



Kent Academic Repository

Dhammapeera, Phot, Brunskill, Chloe, Hellerstedt, Robin and Bergström, Zara M. (2024) *Counterfactual imagination impairs memory for true actions: EEG and behavioural evidence*. *Cognitive neuroscience*, 15 (1). pp. 12-23. ISSN 1758-8928.

Downloaded from

<https://kar.kent.ac.uk/105156/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.1080/17588928.2024.2315814>

This document version

Publisher pdf

DOI for this version

Licence for this version

CC BY-NC-ND (Attribution-NonCommercial-NoDerivatives)

Additional information

Versions of research works

Versions of Record

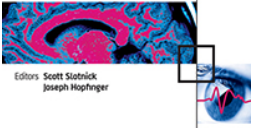
If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal**, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).



Counterfactual imagination impairs memory for true actions: EEG and behavioural evidence

Phot Dhammapeera, Chloe Brunskill, Robin Hellerstedt & Zara M. Bergström

To cite this article: Phot Dhammapeera, Chloe Brunskill, Robin Hellerstedt & Zara M. Bergström (2024) Counterfactual imagination impairs memory for true actions: EEG and behavioural evidence, *Cognitive Neuroscience*, 15:1, 12-23, DOI: [10.1080/17588928.2024.2315814](https://doi.org/10.1080/17588928.2024.2315814)

To link to this article: <https://doi.org/10.1080/17588928.2024.2315814>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 16 Feb 2024.



[Submit your article to this journal](#)



Article views: 154






[View related articles](#)



[View Crossmark data](#)

Counterfactual imagination impairs memory for true actions: EEG and behavioural evidence

Phot Dhammapeera ^{a,b,*}, Chloe Brunskill ^{a,*}, Robin Hellerstedt ^{a,c} and Zara M. Bergström ^a

^aSchool of Psychology, University of Kent, Canterbury, UK; ^bFaculty of Psychology, Chulalongkorn University, Bangkok, Thailand; ^cCentre for Biomedical Technology, Universidad Politécnica de Madrid, Madrid, Spain

ABSTRACT

Imagined events can be misremembered as experienced, leading to memory distortions. However, less is known regarding how imagining counterfactual versions of past events can impair existing memories. We addressed this issue, and used EEG to investigate the neurocognitive processes involved when retrieving memories of true events that are associated with a competing imagined event. Participants first performed simple actions with everyday objects (e.g., rolling dice). A week later, they were shown pictures of some of the objects and either imagined the same action they had originally performed, or imagined a counterfactual action (e.g., stacking the dice). Subsequent tests showed that memory for performed actions was reduced after counterfactual imagination when compared to both veridical imagination and a baseline condition that had not been imagined at all, providing novel evidence that counterfactual imagination impairs true memories beyond simple forgetting over time. ERPs and EEG oscillations showed evidence of separate processes associated with memory retrieval versus post-retrieval processes that were recruited to support recall of memories that were challenging to access. The findings show that counterfactual imagination can cause impairments to sensorimotor-rich event memories, and provide new evidence regarding the neurocognitive mechanisms that are recruited when people need to distinguish memories of imagined versus true events.

ARTICLE HISTORY

Received 19 July 2023
Revised 21 December 2023
Published online 16 February 2024

KEYWORDS


Episodic retrieval;
counterfactual thinking;
imagination inflation; source
monitoring; interference
resolution; EEG oscillations;
event-related potentials

Introduction

When remembering a previous event, we sometimes reflect on how it could have turned out differently. For example, after a job interview, we might in hindsight think about better answers to certain questions by mentally constructing a counterfactual imagined version of the event. Such episodic counterfactual thinking can help people make better future decisions (Roese & Epstude, 2017) and reduce negative emotions associated with unpleasant memories (De Brigard et al., 2019). However, counterfactual thinking can also induce memory distortions, causing people to mistake the imagined version of events for a real autobiographical memory (Gerlach et al., 2014; see also Dhammapeera et al., 2020). Here, we investigated whether counterfactual imagination after an event influences people's original memories of what truly happened. We also used EEG to better understand the neurocognitive mechanisms that enable people to distinguish an imagined version of an event from the corresponding true memory.

Imagined and perceived events can be difficult to differentiate between in our long-term memory, leading people to mistake imagined events for real autobiographical memories (Schacter & Addis, 2020). After imagining themselves performing simple actions, participants commonly erroneously think that they performed actions they had only previously imagined (Goff & Roediger, 1998; Thomas et al., 2003). Because interacting with the world typically results in very strong, accurate memories (Roberts et al., 2022), this research demonstrates the potency of imagination as a source of memory distortion. Vivid imagination is suggested to create memory traces that are rich enough in perceptual detail to resemble true memories, which causes source monitoring (or 'reality monitoring') errors during subsequent retrieval (Johnson et al., 1993). This literature however does not tell us whether counterfactually re-imagining what happened in a past event could *supplant* an existing true memory, nor the brain mechanisms that are engaged when such memory errors occur.

