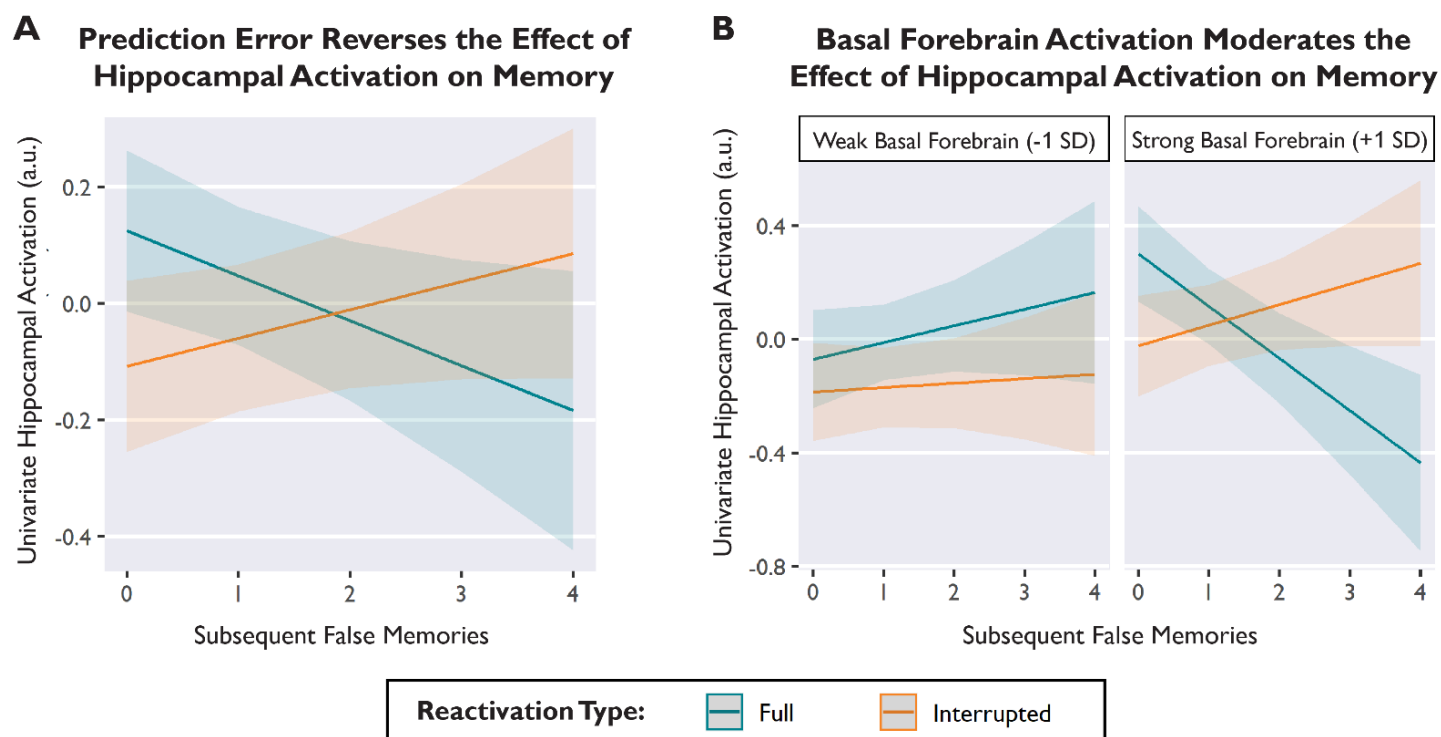2   forebrain activation, and found no significant effects (Table S10).

Figure 3. Prediction error reversed the relationship between average hippocampal activation (arbitrary units, a.u.) and subsequent memory, and this effect depended on concurrent basal forebrain activation. A) After Full-length videos, hippocampal activation was associated with memory *preservation*, predicting fewer false memories (blue). After Interrupted videos, hippocampal activation was associated with memory *updating*, predicting more false memories (orange). B) The effect of prediction error on hippocampal activation and memory was observed only when basal forebrain activation was strong (right). When basal forebrain activation was weak, hippocampal activation was unrelated to memory (left). Basal forebrain activation is binned (weak vs. strong) for visualization, but statistical models used a continuous variable. Lines depict model-predicted estimates, and shaded bands depict the 95% confidence interval. Model-derived estimates are shown instead of individual data points in order to show within-subjects effects, while controlling for subject and stimulus variance.

Multivariate fMRI Results

3   Overall, our univariate results suggest that prediction error changed the role of the

4   hippocampus: the same magnitude of univariate hippocampal activation predicted opposing

1   effects on memory depending on whether events were expected or surprising. Moreover, this

2   effect of hippocampal activation on memory depended on concurrent basal forebrain activation,

3   consistent with the idea that acetylcholine regulates hippocampal processing. On the basis of our

4   univariate findings, we propose that during video playback, the hippocampus continually

5   generates predictions and sustains representations during an ongoing episode (63–65). If no

6   prediction error is detected, these representations should be sustained, and the hippocampus

7   should preserve and reinforce the memory (i.e., decreasing false memories). If a prediction error

8   is detected, then the hippocampus should abandon ongoing predictions and prepare to update a

9   memory (i.e., increasing false memories) (28, 29). Therefore, we hypothesized that prediction

10 errors would disrupt sustained representations in the hippocampus, and that disrupting

11 hippocampal representations would lead to memory updating. Furthermore, we predicted that

12 activation in the basal forebrain and/or VTA would link hippocampal representations to memory

13 outcomes, via neuromodulation of hippocampal processing (24–29, 37, 52, 53).

14 ***Prediction Errors Disrupted Sustained Representations in the Hippocampus***

15 Past studies in rodents and humans have used *autocorrelation* measures, which quantify

16 similarity across neural patterns, to investigate hippocampal representations during naturalistic

17 tasks (57, 58). *Temporal autocorrelation* is an index of multivariate information that is preserved

18 over time; this measures moment-to-moment overlap of activation patterns (57, 64, 66).

19 Intracranial recordings in humans have shown that temporal autocorrelation in the hippocampus

20 ramps up over the course of familiar episodes (64). Ramping autocorrelation reflects sustained

21 neural representations, consistent with the hippocampus generating predictions and anticipating

22 upcoming stimuli (63, 64). To test whether hippocampal representations were sustained or

23 disrupted over time, we calculated temporal autocorrelation by correlating the activation of all

1   voxels within the hippocampus at timepoint *T* with the activation pattern at timepoint *T+1 sec*

2   (Methods, *Multivariate fMRI Analyses*).

3   First, we tested whether autocorrelation increased during video playback. We used linear

4   mixed effects regression to predict *hippocampal autocorrelation* (averaged over 5s bins) from

5   the variables *video segment, hemisphere, axis,* and all interaction terms. We found that

6   hippocampal autocorrelation increased linearly as videos progressed ($\beta = 0.024$, 95% CI [0.01,

7   0.04], $t = 3.37$, $p = .003$), suggesting that episodic representations were sustained and stabilized

8   during video playback (Figure 4A) (64). There were no significant effects of hemisphere or axis.