CONTACT Dr Zara Bergström  z.m.bergstrom@kent.ac.uk  School of Psychology, University of Kent, Canterbury CT2 7NP, UK
*joint first authors.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/17588928.2024.2315814>

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

fMRI research has shown that episodic counterfactual thinking activates highly overlapping brain networks to veridical episodic memory retrieval (De Brigard & Parikh, 2019), suggesting that similar neurocognitive mechanisms are engaged when reconstructing an accurate memory and when drawing on memory in order to construct an imagined event. Previous research has however not investigated the neural mechanisms that are engaged when counterfactual imagination disrupts a person's ability to subsequently recall the true memory, a situation which not only involves episodic reconstruction but also requires people to evaluate the truthfulness of memories and to select the correct over the imagined memory. It is likely that such situations involve prefrontally mediated processes that contribute to reality monitoring (Simons et al., 2017) and resolving retrieval competition (Anderson & Hulbert, 2021). The temporal resolution of fMRI however means it is not suitable for separating different stages of retrieval, such as the initial activation of a memory versus the subsequent post-retrieval operations that are used to evaluate the source and accuracy of that memory. Therefore, in the current study, we used EEG to investigate the neurocognitive processes that enable people to distinguish between counterfactually imagined versus true memories during retrieval.

EEG evidence has revealed different memory processes that occur at early and late stages of retrieval with near real-time temporal resolution. Upon encountering a cue that overlaps with a stored memory, rapid and relatively automatic neurocognitive processes such as priming and familiarity are reflected in EEG within the first few hundred milliseconds (Rugg & Curran, 2007). Such item-specific processes are then followed by recollection of the associated episodic context from around 500 ms post-cue, via reconstructive processes involving pattern completion in the hippocampus (Staresina & Wimber, 2019). Recollection is typically reflected by a parietal ERP positivity around 500–800 ms (Rugg & Curran, 2007) and an increase in theta (4–8 Hz) oscillation power around the same time (Herweg et al., 2020). After the initial theta increase, EEG oscillations typically also show a decrease in alpha/beta power (~800–1500 ms), which is thought to reflect activation of cortical memory content during retrieval (Griffiths et al., 2019; Martín-Buro et al., 2020).

Additional post-retrieval processes occur after initial activation of stored memory content, and are correlated with relatively late (after 1 s post-cue onwards) EEG effects. In ERPs, late frontal positive slow drifts and posterior negativities are typically enhanced when memory retrieval requires increased and/or extended monitoring and control during retrieval competition (Bergström

et al., 2013; Hayama et al., 2008; Hellerstedt et al., 2016; Johansson et al., 2007; Mecklinger et al., 2016). In the oscillatory domain, an increase in alpha/beta power is associated with cognitive control across a variety of situations such as high working memory load, suppression of distracting input, and inhibition of competing memories during selective memory recall (Jensen & Mazaheri, 2010; Waldhauser et al., 2012). Here, we investigated if late ERP and EEG effects related to retrieval monitoring and control are observed when people need to distinguish between true and imagined memories.

In the current study, participants first performed real actions with everyday objects, such as rolling dice and putting on sunglasses (based on Brandt et al., 2014), thereby encoding strong and sensorimotor rich memories. One week later, they returned to the lab and were shown a series of photographs of the objects together with a brief text description of an action, and were asked to imagine performing each listed action with the object. For some objects, participants were asked to counterfactually imagine a novel action (for example, stacking the dice) whilst for other objects they imagined the originally performed action (thereby rehearsing it). We then tested memory for the original action from day one with both cued recall and associative recognition tests. The cued recall test presented only the object pictures as cues, and we therefore expected counterfactual imagination to cause retrieval competition and a need for selecting between original and imagined action memories on this test. The recognition test included the object pictures paired with either the original action text or a completely new action text (different from the counterfactually imagined action), and was therefore designed to be less susceptible to retrieval competition than the cued recall test. Hence, the recognition test should be more indicative of whether the original memories were available or forgotten (Anderson, 2003; Hicks & Starns, 2004). EEG was recorded during the cued recall test. Memory performance, ERPs and EEG oscillations were compared across counterfactually imagined and rehearsed conditions against a baseline condition with cues that had not been presented in the imagination task.

We expected counterfactual imagination to impair memory for the true action more than simple forgetting over time, which would be evident by lower memory accuracy in the imagination condition when compared to baseline. If this impairment was caused by reduced availability of the original memory, parallel reductions should be observed on both the cued recall and recognition tests. In contrast, if the original memories were intact but difficult to recall because of competition from the imagined

memories, we expected the impairment to be restricted to the cued recall test. We also measured participants' self-reported confidence in their responses in order to determine whether counterfactual imagination reduced memory confidence, or whether counterfactually imagined actions would be misremembered with high confidence similar to other imagination-induced false memories (e.g., Goff & Roediger, 1998).

In terms of EEG activity during the cued recall test, we expected typical markers of initial memory retrieval to be enhanced in both the rehearsed and imagined conditions compared to the baseline condition, due to enhanced cue familiarity followed by accurate recall in the rehearsed condition and accurate *and/or* erroneous recall in the imagined condition. Such effects would be indexed by enhanced positive ERP effects peaking approximately 300–800 ms (typical effects related to item recognition and contextual recollection; Rugg & Curran, 2007), an increase in theta power around the same time (Herweg et al., 2020), and a subsequent reduction in alpha/beta power (Griffiths et al., 2019; Martín-Buro et al., 2020). We also expected post-retrieval control and monitoring processes to be reflected in enhanced late ERP slow drifts and alpha/beta power increases (Hayama et al., 2008; Jensen & Mazaheri, 2010; Johansson et al., 2007; Waldhauser et al., 2012). If these late EEG effects index processes that resolve retrieval competition between competing true and imagined memories, they should be particularly enhanced following counterfactual imagination.