9   All parameter estimates are provided in Table S10A.

10   Next, we tested whether prediction error disrupted this ramping autocorrelation. To

11   analyze Post-Video *change in autocorrelation*, we calculated a difference score for each trial by

12   subtracting the average autocorrelation value from the 5s bin immediately *before* video offset

13   from the average autocorrelation value from the 5s video immediately *after* video offset. We

14   then used linear mixed effects regression to predict average *change in autocorrelation* from the

15   variables *reactivation type*, *hemisphere*, *axis*, and interaction terms. There was a significant main

16   effect of reactivation type ($\beta = 0.05$, 95% CI [0.02, 0.08], $t = 2.97$, $p = .006$), such that

17   autocorrelation increased after the offset of Full videos but not Interrupted videos (Figure 4C). In

18   other words, prediction errors disrupted the continuity of hippocampal representations. This

19   Post-Video divergence is visualized in Figure 4B. There were no significant interactions with

20   hemisphere or axis; all parameter estimates are provided in Table S10B.
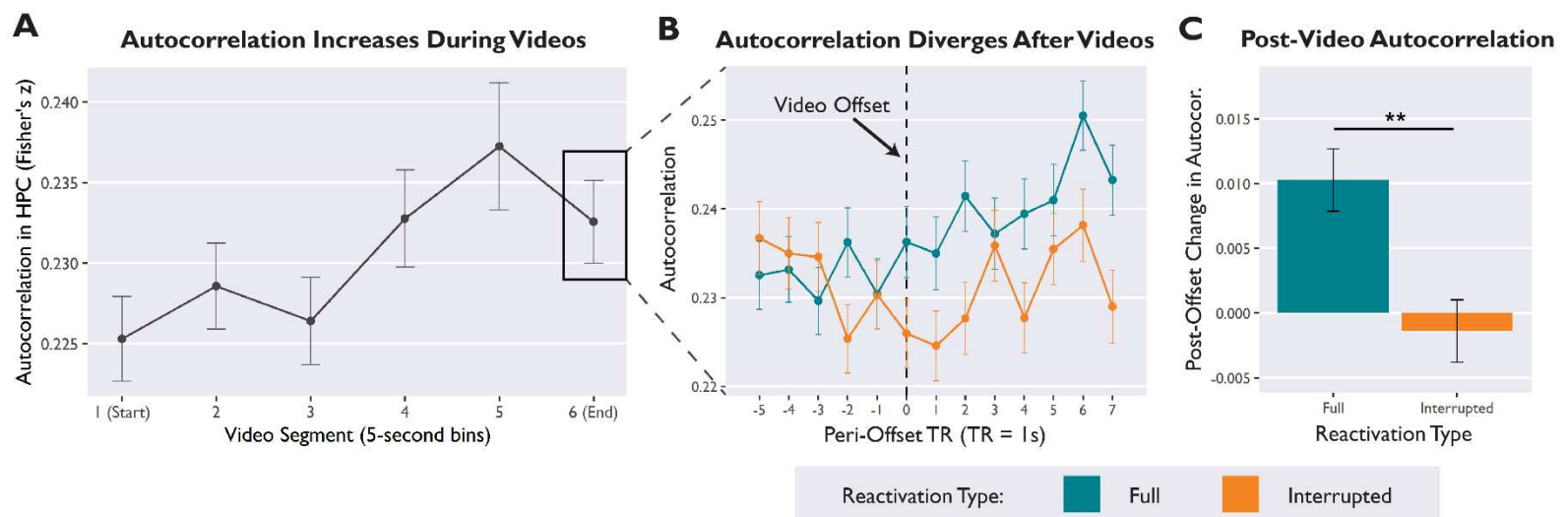
Figure 4. Hippocampal representations are sustained or disrupted over time, depending on whether or not episodes align with expectations. A) Temporal autocorrelation in the hippocampus gradually increased over the course of a video, suggesting that episodic representations were sustained over time. Autocorrelation values were averaged over 5-second bins of video playback. B) Autocorrelation trajectories for Full and Interrupted videos diverged during the Post-Video period. Plot visualizes second-by-second autocorrelation values in the hippocampus, time-locked to the moment of video offset (black dotted line). C) Average Post-Video change in autocorrelation (average autocorrelation scores for the 5-sec bin immediately *after* video offset, minus average autocorrelation for the bin immediately *before* offset). Hippocampal representations were sustained after Full videos, but disrupted after Interrupted videos. Error bars depict SEM.

1   ***Disruption of Hippocampal Autocorrelation Predicted False Memories***

2   Next, we tested whether disruption of hippocampal representations predicted memory

3   updating. Using linear mixed effects regression, we predicted subsequent *false memories* from

4   the variables *reactivation type*, Post-Video *change in autocorrelation*, *hemisphere*, *axis*, and all

5   relevant interactions. We also included a continuous covariate for univariate hippocampal

6   activation (thus controlling for any autocorrelation effects that may be a consequence of

7   univariate activation). There was a significant interaction between reactivation type and change

1   in autocorrelation predicting false memories ($\beta = 0.04$, 95% CI [0.01, 0.06], $t = 3.09$, $p = .002$)

2   (Figure 5A). After Interrupted videos, disrupting hippocampal representations led to memory

3   updating ($\beta = -0.06$, $z = -2.65$, $p = .008$). Conversely, after Full videos, hippocampal

4   autocorrelation was not related to memory ($\beta = 0.02$, $z = 0.88$, $p = .377$). There were no

5   significant interactions with hemisphere or axis. All other parameters are reported in Table

6   S11A.

7   ***Basal Forebrain Activation Links Hippocampal Autocorrelation to Memory***

8   What determines whether hippocampal representations are sustained or disrupted? To

9   investigate candidate neuromodulatory mechanisms, we extended the model described above by

10   adding parameters for Post-Video univariate basal forebrain activation and interaction terms. The

11   model included all relevant interaction terms, reported in full in Table S11B. Paralleling our

12   univariate findings, we found a significant three-way interaction among basal forebrain

13   activation, reactivation type, and hippocampal autocorrelation that predicted subsequent false

14   memories ($\beta = 0.03$, 95% CI [0.01, 0.05], $t = 2.74$, $p = .006$) (Figure 5B). In other words,

15   prediction errors disrupted hippocampal representations and led to memory updating, but only

16   when the basal forebrain was also strongly activated (at +1SD basal forebrain activation, Full: β

17    $= 0.08$, $z = 3.24$, $p = .001$; Interrupted: $\beta = -0.06$, $z = -2.20$, $p = .028$) (Figure 5B, right).

18   Next, we tested whether VTA activation was related to hippocampal autocorrelation and

19   memory. We modified the model described above (predicting subsequent false memories) by

20   replacing the basal forebrain activation variable (and interaction terms) with the VTA activation

21   variable. In contrast to our basal forebrain findings, the three-way interaction among VTA

22   activation, reactivation type, and hippocampal autocorrelation was not significant ($\beta = -0.003$, $t =$

23   -0.26, $p = .794$, 95% CI [-0.03, 0.02]). All parameter estimates are reported in Table S11C.

1   Overall, we found that our autocorrelation results paralleled our univariate findings: basal

2   forebrain activation, but not VTA activation, was crucial for connecting hippocampal

3   representations to memory outcomes.

4   ***Autocorrelation Control Analyses***

5   Lastly, to determine the anatomical specificity of our autocorrelation findings, we tested

6   two control regions: inferior lateral occipital cortex (LOC) and white matter. We predicted that

7   autocorrelation in LOC would be sensitive to all video offsets because of the change in visual

8   input, but *not* sensitive to prediction error. In contrast, physiological noise from white matter

9   should not be sensitive to either video offsets or prediction errors. Autocorrelation in LOC

10   significantly increased after videos ($t(23) = 9.47$, $p < .001$, Cohen's $d = 1.37$), but did not differ

11   by reactivation type ($t(23) = -0.05$, $p < .96$, $d = 0.01$). Autocorrelation in white matter did not

12   change post-offset ($t(23) = 0.99$, $p = .329$, $d = 0.14$) and did not differ by reactivation type ($t(23)$

13   $= 0.86$, $p = .40$, $d = 0.18$). In summary, these control analyses indicated that our autocorrelation

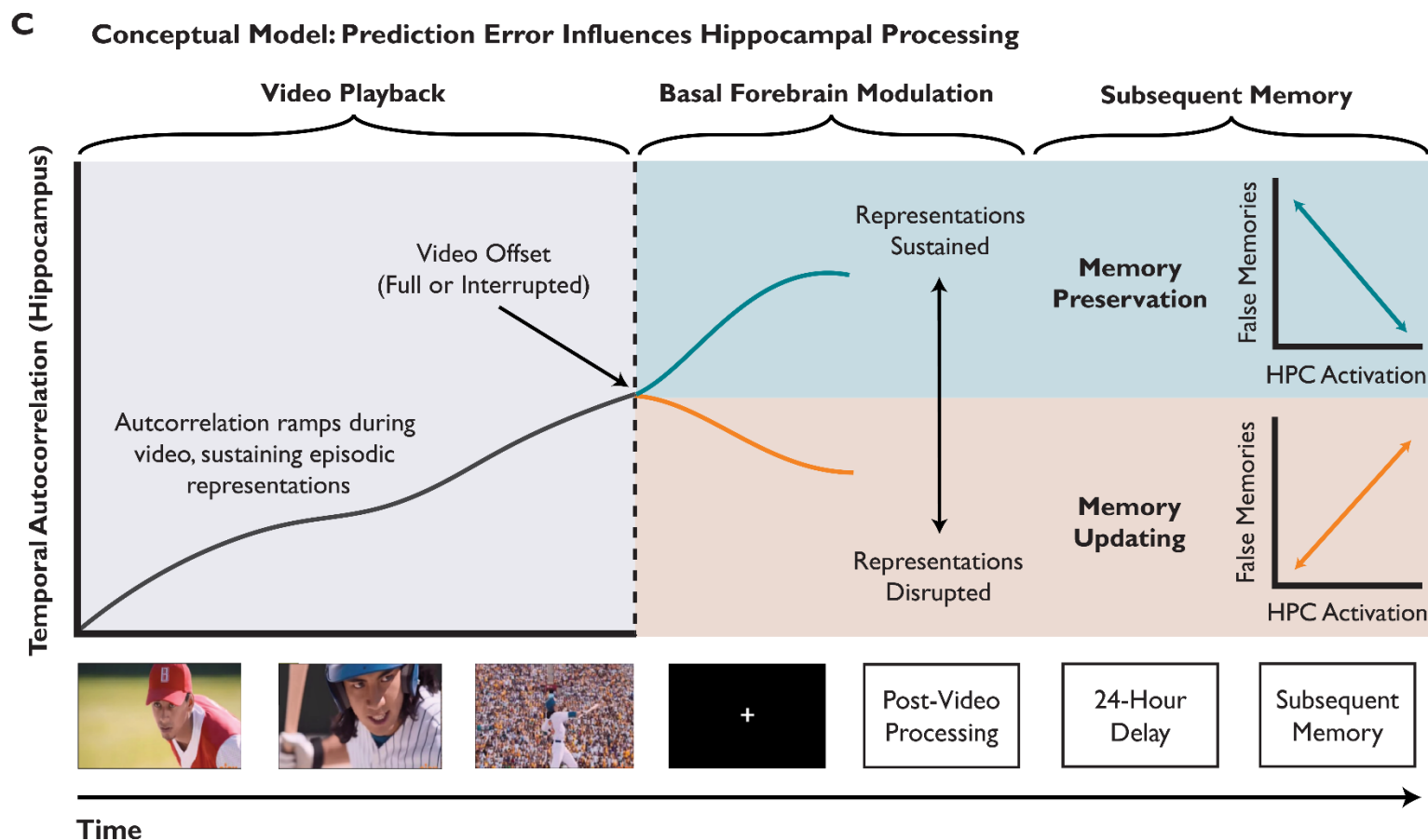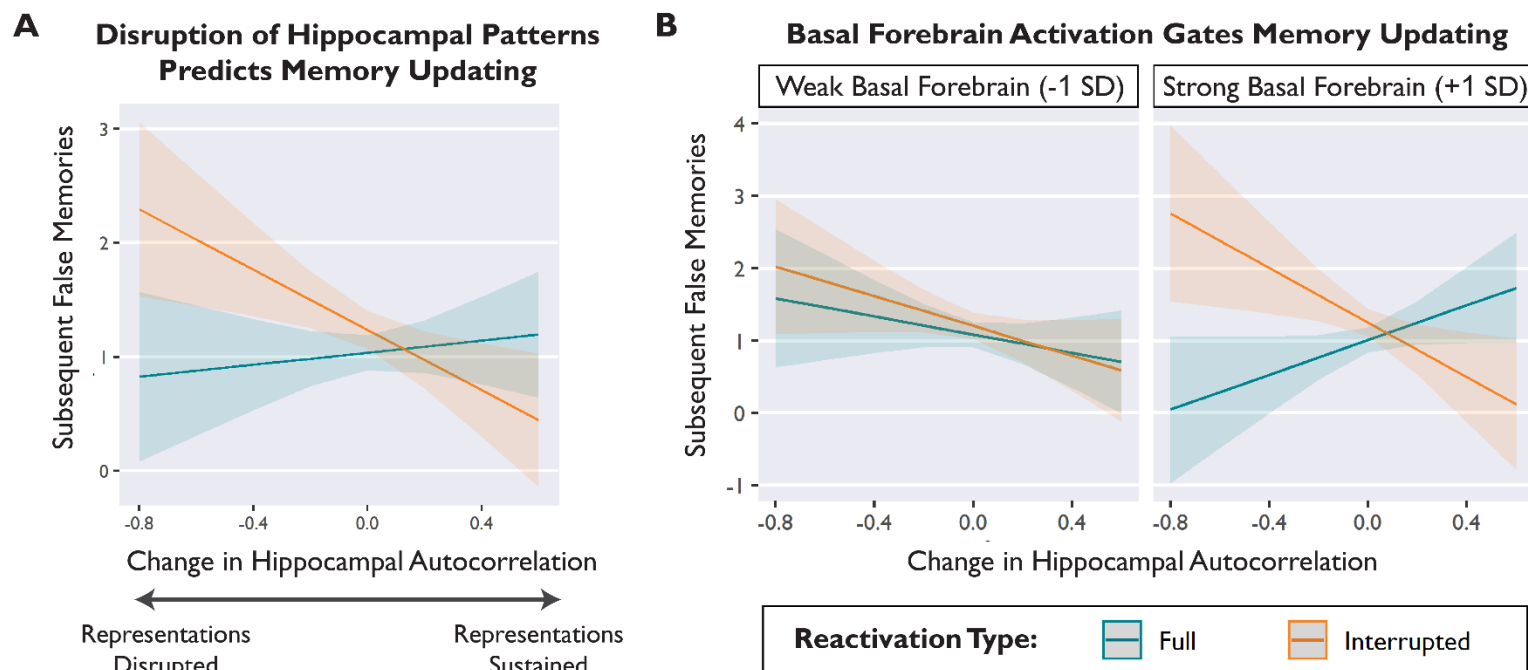14   findings were not a brain-wide phenomenon.

Figure 5. Prediction errors elicited by Interrupted videos disrupt sustained hippocampal representations, and this disruption predicts memory updating. A) Estimated values from a linear regression model predicting subsequent false memories from the interaction of reactivation type and change in autocorrelation. After Interrupted videos, decreases in autocorrelation were related to increased memory updating. B) The effect of prediction error on hippocampal autocorrelation and subsequent memory depended on concurrent basal forebrain activation. Basal forebrain activation was binned (weak vs. strong) for visualization, but statistical models used a continuous variable. Shaded bands depict 95% confidence intervals around the regression line. In panels A and B, model-predicted estimates are depicted instead of individual data points in order to show within-subject effects, while controlling for subject and stimulus variability.
C) Conceptual schematic depicting the effect of prediction error on hippocampal representations and subsequent memory. During a video, the hippocampus sustains episodic representations over time, consistent with generating ongoing predictions. After video offset, preservation or disruption of these representations indicates whether the hippocampus switches to an internal processing mode, whereby univariate activation preserves memories, or an external processing mode, whereby univariate activation updates memories. The link between prediction error and memory outcomes depended on co-activation of the hippocampus and basal forebrain during the post-video period.