Methods

Participants

The final sample consisted of 36 participants ($M_{age} = 20$, $SD_{age} = 1.79$, range = 18–25, 25 females), who were right-handed, had English as their first language, had no psychiatric or neurological diagnosis and were not taking any psychoactive medication. This sample size was selected to achieve full counterbalancing of stimuli across conditions, and provided $>.80$ power to detect a medium effect size ($d = 0.5$) at $\alpha = .05$ (when testing for differences between pairwise conditions using paired t-tests). All participants were students at the University of Kent and were compensated either with payment (£25) or with course credits. All participants gave informed consent and the research was approved by the School of Psychology ethics committee at the University of Kent.

Materials

Stimuli consisted of 120 real everyday objects or object sets (see Figure 1), for example a pair of dice, a rubber duck, and a mobile phone. For each object set, three short sentences describing possible actions were generated, for example 'roll the dice,' 'stack the dice' or 'place dice with 5s up.' During counterfactual imagination and subsequent memory tests, photographs of the objects were used as cues together with the short sentences. Assignment of objects and associated sentences to the different experimental conditions was fully counterbalanced across participants.

Design and procedure

Encoding phase

The first day, participants encoded all 120 object-action pairs by performing each action with the relevant object. The experimenter first took each object out of a box, placed it at the center of the table and read aloud an action sentence. Participants then performed the action with the object. To ensure incidental encoding, participants were told that the experiment was about mental processes involved in performing everyday actions and were not told about the upcoming memory tests.

Manipulation phase

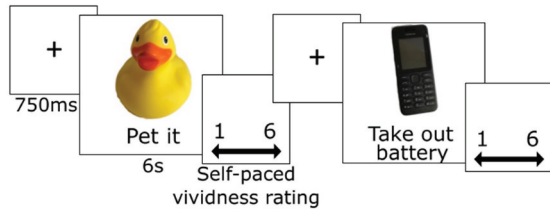
One week later, participants returned to the lab and completed the remaining tasks with EEG recordings. In the imagination task, they were shown photographs of 80 of the original 120 objects on a computer screen, each with an action sentence underneath it. For the 40 objects in the rehearsed condition, the same action as performed in session one was shown, and for the 40 objects in the imagination condition a new, counterfactual action was shown. The final 40 objects were assigned to the baseline condition and so were not included in this task. Participants were informed that some actions might match the action they had performed in session one, but that it was not relevant to the task so they should attempt not to think back to the first session and only focus on the action shown on the screen. To encourage vivid imagination, participants were told to think about as many of their senses as they could (e.g., what would the object feel like/smell like/sound like) and to imagine how they would move their body for each action.

Each trial started with a fixation cross for 750 ms, followed by the object picture-action sentence pairs for 6000 ms and ended with a self-paced rating scale where participants used the keyboard to rate how vivid their imagination of the action felt on a scale from one ('absence of imagination') to six ('comparable to real-life'). The 80 object-action pairs were repeated three

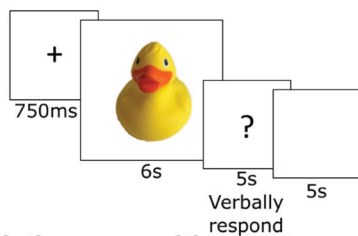
Encoding Phase



Manipulation phase



Cued recall test with EEG recording



Associative recognition test

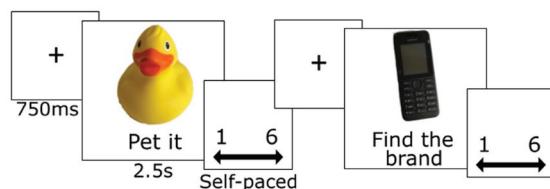


Figure 1. Overview of methodology with example stimuli.

times in a randomized order. For the second and third repetitions the sentences were shortened to reduce eye movement noise in the EEG recording.

Cued recall test

Next, participants were tested on their memory for the original action associated with each of the 120 objects from day 1 (resulting in three conditions; Imagined, Rehearsed, and Baseline). Participants were instructed to ignore anything they had read/imagined in the manipulation task and to only think of actions they had performed in the week previously. After a fixation cross for 750 ms, the photo of one object was shown for 6000 ms, followed by a question mark for 5000 ms which prompted participants to verbally describe the action they had performed with that object and rate their confidence on a scale of one ('not sure at all') to six ('definitely remember'). The trial ended with a blank screen

for 5000 ms. Audio was recorded for this task and verbal responses were coded by the experimenters.

Associative recognition test

Finally, participants were given an associative recognition test for all 120 object picture-action sentence pairs again that either included the action from the first session ('old') or an entirely new action, not previously enacted nor imagined during the manipulation phase ('new'). Each trial started with a fixation cross for 750 ms, the object-action pairs were then shown for 2500 ms, followed by a self-paced keyboard press response on a scale from one ('definitely new') to six ('definitely old') to capture both their recognition response and confidence.

A questionnaire was administered after all tasks to assess participants compliance with task instructions, which was generally high and no participants were excluded due to lack of compliance (see Supplementary file).