Discussion

1   Here, we show that prediction errors modulate the function of the hippocampus and allow

2   memories to be modified, consistent with an adaptive updating mechanism. In our fMRI

3   paradigm, we elicited mnemonic prediction errors by interrupting familiar narrative videos

4   immediately before the expected conclusions. Prediction errors reversed the relationship between

5   univariate hippocampal activation and subsequent memory: After expected video endings,

6   hippocampal activation was associated with memory preservation, but after prediction errors,

7   hippocampal activation was associated with memory updating. Tracking the stability of

8   hippocampal representations revealed that prediction errors disrupted activation patterns; this

9   pattern disruption predicted memory updating. Crucially, the association between hippocampal

10   activation (both univariate and multivariate) and memory outcomes depended on concurrent

11   basal forebrain activation during the Post-Video period. We conclude that prediction error,

1   coupled with basal forebrain modulation, prompts the hippocampus to abandon ongoing

2   predictions and prepare to update a memory by incorporating other information present in the

3   environment (Figure 5).

Prediction Errors Disrupt Hippocampal Representations and Update Memories

5   Past studies of mnemonic prediction errors have reported an increase in univariate

6   hippocampal activation, but have not examined whether this neural signal affects memory (17–

7   19, 67). For the first time, we show that after prediction errors, hippocampal activation leads to

8   memory updating. Crucially, we demonstrate that univariate measures are *insufficient* for

9   understanding the effect of prediction error on the hippocampus, because the same amount of

10   hippocampal activation can exert opposing effects on memory. Prediction error reversed the

11   relationship between hippocampal activation and subsequent memory, consistent with a shift in

12   hippocampal processing (Figure 3). After expected endings (Full videos), hippocampal activation

13   protected against false memories, consistent with the idea that the hippocampus reinforces

14   memory after the conclusion of an episode (59, 60). In contrast, after surprising endings

15   (Interrupted videos), hippocampal activation predicted *more* false memories, consistent with the

16   idea that prediction errors can destabilize memories and enable updating (1–3).

17   To test the idea that prediction errors influence hippocampal processing, we tracked

18   hippocampal activation patterns to examine how episodic representations were sustained or

19   disrupted over time. We used *temporal autocorrelation* (the moment-to-moment overlap of

20   activation patterns) as a measure of continuity in hippocampal representations (57, 64, 66). As

21   narratives progressed, autocorrelation increased, reflecting stability and continuity; this increase

22   in autocorrelation suggested that the hippocampus generated predictions (63, 64) and sustained

23   episodic representations over time (59, 68) (Figure 4A). Crucially, prediction errors disrupted the

1   stability of hippocampal representations (Figure 4B, 4C), and this disruption predicted the degree

2   of memory updating (Figure 5A). Overall, we propose that disruption of hippocampal

3   representations indicates a shift in processing: Prediction error prompts the hippocampus to

4   abandon ongoing predictions and prepare to update a memory (Figure 5C).

5   Our univariate findings also diverge from past studies of hippocampal prediction error

6   responses, which have shown that hippocampal activation increases after prediction errors (17–

7   19, 67). In contrast, our single-trial analyses showed no significant main effect of reactivation

8   type on hippocampal activation; the effect of prediction error was revealed only when examining

9   the link between hippocampal activation and subsequent memory. Moreover, our whole-brain

10   mass univariate analyses revealed a significant cluster in left hippocampus that was *less* activated

11   after Interrupted videos than Full videos (Supplementary Information, *Whole-Brain Analysis*).

12   Crucially, our task elicits surprise by omitting expected endings, comparable to a *negative*

13   prediction error. Previous studies have elicited surprise by replacing expected stimuli with novel

14   stimuli, thus adding new information, comparable to a *positive* prediction error. Therefore, this

15   reversal of activation is consistent with prior studies of reward (13) and information prediction

16   errors (4, 69), which have shown *increases* in neural activation after positive prediction errors

17   and *decreases* after negative prediction errors. Overall, our results suggest that hippocampal

18   responses may depend on the stimuli and type of surprise, and demonstrate that prediction errors

19   change the effect of hippocampal activation on memory.

20   **Basal Forebrain Activation Relates to Hippocampal Processing**

21   Past studies have suggested that either cholinergic (27–29, 41–43, 52) or dopaminergic

22   (16, 24, 53) modulation could regulate hippocampal processing after prediction errors. However,

23   mixed evidence supporting both hypotheses has left the question unresolved (25, 26, 31, 37, 40).

1   Here, we investigated whether activation of the basal forebrain or the VTA could explain the

2   relationship between hippocampal activation after prediction error and subsequent memory. We

3   found that the effect of prediction error on memory depended on co-activation of the

4   hippocampus and basal forebrain, suggesting that connectivity between these regions is

5   important for shifting hippocampal processing. Hippocampal activation was associated with

6   memory updating after prediction errors, but only when the basal forebrain was also activated.

7   Likewise, disrupting hippocampal representations led to memory updating after prediction errors,

8   but only when the basal forebrain was also activated. Although fMRI cannot provide direct

9   evidence of neuromodulation, our results are consistent with the idea that cholinergic modulation

10   from the basal forebrain (27, 41–43, 45) gates memory updating and influences hippocampal

11   processing.

12   Our findings are also relevant to the functional relationship between the VTA and

13   hippocampus. Although we found a robust positive correlation between VTA and hippocampal

14   activation during the Post-Video period, VTA activation was unrelated to prediction error and

15   did not link hippocampal activation with memory updating. These findings are consistent with

16   our prior proposal that connectivity between the VTA and hippocampus reflects modulation of

17   hippocampal learning states by sustained VTA activity (16, 69) rather than phasic VTA

18   responses (25, 70–72). However, our paradigm was optimized for detecting memory updating

19   instead of midbrain prediction error responses. It is also possible that prediction error signals

20   could be transmitted to the hippocampus from the locus coeruleus (74, 75). Analyses of locus

21   coeruleus in the current data did not reveal any relationships. Future research could disambiguate

22   the roles of the basal forebrain, VTA, and locus coeruleus by examining both event-related and

23   sustained connectivity with the hippocampus and their consequences for memory.

1   **Prediction Error Both Strengthens and Updates Memories**

2   Our behavioral results demonstrated a novel dissociation: prediction error both

3   strengthened and updated memories, but over distinct timecourses and likely via different

4   mechanisms (Figure 2). Prediction error increased the number of correct details recalled, both

5   immediately and after a one-day delay. This finding is consistent with recent evidence that

6   mnemonic prediction errors can promote detailed memories immediately (76), possibly by way

7   of enhanced attention and pattern separation. In contrast, we found that prediction error increased

8   false memories only after a delay, supporting that idea that reconsolidation specifically enables

9   memory updating or distortion (1, 54, 77). Overall, our finding that prediction error increased

10   both correct details and false memories supports the idea that surprise drives adaptive updating.