EEG recording and analysis

EEG was recorded between 0.1-70 Hz at a 500 Hz sampling rate from 29 scalp electrodes placed in the extended 10–20 system using an EasyCap (Brain Products GmbH), referenced against an average reference. Vertical and horizontal EOG was also recorded. Pre-processing was completed using the EEGLAB toolbox for MATLAB (Delorme & Makeig, 2004). Data were first re-referenced to the average of mastoid channels. Epochs from –600 ms to 3500 ms were extracted, time-locked to the presentation of the cued recall stimuli, and baseline corrected using the 200 ms pre-stimulus interval. Excessively noisy segments and/or channels were removed prior to independent component analysis (ICA) decomposition. We manually inspected the ICA results and removed components relating to eye movements and other forms of noise. The data was then low-pass filtered at 40 Hz, segments of data and/or channels with remaining excessive noise were removed, and any deleted channels were interpolated.

Pre-processed epochs from the cued recall test¹ were split into conditions regardless of accuracy, with similar trial numbers for Rehearsed (Range = 34–40, $M = 38$, $SD = 2$), Imagined (Range = 31–40, $M = 38$, $SD = 3$) and Baseline (Range = 32–40, $M = 39$, $SD = 2$) conditions. All

¹We also analyzed EEG recorded during the imagination task, by comparing ERPs and oscillations associated with counterfactual imagination versus imagination of a performed action using the same statistical methods and parameters as in the main reported analysis. There were no significant differences between conditions, consistent with the notion that counterfactual imagination and veridical memory retrieval engages highly overlapping neural processes (e.g., Schacter & Addis, 2020). The associative recognition test had too few trials per condition for clean EEG data since cues were shown with action statements resulting in excessive noise from eye-movements, and those conditions also required splitting trials by old/new status of the action statement. Hence this paper reports the cued recall results only.

trials were included to avoid ‘diluting’ effects of counterfactual imagination on subsequent retrieval processes, since such effects may be largest on trials where participants make a cued recall error (if those require additional retrieval monitoring/control). Including all trials also ensured stable estimates of effects and equated signal-to-noise ratios across all conditions. These epochs were averaged into ERPs for each condition, and were also submitted to time-frequency decomposition to extract oscillatory power, using the Fieldtrip toolbox for MATLAB (Oostenveld et al., 2011). Induced and evoked oscillatory power across 4–30 Hz was estimated using a wavelet transform, and represented on a decibel scale against an extended baseline between –825 and –375 ms pre-stimulus (see supplementary file).

EEG statistical analysis

We used nonparametric cluster-based permutation tests (Maris & Oostenveld, 2007) for statistical analysis of both ERPs and EEG oscillation power, in order to conduct a data-driven search for pairwise differences between conditions whilst controlling for multiple comparisons. This statistical approach is a commonly used method when analyzing EEG data from new tasks, like the one in our experiment, for which no specific time-windows, frequencies, or electrodes of interest can be selected a priori based on previous literature. For each comparison, ERP/power differences between conditions that extended across time and space (and frequency for oscillatory power) were identified as clusters and tested for significance using permutation resampling from the data, thereby protecting against inflated false positives (see supplementary file).

Results

Behavioural results

Imagination vividness

We first tested whether participants’ subjective imagination vividness during the manipulation phase predicted false memory for the counterfactually imagined action. Participants rated the vividness of their imagination significantly higher for rehearsed actions that would subsequently be recalled correctly ($M = 4.82$, $SD = 0.77$) than for counterfactually imagined actions, both when they would later correctly recall the original actions ($M = 4.42$,

$SD = 0.92$; $t(35) = 3.12$, $p = 0.003$, $d = 0.46$); Cohen’s d calculated for all analyses as the difference between means divided by the pooled standard deviation (Dunlap et al., 1996); and when they would later incorrectly recall the counterfactual action ($M = 4.60$, $SD = 0.84$; $t(35) = 2.08$, $p = 0.044$, $d = 0.27$). Counterfactually imagined actions that would later be recalled in error were rated as (trend level) more vivid than counterfactually imagined actions where participants would later recall the correct original action ($t(35) = 1.98$, $p = 0.055$, $d = 0.20$; see supplementary Fig. S1).

Cued recall test performance

Next, we tested the critical predictions that counterfactual imagination would impair participants’ later ability to recall the correct action. Figure 2 (top row) shows cued recall accuracy and mean confidence scores (averaged across accurate and inaccurate responses) for all conditions. Unsurprisingly, participants were better able to recall the original actions after rehearsing them compared to both after counterfactual imagination ($t(35) = 15.52$, $p < .001$, $d = 3.42$) and in the baseline condition ($t(35) = 17.26$, $p < .001$, $d = 2.72$). Importantly, cued recall accuracy was significantly reduced after counterfactual imagination compared to baseline ($t(35) = -3.80$, $p < .001$, $d = -0.65$). A supplementary analysis showed that this impairment was primarily caused by erroneous recall of the counterfactually imagined actions instead of original actions (see supplementary file), whereas failures to respond with any action occurred more often in baseline than the imagined condition ($t(35) = 3.64$; $p < .001$; $d = 0.45$). Parallel with the accuracy results, participants also rated their recall confidence higher for the rehearsed condition compared to both the imagined condition ($t(35) = 7.55$, $p < .001$, $d = 0.94$) and the baseline condition ($t(35) = 8.58$, $p < .001$, $d = 1.14$). However, confidence was not significantly different between the imagined and baseline conditions ($t(35) = 1.09$, $p = 0.29$, $d = 0.13$).