11   In the present study, we used false memories as an index of memory updating. In the real

12   world, however, memory updating can be adaptive: new information is not "false" *per se*, but a

13   relevant addendum to prior knowledge. In our paradigm, interference from other stimulus videos

14   likely produced false memories because information was integrated across videos. Previously,

15   we found that prediction errors selectively updated memories with semantically-related

16   information from new videos that were specifically chosen to interfere with reactivated

17   memories (39). Here, we showed that videos that shared greater semantic similarity with the rest

18   of the stimulus set produced more false memories (Table S4). This finding suggests that

19   prediction error increases source confusion or integration of information across related

20   memories. Reconsolidation studies have shown that memory updating occurs when interference

21   is introduced after a memory trace is reactivated (1, 54, 78). In our paradigm, interference could

22   arise from semantically-related details that result from the subject recalling related memories, or

23   visual input from subsequent videos during the task. This finding accords with reconsolidation

1   research (1, 54, 78) and computational models of event segmentation (79, 80), which have both

2   shown that interference among related episodes can produce false memories and/or source

3   confusion after prediction error. However, memory updating is beneficial in other situations that

4   require integrating old and new knowledge, or correcting erroneous information.

5   **Limitations and Future Directions**

6   Our experimental design was inspired by reconsolidation theory, but evidence for cellular

7   reconsolidation processes in humans is lacking. Numerous behavioral studies have used

8   reconsolidation-like paradigms to demonstrate memory malleability (4, 39, 54, 78), but it

9   remains unknown whether the synaptic mechanisms of reconsolidation are consistent across

10   animals and humans (1). We found that the behavioral effect of prediction error on memory

11   updating required a delay, consistent with reconsolidation processes that rely on protein

12   synthesis. We are unable to directly test whether the relationship between neural activation and

13   memory updating required a delay, because the fMRI participants always completed the delayed

14   memory test. Overall, our findings are broadly relevant to research on prediction error and

15   memory even though the synaptic mechanisms remain unknown; reconsolidation theory offers

16   one plausible framework for our results. Another limitation is that the present data lacked the

17   spatial resolution required to segment hippocampal subfields, because we prioritized temporal

18   resolution over spatial resolution in order to track rapid changes in hippocampal representations

19   during and after naturalistic episodes.

20   In the present study, we elicited surprise by interrupting videos before their expected

21   narrative endings. However, we were not able to directly measure subjective surprise, because

22   asking participants to rate surprise after each video would have disrupted other cognitive

23   processes and revealed the goal of the manipulation. One possibility is that the strength of

1   surprise diminished over the course of the experiment; the first interrupted video was likely the

2   most surprising. To address this limitation, we designed our experiment to maintain surprise as

3   much as possible: Full-length videos were presented twice during encoding to set strong

4   expectations; Interrupted videos violated strongly-expected action-outcome contingencies (e.g., a

5   baseball batter halted mid-swing); we pseudorandomized our trials to ensure that participants

6   could not anticipate whether the next video would be Full or Interrupted; and participants did not

7   know *when* each video would be interrupted. Consistent with these design considerations, we

8   found that the effect of prediction error on false memories did not interact with trial number,

9   suggesting that surprise did not meaningfully diminish over the course of the experiment (Table

10   S2).

11   Animal and human reconsolidation studies have previously used *incomplete reminders*

12   (e.g., a conditioned stimulus without the expected outcome) to elicit prediction error (1, 3, 78).

13   Here, we mimicked this approach by interrupting narrative videos and omitting the expected

14   endings. Incomplete reminders may be particularly effective because they elicit memory

15   reactivation, and memory reactivation supports plasticity (81–83). After Interrupted videos,

16   participants may actively retrieve the missing endings; this memory reactivation could contribute

17   to memory malleability. As discussed in our prior review (1), both prediction error and memory

18   reactivation may be important for memory updating. These mechanisms could interact. Memory

19   reactivation may be a prerequisite for generating a prediction, and experiencing a prediction error

20   may prompt further memory reactivation. Future studies could investigate the role of memory

21   reactivation, such as by asking participants to report their ability to recall the missing endings, or

22   by testing encoding-retrieval pattern similarity.

1   Lastly, a limitation of the present study is that we were unable to determine the temporal

2   distribution or the source of each individual false memory. Most details that participants recalled

3   (both correct and incorrect) pertained to the entire video, such as perceptual details about the

4   characters and setting. Many of these perceptual details were also shared across multiple videos

5   (e.g., several characters with red shirts), making it impossible to determine the source of each

6   specific detail. Interestingly, we found that the omitted endings on Interrupted trials were rarely

7   associated with false memories. Recall of the narrative content of the video endings was very

8   accurate: in the Reconsolidation group, only 2% of all false memories specifically pertained to

9   the missing ending. Because the salient action-outcome contingencies define the videos, even

10   when interrupted, these central details were not likely to be affected. Instead, we found that

11   prediction error was more likely to induce a holistic distortion of perceptual content from

12   throughout each video (e.g., details about the setting or a character).

13   **Conclusion**

14   The brain continually generates predictions based on past experiences. When

15   expectations do not align with reality, memories should be updated with relevant new

16   information. We propose that prediction errors prompt the hippocampus to abandon ongoing

17   predictions and update memories by incorporating relevant details from subsequent experiences.

18   In this way, surprise modulates hippocampal processing and determines the fate of episodic

19   memories. This theoretical framework of memory updating bears implications for eyewitness

20   testimony, education, and conditions like Post-Traumatic Stress Disorder. Beyond memory

21   research, our results offer new insights for theories on the whole-brain predictive processes that

22   govern attention, perception, action, and decision-making.

Author Contributions: AHS and MDB developed the study design. AHS programmed the study, collected data, conducted analyses, and drafted the manuscript. GMM contributed substantially to data collection and IKB contributed to autocorrelation analyses. MDB and RAA contributed to the analysis approach and interpretation of results. All authors contributed to revising the manuscript and approved the final version.

Declaration of Interests: The authors have no competing interests to declare.

1 <div align="center">**Methods**</div>

2 **Data, Code, and Materials**

3   Brief descriptions of the stimulus videos are provided in Table S12. The full set of

4 stimulus videos, along with derivative data and code necessary to reproduce results, are provided

5 online in the project repository hosted by the Open Science Framework (https://osf.io/xb7sq/).

6 **Participants**

7   We recruited 55 paid participants from the University of Toronto community

8 (Reconsolidation group: $70, Immediate control group: $40). Seven participants were excluded

9 (see *Exclusions*), yielding a final sample of 48 participants. The sample size was determined *a*

10 *priori* to satisfy the following conditions: (1) reproduce the sample size and achieve at least 90%

11 power to detect the interaction effect found in prior work ($\eta_p^2 = 0.17$) (39), and (2) evenly

12 allocate participants to 6 pseudorandomized trial order lists. Participants were healthy young

13 adults (age: $M = 22.42$, $SD = 2.41$, range [18, 30]; gender: 75% female, 25% male). Inclusion

14 criteria were as follows: between the ages of 18-30, normal or corrected-to-normal vision and

15 hearing, no history of neurological or psychiatric disorders, and fluency in English. fMRI

16 participants were all right-handed. All participants provided informed consent, and the study was

17 approved by the University of Toronto Institutional Review Board, Protocol #00035787.

18   In consideration of the effects of sleep on consolidation, we also asked participants to

19 report approximate hours of sleep over the course of the study. Participants slept an average of

20 7.28 hours ($SD = 1.31$) between the Day 1 and Day 2 sessions, and Reconsolidation group

21 participants slept an average of 7.02 hours ($SD = 1$) between the Day 2 and Day 3 sessions.

22

23

Stimuli

Stimulus videos were sourced from movies, TV, and YouTube clips. We chose 70 videos that featured distinct narrative events (duration $M = 30$ sec, $SD = 7$ sec). Semantic similarity varied across videos (e.g., several videos featured sporting events), but there were no overlapping scenes or characters. During pilot testing, we ensured that the videos would be infrequently recognized by our participants. The 70 videos used in the experiment are described in Table S12 and publicly available on the Open Science Framework (https://osf.io/xb7sq/). The Interrupted version of each video ended abruptly at the narrative climax, omitting the salient ending and violating expectations (duration $M = 25$ sec, $SD = 4$ sec). Importantly, the duration of the videos varied considerably, yielding overlapping distributions of Full and Interrupted videos (Figure S3). This variability ensured that our neural results could not be attributed to systematic timing differences between the Full and Interrupted videos (e.g., neural adaptation after given durations of visual stimulation).