Associative recognition performance

Associative recognition responses on the scale from one (‘definitely new’) to six (‘definitely old’) were recoded to separately measure categorical old/new response accuracy,² versus confidence independent of accuracy. Figure 2 (bottom row) shows the accuracy and

²We present analyses of raw accuracy scores because they capture recognition performance separately for old and new actions which is most informative in this design. A supplementary analysis (see supplementary file) investigated whether the manipulation conditions differentially influenced participants’ ability to discriminate between old versus new actions separately from their response bias (general tendency to respond old vs. new independently from discrimination). In summary, this analysis showed that the imagined condition was associated with a stricter response bias (a tendency to respond ‘new’) compared to the rehearsed and baseline conditions (which did not differ). The rehearsed condition was associated with higher discrimination performance than imagined and baseline conditions. However, discrimination was still significantly reduced for the imagined condition compared to baseline, consistent with impaired recognition ability following counterfactual imagination when the influence of response bias was removed.

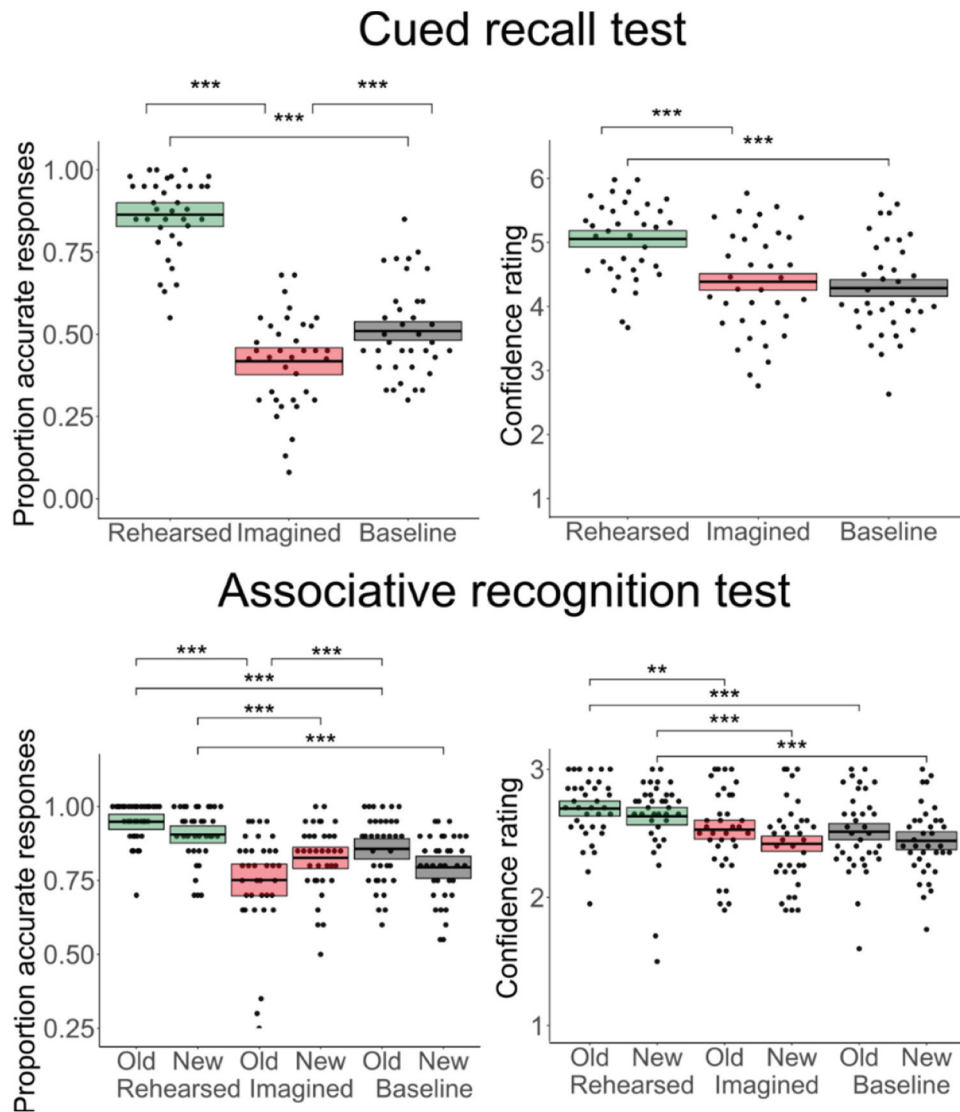


Figure 2. Memory test results. Proportion accurate responses (left panels) and mean confidence ratings regardless of accuracy (right panels) for the cued recall (top panels) and associative recognition tests (bottom panels). Dots show individual scores, the black line shows the mean score, and boxes indicate 95% confidence intervals. Significant differences indicated by brackets, *** $p < .001$, ** $p < .01$.

confidence scores (averaged across accurate and inaccurate responses) on this test.

Initial ANOVAs revealed a significant 3×2 interaction between manipulation condition (Rehearsed/Imagined/Baseline) and action type (Old/New) for accuracy

($F(2,70) = 8.27$, $p < .001$, $\eta p^2 = 0.19$) but not for confidence ($F(2,70) = 0.50$, $p = 0.61$, $\eta p^2 = 0.01$). Participants were significantly more accurate at recognizing old actions for the rehearsed condition than both imagined and baseline conditions (Table 1). Recognition of old

Table 1. Pairwise comparisons of accuracy and confidence on the associative recognition test.