Procedure

During the **Encoding session**, participants viewed all 70 videos in full-length form (randomized order). Each video was presented twice in a row to ensure that participants had strong expectations about the narrative outcomes for each video, a prerequisite for eliciting prediction error later.

During the **Reactivation session**, participants viewed each video again a single time (35 Full videos, 35 Interrupted videos). Videos were played in a pseudorandom order (six trial order lists, counterbalanced across participants) such that there were never more than two consecutive Interrupted videos. Participants could not reliably anticipate whether each video would be Interrupted, or where the interruption might occur. thus preventing participants from anticipating

1   which videos would be interrupted. We also performed eye-tracking during the Encoding and

2   Reactivation sessions for participants in both the Reconsolidation and Immediate groups

3   (EyeLink v.1000+, SR-Research). Eye-tracking was used to monitor alertness during the task,

4   but these data are not discussed further.

5   Lastly, the **Test session** involved a structured interview-style recall test about details

6   from each of the videos. Participants were cued with the name of each video and prompted to

7   recall the narrative. The experimenter then probed the participant for more information with a

8   pre-determined list of questions (e.g., "Can you describe the setting or context of the video?",

9   "Can you describe what the character looked like? Do you remember gender, age range, hair

10   color, or clothing?"). Participants were instructed to answer based on their memory of the Full-

11   length videos that had been originally presented during encoding. Because we were interested in

12   false memories as a measure of memory modification, we instructed participants not to guess and

13   permitted them to skip details they could not recall.

14   Overall, the experiment took place over three days for participants in the Reconsolidation

15   group (24-hour delays between Encoding, Reactivation, and Test), or over two days for

16   participants in the Immediate control group (24-hour delay between Encoding and Reactivation,

17   no delay between Reactivation and Test). Only the Reconsolidation group underwent

18   neuroimaging.

19   Consistent with past reconsolidation studies (84–86), we maintained consistent contextual

20   factors between Encoding, Reactivation, and Test sessions. Reconsolidation group participants

21   completed the encoding session in a mock scanner (shell of a retired 1.5T Siemens Avanto

22   scanner), while recorded MRI sounds were played in the background. The mock scanner room

23   was adjacent to the real scanner room, and was very similar in size and appearance.

1   Reconsolidation group participants completed the Reactivation session in the real fMRI scanner

2   and the Test session at a desk in the mock scanner room. Participants in the Immediate control

3   group completed all three sessions in the same behavioral testing room. In both groups,

4   participants completed all three sessions with the same experimenter.

5   **Statistical Analysis**

6   For both behavioral and neural data, we conducted trial-wise analyses with linear mixed-

7   effects regression models. Detailed information about the random effects specification is

8   provided in the Supplementary Information (*Supplementary Methods, Linear Mixed-Effects*

9   *Regression*).

10   ***Univariate fMRI Analyses***

11   Whole-brain mass univariate results are reported in the Supplementary Information

12   (*Whole-Brain Analysis*, Figure S2, Table S5). Details about preprocessing steps and region-of-

13   interest masks are provided in the Supplementary Information (*Supplementary Methods, fMRI*

14   *Preprocessing, Region of Interest Masks*). The primary findings reported in the main text reflect

15   a single-trial modelling approach that estimated hippocampal responses to each video during the

16   task. In order to isolate responses on each trial, we employed the Least Squares-Single approach

17   and constructed a separate GLM for each trial (87, 88). We modelled each trial as a 2s impulse in

18   the post-video period, convolved with the canonical double-gamma hemodynamic response

19   function and phase-shifted 2s after video offset. This 2s shift targets the peak hippocampal

20   response previously identified in studies of post-video processing (59, 60). Within each GLM,

21   the target trial (2s event) was isolated as one regressor, and all other events were modelled with a

22   separate regressor for each type of event (e.g., video playback, video name cues, other fixation

23   periods). This approach yielded whole-brain parameter estimates for each trial, in native space.

1   For each trial, we masked the processed data and averaged across voxels within each ROI to

2   yield an average activation value.

3   ***Multivariate fMRI Analyses***

4   Multivariate temporal autocorrelation analyses (57, 64) were conducted on preprocessed

5   data (prior to single-trial GLM analysis). We extracted the whole-run timeseries from every

6   voxel within each ROI using the *fslmeants* utility. For control analyses (white matter and LOC

7   ROIs), autocorrelation was calculated on 200 contiguous voxels, approximately matching the

8   size of the hippocampal ROIs. Temporal autocorrelation was defined as the Pearson product-

9   moment correlation between all voxel activation values at timepoint T and timepoint T+1s. This

10   method yielded an autocorrelation value for every second of each functional run, excluding the

11   final TR. Autocorrelation values were standardized (Fisher's *z*) prior to statistical analysis.

12   Next, we aligned multivariate timeseries data with event onset and duration markers.

13   Comparable to past research, we phase-shifted the timeseries by 4 seconds in order to account for

14   HRF lag (89). This manual shifting is necessary because event onset regressors have not yet been

15   convolved with the HRF (unlike in standard GLM analyses used for our univariate analyses).

16   Note that our univariate analyses involved an additional 2s shift in addition to the standard HRF

17   shift; this allowed us to target the peak post-video hippocampal response on each trial, but was

18   not necessary for the autocorrelation analyses that yield TR-by-TR values.

19   After alignment, we calculated average autocorrelation values that were time-locked to

20   events. For statistical analyses, autocorrelation values were averaged across 5-second bins during

21   and after each video. To analyze signal history over the course of video playback, we related the

22   video segment number (5s bins) to average autocorrelation values. For each video, we included

23   the first five seconds (timepoints 0-4), the next four middle segments (timepoints 5-9, 10-14, 15-

1   19, and 20-24), and the last five seconds (variable depending on the length of the video). This

2   binning scheme spans the average video length of 30 seconds; additional middle segments from

3   videos that were longer than 30 seconds were omitted. Lastly, to analyze Post-Video changes in

4   autocorrelation, we calculated trial-by-trial difference scores by subtracting the average

5   autocorrelation value for the 5-second bin immediately *before* video offset from the average

6   autocorrelation value for the 5-second bin immediately *after* video offset. Autocorrelation values

7   and difference scores for each trial were then submitted to linear mixed effects regression.

References

1. A. H. Sinclair, M. D. Barense, Prediction Error and Memory Reactivation: How Incomplete Reminders Drive Reconsolidation. *Trends Neurosci.* **42**, 727–739 (2019).

2. R. N. Henson, P. Gagnepain, Predictive, Interactive Multiple Memory Systems. *Hippocampus* **20**, 1315–1326 (2010).

3. M. T. J. Exton-Mcguinness, J. L. C. Lee, A. C. Reichelt, Updating memories—The role of prediction errors in memory reconsolidation. *Behav. Brain Res.* **278**, 375–384 (2015).

4. A. Pine, N. Sadeh, A. Ben-Yakov, Y. Dudai, A. Mendelsohn, Knowledge acquisition is governed by striatal prediction errors. *Nat. Commun.* **9**, 1–14 (2018).

5. G. Kim, K. A. Norman, N. B. Turk-Browne, Neural differentiation of incorrectly predicted memories. *J. Neurosci.* **37**, 2022–2031 (2017).

6. J. A. Quent, R. N. Henson, A. Greve, A predictive account of how novelty influences declarative memory. *Neurobiol. Learn. Mem.* **179**, 107382 (2021).

7. D. L. Schacter, D. R. Addis, The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philos. Trans. R. Soc. B Biol. Sci.* **362**, 773–786 (2007).