	Accuracy			Confidence		
	<i>t</i>	<i>p</i>	<i>d</i>	<i>t</i>	<i>p</i>	<i>d</i>
Old actions						
Rehearsed vs. Imagined	6.40	< .001	1.55	3.49	.001	0.59
Rehearsed vs. Baseline	4.61	< .001	0.98	4.21	< .001	0.65
Imagined vs. Baseline	-3.97	< .001	0.74	0.42	.68	0.05
New actions						
Rehearsed vs. Imagined	4.20	< .001	0.76	5.10	< .001	0.69
Rehearsed vs. Baseline	5.32	< .001	1.10	4.31	< .001	0.66
Imagined vs. Baseline	1.43	.16	0.28	-0.58	.57	0.08

Note: *dfs* for all comparisons = 35.

actions was significantly lower in the imagined than baseline condition, converging with the cued recall results in showing impaired memory for original actions after counterfactual imagination. Accuracy at identifying new actions was significantly higher in the rehearsed condition than in imagined and baseline conditions, but there was no significant difference in accuracy for new actions between imagined and baseline conditions. For confidence, the same pattern of results was found regardless of whether objects were paired with old or new actions. For both action types, participants rated themselves as significantly more confident in the rehearsed condition than in both imagined and baseline conditions, with no significant difference in confidence between the latter.

Cued recall ERP results

As predicted, the rehearsed and the imagined conditions elicited initial ERP positivities, as shown by very similar, highly significant positive clusters for both the rehearsed-baseline (between 200 and 910 ms, $p < .001$) and imagined-baseline (between 270 and 940 ms, $p = .002$) comparisons (see Figure 3(b)). The clusters had a broad topography but peaked across the central and parietal scalp between ~ 300 and 750 ms, and they encompassed both typical early familiarity-related and later recollection-related ERP effects (see Figure 3(a)).

Following the initial positivities, both rehearsed and imagined conditions elicited negative ERPs compared to the baseline condition. A very sustained and broadly distributed significant negative effect was present in the rehearsed-baseline comparison ($p < .001$), with the cluster lasting between 1270 and 3000 ms and including most of the scalp electrodes except the occipital sites. For the imagined-baseline comparison, a briefer negative effect was evident from a significant negative cluster between 1160 and 2190 ms that peaked across the midline central and parietal scalp ($p < .001$). In the last part of the epoch, the rehearsed condition elicited more negative slow-drift ERPs than the imagined condition, between approximately 2160–3000 ms (split into two clusters: 2160–2430 ms, $p = 0.01$; and 2440–3000 ms, $p = .002$) with a central and right frontal scalp distribution. Therefore, whereas the negative effect for the rehearsed condition was very sustained, the imagined condition was similar to the baseline condition toward the end of the epoch, with both counterfactual imagination and baseline conditions eliciting more positive ERPs than the rehearsed condition.

Cued recall EEG oscillation results

Both rehearsed and imagined conditions elicited an initial reduction in oscillatory power compared to the baseline condition from around 0.5–1.5 s after cue presentation (see Figure 4). In the rehearsed-baseline comparison, a significant negative cluster ($p < .01$) spanned

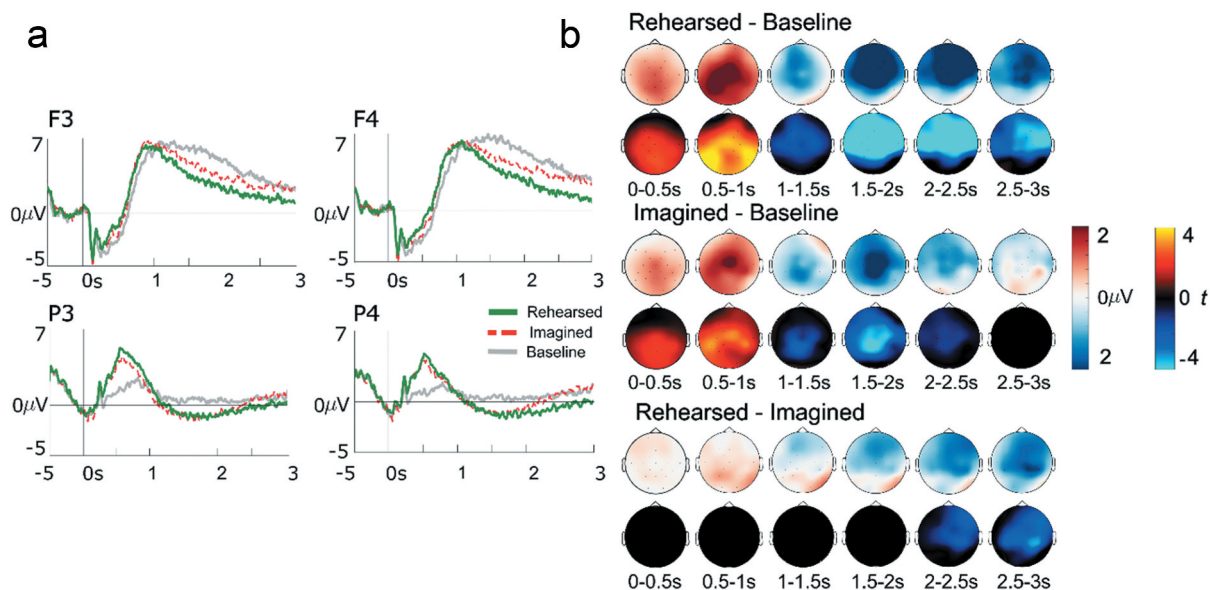


Figure 3. Cued recall ERP results. a. grand average ERP waveforms from left and right frontal (F3 and F4) and left and right parietal (P3 and P4) electrodes, time-locked to cue presentation in the cued recall test. b. topographical maps of the amplitude differences (top rows) and t-values for the differences (bottom rows) for all pairwise comparisons in the cued recall test. Only significant clusters are shown in the t-statistical maps.

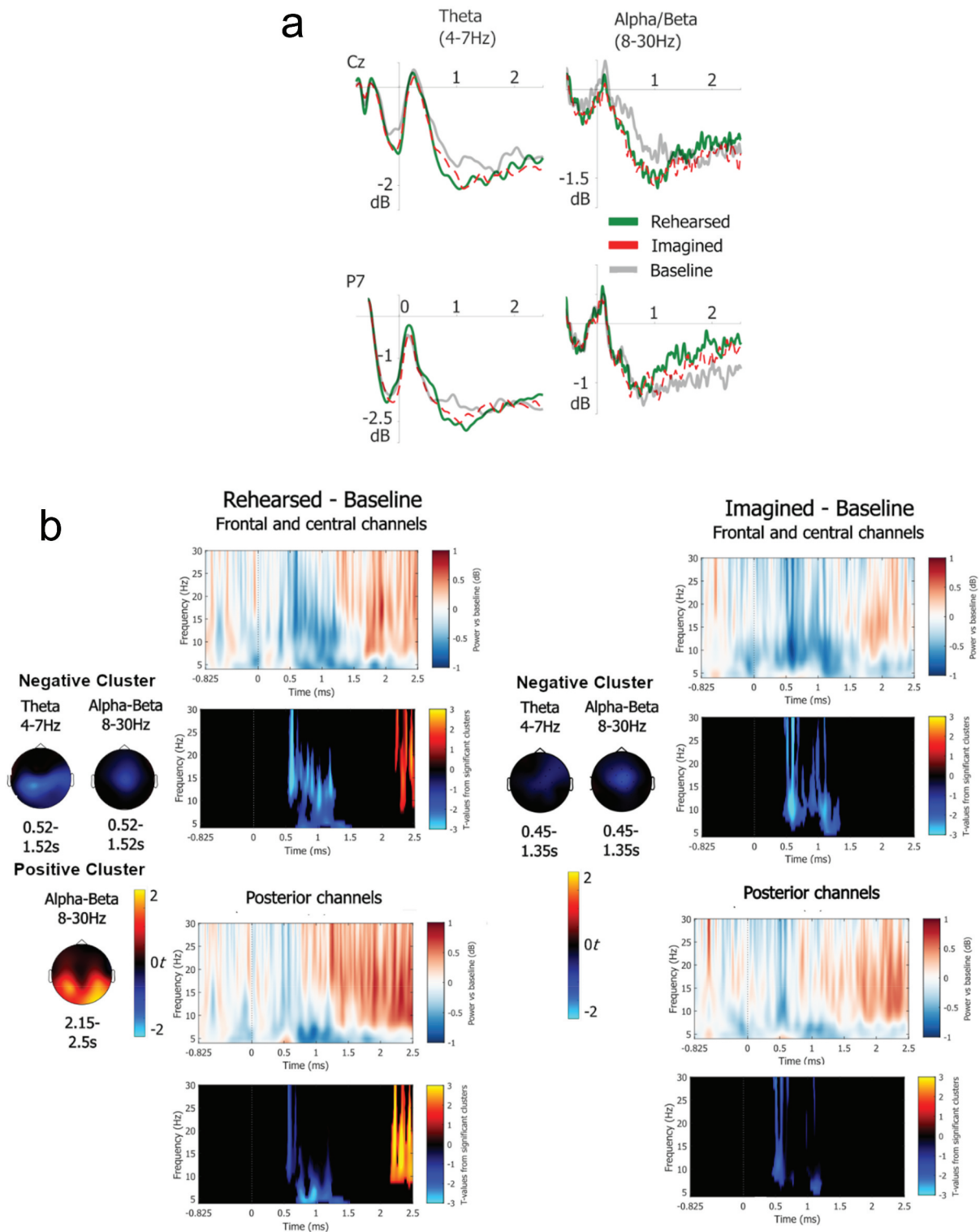


Figure 4. Cued recall EEG oscillation results. Results are visualised separately for theta (4-7Hz) and Alpha/Beta bands (8-30Hz) for illustration purposes. **a.** line plots showing mean power over time at a central (Cz) and left parietal (P7) site. **b.** time-frequency plots of power differences (top rows, blue/white/red colourmap) and t-values for the differences (bottom rows, cold/black/hot colour map) for pairwise comparisons averaged across frontal and central (F3, Fz, F4, FC1, FC2, C3, Cz, C4, CP1, CP2; top row) and posterior electrodes (CP5, CP1, CP2, P7, P3, Pz, P4, P8, O1, O2; bottom row) that showed significant effects. Topographies of significant clusters are shown on the left for each pairwise comparison. Only significant clusters are shown in the t-statistic plots.

from 530 to 1520 ms across all analyzed frequencies (4–30 Hz). Whereas this difference in the theta band (4–7 Hz) had a more posterior distribution, the alpha/beta band (8–30 Hz) topography peaked across midline frontal and central electrodes. A similar significant negative cluster ($p = 0.02$) was found for the imagined condition compared to the baseline condition, shifted slightly earlier at 450–1350 ms. This effect was strongest in the alpha/beta frequency band (specifically 8–20 Hz) with a mid fronto-central topography, but it also included frequencies in the theta band.