8. L. F. Barrett, W. K. Simmons, Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* **16**, 419–429 (2015).

9. K. Friston, The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138 (2010).

10. G. B. Keller, T. D. Mrsic-Flogel, Predictive Processing: A Canonical Cortical Computation. *Neuron* **100**, 424–435 (2018).

11. M. W. Spratling, A review of predictive coding algorithms. *Brain Cogn.* **112**, 92–97 (2017).

12. R. S. Sutton, A. G. Barto, *Reinforcement learning: An introduction*, 2nd Ed. (MIT Press, 1998).

13. M. Watabe-Uchida, N. Eshel, N. Uchida, Neural circuitry of reward prediction error. *Annu. Rev. Neurosci.* **40**, 373–394 (2017).

14. N. E. Wheeler, *et al.*, Ideology and predictive processing: coordination, bias, and polarization in socially constrained error minimization. *Curr. Opin. Behav. Sci.* **34**, 192–198 (2020).

15. N. Hindy, E. Avery, N. Turk-Browne, Hippocampal-neocortical interactions sharpen over time for predictive actions. *Nat. Commun.* **10** (2019).

16. D. Shohamy, R. A. Adcock, Dopamine and adaptive memory. *Trends Cogn. Sci.* **14**, 464–472 (2010).

17. J. Chen, P. A. Cook, A. D. Wagner, Prediction strength modulates responses in human area CA1 to sequence violations. *J. Neurophysiol.* **114**, 1227–1238 (2015).

18. K. Duncan, N. Ketz, S. J. Inati, L. Davachi, Evidence for area CA1 as a match/mismatch detector: A high-resolution fMRI study of the human hippocampus. *Hippocampus* **22**, 389–398 (2012).

19. D. Kumaran, E. A. Maguire, An unexpected sequence of events: mismatch detection in the human hippocampus. *PLoS Biol.* **4**, e424 (2006).

20. O. Bein, K. Duncan, L. Davachi, Mnemonic prediction errors bias hippocampal states. *Nat. Commun.* **11**, 3451 (2020).

21. J. Chen, R. K. Olsen, A. R. Preston, G. H. Glover, A. D. Wagner, Associative retrieval processes in the human medial temporal lobe: Hippocampal retrieval success and CA1 mismatch detection. *Learn. Mem.* **18**, 523–528 (2011).

22. K. Duncan, C. Curtis, L. Davachi, Distinct memory signatures in the hippocampus: intentional states distinguish match and mismatch enhancement signals. *J. Neurosci.* **29**, 131–139 (2009).

23. K. C. Dickerson, J. Li, M. R. Delgado, Parallel contributions of distinct human memory systems during probabilistic learning. *NeuroImage* **55**, 266–276 (2011).

24. J. E. Lisman, A. A. Grace, The Hippocampal-VTA Loop: Controlling the Entry of Information into Long-Term Memory. *Neuron* **46**, 703–713 (2005).

25. J. E. Lisman, N. A. Otmakhova, Storage, recall, and novelty detection of sequences by the hippocampus: elaborating on the SOCRATIC model to account for normal and aberrant effects of dopamine. *Hippocampus* **11**, 551–568 (2001).

26. J. Schomaker, M. Meeter, Short- and long-lasting consequences of novelty, deviance and surprise on brain and cognition. *Neurosci. Biobehav. Rev.* **55**, 268–279 (2015).

27. M. Hasselmo, B. Wyble, G. Wallenstein, Encoding and Retrieval of Episodic Memories: Role of Cholinergic and GABAergic Modulation in the Hippocampus. *Hippocampus* **6** (1996).

28. C. J. Honey, E. L. Newman, A. C. Schapiro, Switching between internal and external modes: A multiscale learning principle. *Netw. Neurosci.* **1**, 339–356 (2017).

29. M. Meeter, J. M. J. Murre, L. M. Talamini, Mode shifting between storage and recall based on novelty detection in oscillating hippocampal circuits. *Hippocampus* **14**, 722–741 (2004).

30. B. E. Sherman, N. B. Turk-Browne, Statistical prediction of the future impairs episodic encoding of the present. *Proc. Natl. Acad. Sci.* **8** (2020).

31. M. F. Carr, L. M. Frank, A single microcircuit with multiple functions: state dependent information processing in the hippocampus. *Curr. Opin. Neurobiol.* **22**, 704–708 (2012).

32. M. M. Chun, J. D. Golomb, N. B. Turk-Browne, A taxonomy of external and internal attention. *Annu. Rev. Psychol.* **62**, 73–101 (2011).

33. J. Debiec, J. E. LeDoux, K. Nader, Cellular and systems reconsolidation in the hippocampus. *Neuron* **36**, 527–538 (2002).

34. J. L. C. Lee, Reconsolidation: maintaining memory relevance. *Trends Neurosci.* **32**, 413–420 (2009).

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

36. K. Nader, E. O. Einarsson, Memory reconsolidation: an update. *Ann N Acad Sci* **1191**, 27–41 (2010).

37. C. E. Wideman, K. H. Jardine, B. D. Winters, Involvement of classical neurotransmitter systems in memory reconsolidation: Focus on destabilization. *Neurobiol. Learn. Mem.* **156**, 68–79 (2018).

38. C. Forcato, P. Argibay, M. Pedreira, H. Maldonado, Human reconsolidation does not always occur when a memory is retrieved: The relevance of the reminder structure. *Neurobiol. Learn. Mem.* **91**, 50–57 (2009).

39. A. H. Sinclair, M. D. Barense, Surprise and destabilize: Prediction error influences episodic memory reconsolidation. *Learn. Mem.* **25**, 369–381 (2018).

40. H. Tarder-Stoll, M. Jayakumar, H. R. Dimsdale-Zucker, E. Günseli, M. Aly, Dynamic internal states shape memory retrieval. *Neuropsychologia* **138**, 107328 (2020).

41. E. L. Newman, S. N. Gillet, J. R. Climer, M. E. Hasselmo, Cholinergic blockade reduces theta-gamma phase amplitude coupling and speed modulation of theta frequency consistent with behavioral effects on encoding. *J. Neurosci.* **33**, 19635–19646 (2013).

42. L. M. Giocomo, M. E. Hasselmo, Neuromodulation by glutamate and acetylcholine can change circuit dynamics by regulating the relative influence of afferent input and excitatory feedback. *Mol. Neurobiol.* **36**, 184–200 (2007).

43. M. E. Hasselmo, The role of acetylcholine in learning and memory. *Curr. Opin. Neurobiol.* **16**, 710–715 (2006).

44. C. Kemere, M. F. Carr, M. P. Karlsson, L. M. Frank, Rapid and Continuous Modulation of Hippocampal Network State during Exploration of New Places. *PLOS ONE* **8**, e73114 (2013).

45. A. L. Decker, K. Duncan, Acetylcholine and the complex interdependence of memory and attention. *Curr. Opin. Behav. Sci.* **32**, 21–28 (2020).

46. N. Bunzeck, E. Düzel, Absolute Coding of Stimulus Novelty in the Human Substantia Nigra/VTA. *Neuron* **51**, 369–379 (2006).

47. B. C. Wittmann, N. Bunzeck, R. J. Dolan, E. Düzel, Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *NeuroImage* **38**, 194–202 (2007).

48. V. P. Murty, R. A. Adcock, Enriched encoding: reward motivation organizes cortical networks for hippocampal detection of unexpected events. *Cereb. Cortex N. Y. N 1991* **24**, 2160–8 (2014).

49. D. Shohamy, A. D. Wagner, Integrating Memories in the Human Brain: Hippocampal-Midbrain Encoding of Overlapping Events. *Neuron* **60**, 378–389 (2008).

50. R. A. Adcock, A. Thangavel, S. Whitfield-Gabrieli, B. Knutson, J. D. E. Gabrieli, Reward-Motivated Learning: Mesolimbic Activation Precedes Memory Formation. *Neuron* **50**, 507–517 (2006).

51. A. Tompary, K. Duncan, L. Davachi, Consolidation of associative and item memory is related to post-encoding functional connectivity between the ventral tegmental area and different medial temporal lobe subregions during an unrelated task. *J. Neurosci.* **35**, 7326–7331 (2015).

52. K. H. Jardine, *et al.*, Activation of cortical M1 muscarinic receptors and related intracellular signaling is necessary for reactivation-induced object memory updating. *Sci. Rep.* **10**, 9209 (2020).

53. J. I. Rossato, *et al.*, State-dependent effect of dopamine D1/D5 receptors inactivation on memory destabilization and reconsolidation. *Behav. Brain Res.* **285**, 194–199 (2015).

54. A. Hupbach, R. Gomez, L. Nadel, "Episodic memory reconsolidation: An update" in *Memory Reconsolidation*, C. M. Alberini, Ed. (Elsevier Academic Press, 2013), pp. 233–247.

55. R. A. Cooper, M. Ritchey, Progression from feature-specific brain activity to hippocampal binding during episodic encoding. *J. Neurosci.* **40**, 1701–1709 (2020).

56. Z. M. Reagh, A. I. Delarazan, A. Garber, C. Ranganath, Aging alters neural activity at event boundaries in the hippocampus and Posterior Medial network. *Nat. Commun.* **11**, 1–12 (2020).

57. I. K. Brunec, *et al.*, Multiple Scales of Representation along the Hippocampal Anteroposterior Axis in Humans. *Curr. Biol.* **28**, 2129-2135.e6 (2018).

58. K. B. Kjelstrup, *et al.*, Finite scale of spatial representation in the hippocampus. *Science* **321**, 140–143 (2008).

59. A. Ben-Yakov, N. Eshel, Y. Dudai, Hippocampal Immediate Poststimulus Activity in the Encoding of Consecutive Naturalistic Episodes. *J. Exp. Psychol. Gen.* **142**, 1255–1263 (2013).

60. A. Ben-Yakov, Y. Dudai, Constructing Realistic Engrams: Poststimulus Activity of Hippocampus and Dorsal Striatum Predicts Subsequent Episodic Memory. *J. Neurosci.* **31**, 9032–9042 (2011).

61. C. Baldassano, *et al.*, Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron* **95**, 709-721.e5 (2017).

62. M. Zubair, *et al.*, Divergent Whole Brain Projections from the Ventral Midbrain in Macaques. *Cereb. Cortex N. Y. N 1991* **31**, 2913–2931 (2021).

63. R. U. Haque, S. K. Inati, A. I. Levey, K. A. Zaghloul, Feedforward prediction error signals during episodic memory retrieval. *Nat. Commun.* **11**, 6075 (2020).

64. R. Paz, *et al.*, A neural substrate in the human hippocampus for linking successive events. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 6046–6051 (2010).

65. A. Tambini, L. Davachi, Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 19591–6 (2013).

66. R. Raut, A. Snyder, M. Raichle, Hierarchical dynamics as a macroscopic organizing principle of the human brain. *PNAS* **117**, 20890–20897 (2020).

67. D. Kumaran, E. A. Maguire, Match-mismatch processes underlie human hippocampal responses to associative novelty. *J. Neurosci.* **27**, 8517–8524 (2007).

68. S. DuBrow, L. Davachi, Temporal binding within and across events. *Neurobiol. Learn. Mem.* **134**, 107–114 (2016).

69. E. S. Bromberg-Martin, O. Hikosaka, Lateral habenula neurons signal errors in the prediction of reward information. *Nat. Neurosci.* **14**, 1209–1218 (2011).

70. V. P. Murty, I. C. Ballard, R. A. Adcock, Hippocampus and Prefrontal Cortex Predict Distinct Timescales of Activation in the Human Ventral Tegmental Area. *Cereb. Cortex N. Y. NY* **27**, 1660–1669 (2017).

71. J. Lisman, A. A. Grace, E. Duzel, A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends Neurosci.* **34**, 536–547 (2011).

72. A. A. Grace, S. B. Floresco, Y. Goto, D. J. Lodge, Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends Neurosci.* **30**, 220–227 (2007).

73. C. G. McNamara, Á. Tejero-Cantero, S. Trouche, N. Campo-Urriza, D. Dupret, Dopaminergic neurons promote hippocampal reactivation and spatial memory persistence. *Nat. Neurosci.* **17**, 1658–1660 (2014).

74. C. G. McNamara, D. Dupret, Two sources of dopamine for the hippocampus. *Trends Neurosci.* **40**, 383–384 (2017).

75. K. A. Kempadoo, E. V. Mosharov, S. J. Choi, D. Sulzer, E. R. Kandel, Dopamine release from the locus coeruleus to the dorsal hippocampus promotes spatial learning and memory. *Proc. Natl. Acad. Sci.* **113**, 14835–14840 (2016).

76. O. Bein, N. A. Plotkin, L. Davachi, Mnemonic prediction errors promote detailed memories (2021) https:/doi.org/10.31234/osf.io/b5z8a (May 11, 2021).

77. K. Nader, Memory traces unbound. *Trends Neurosci.* **26**, 65–72 (2003).

78. C. Forcato, M. L. C. Rodríguez, M. E. Pedreira, H. Maldonado, Reconsolidation in humans opens up declarative memory to the entrance of new information. *Neurobiol. Learn. Mem.* **93**, 77–84 (2010).

79. N. T. Franklin, K. A. Norman, C. Ranganath, J. M. Zacks, S. J. Gershman, Structured Event Memory: A neuro-symbolic model of event cognition. *Psychol. Rev.* **127**, 327–361 (2020).

80. G. A. Radvansky, J. M. Zacks, Event boundaries in memory and cognition. *Curr. Opin. Behav. Sci.* **17**, 133–140 (2017).

81. V. J. H. Ritvo, N. B. Turk-Browne, K. A. Norman, Nonmonotonic Plasticity: How Memory Retrieval Drives Learning. *Trends Cogn. Sci.* **23**, 726–742 (2019).

82. D. Stawarczyk, C. Wahlheim, J. Etzel, A. Snyder, J. Zacks, Aging and the encoding of event changes: The role of neural activity pattern reinstatement. *bioRxiv* (2020) https:/doi.org/https://doi.org/10.1101/809806.

83. D. Stawarczyk, M. A. Bezdek, J. M. Zacks, Event Representations and Predictive Processing: The Role of the Midline Default Network Core. *Top. Cogn. Sci.* (2019) https:/doi.org/10.1111/tops.12450.

84. A. M. Capelo, P. B. Albuquerque, S. Cadavid, Exploring the role of context on the existing evidence for reconsolidation of episodic memory. *Memory* **27**, 280–294 (2019).

85. A. Hupbach, R. Gomez, L. Nadel, Episodic memory updating: The role of context familiarity. *Psychon. Bull. Rev.* **18**, 787–797 (2011).

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

87. J. A. Mumford, B. O. Turner, F. G. Ashby, R. A. Poldrack, Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage* **59**, 2636–2643 (2012).

88. J. A. Mumford, T. Davis, R. A. Poldrack, The impact of study design on pattern estimation for single-trial multivariate pattern analysis. *NeuroImage* (2014) https:/doi.org/10.1016/j.neuroimage.2014.09.026.

89. T. Sadeh, J. Chen, Y. Goshen-Gottstein, M. Moscovitch, Overlap between hippocampal pre-encoding and encoding patterns supports episodic memory. *Hippocampus* (2019) https:/doi.org/10.1002/hipo.23079.