In the last part of the epoch, the direction of differences reversed with the rehearsed condition eliciting increased alpha/beta power compared to the baseline condition, as shown by a significant ($p = .01$) positive cluster from 2150 to 2500 ms including frequencies between 9 and 30 Hz. This effect had a bilateral posterior distribution across occipital and parietal electrodes. The imagined-baseline comparison showed no significant clusters in this time-window, although the difference in alpha/beta power was numerically in the same direction. There were no significant differences between the rehearsed and imagined conditions in any of the analyzed frequency bands and time periods. The oscillation results therefore mirror the ERP results in showing the largest and most sustained differences between rehearsed and baseline conditions.

Discussion

The role of imagination in forming false memories has been well studied (Goff & Roediger, 1998; Schacter & Addis, 2020; Thomas et al., 2003), whilst less is known regarding how imagining a counterfactual version of a previous event influences people's ability to remember what really happened. Here, we found novel evidence that counterfactual imagination impaired subsequent memory for true actions that had been performed with real objects. ERPs and oscillatory power during the subsequent cued recall test showed separable early effects relating to activation of memory content versus later effects which appear to index post-retrieval processes that are engaged when recall attempts are not immediately successful.

Participants encoded memories by performing actions with everyday objects, which is known to result in strong memories that are rich in sensorimotor details (Roberts et al., 2022). One week later they were cued to repeatedly imagine the same action or a new counterfactual action, and then completed surprise cued recall and associative recognition tests. As expected, rehearsing a previously performed action improved subsequent test accuracy and enhanced people's subjective

confidence in their memory responses on both tests, when compared to baseline object-action pairs that had been encoded in the first session but were not shown until the final tests. More importantly, both recall and associative recognition accuracy was lower after counterfactual imagination compared to baseline, showing a counterfactual imagination-induced memory impairment over and above nonspecific sources of forgetting over time. These findings extend on previous evidence that people can mistake counterfactually imagined actions for true memories (Gerlach et al., 2014; see also Dhammapeera et al., 2020) by showing that imagined counterfactuals can also disrupt their ability to remember true memories of performing actions.

Furthermore, since strong impairments were found not only during cued recall (which is highly vulnerable to retrieval competition) but also during associative recognition (which should be less vulnerable to retrieval competition; cf. Anderson, 2003; Hicks & Starns, 2004), our results suggest that the original memories may have become weakened or distorted as a result of counterfactual imagination. That is, whereas memories of imagined actions might compete during cued recall attempts and block access to memories of performed actions without any actual change to the original event memories, the true actions were actually presented on the associative recognition test and participants were required to distinguish those actions from completely new actions that had not been previously shown in the experiment. We used this design so that participants would be able to make recognition decisions based on a general assessment of memory strength (such as a sense of familiarity, see Hicks & Starns, 2004). Therefore, impaired recognition of old actions on this test suggests reduced memory *availability* rather than *accessibility* (Tulving & Pearlstone, 1966). However, this suggestion is tentative and requires further evidence since we cannot rule out that participants may have used a 'recall-to-reject' strategy during the associative recognition test (Rotello & Heit, 2000) by recalling the imagined action associated with the object and then making a source monitoring error by rejecting the true action. In order to determine the extent to which counterfactual imagination induces forgetting of the original event details, future research should test those memories using cues that are completely independent from the imagined event (Anderson, 2003).

Although counterfactual imagination reduced memory accuracy compared to baseline, it did not reduce confidence on either memory test. This pattern suggests that participants lacked full metacognitive awareness that counterfactual imagination had impaired their memory, in line with previous evidence that imagination

can produce high confidence memory distortions (Goff & Roediger, 1998). Furthermore, vividness ratings during the imagination task were higher if participants later erroneously reported the counterfactual action instead of the original action on the cued recall test. These findings suggest a specific role for vivid counterfactual imagination in producing false memories that supplant the true memories of events, consistent with past evidence and theoretical accounts suggesting that vivid imagination can become encoded into memories that appear similar to true experiences (Johnson et al., 1993; Schacter & Addis, 2020; Simons et al., 2017).

The EEG measures recorded during cued recall revealed that prior counterfactual imagination primarily influenced relatively late processes that are likely engaged at a post-retrieval stage. In contrast, early EEG markers of initial activation of memory content were enlarged for cues that had been shown in the imagination task regardless of whether participants had imagined true or counterfactual actions for those cues. In ERPs, an early (from ~200 ms) central and parietal positivity was present in both the rehearsed and imagined conditions compared to the baseline condition, lasting through the first second of the epoch. This positivity likely reflects both initial familiarity of the object cue and subsequent recollection of associated actions, regardless of whether that action had been performed or imagined (Rugg & Curran, 2007; Staresina & Wimber, 2019), consistent with the behavioral finding that participants were least likely to recall any action at all (correct or incorrect) in the baseline condition. Shortly after the cue (~500 ms) there was also a reduction in oscillatory power across all frequencies for the rehearsed and imagined conditions relative to the baseline condition. Alpha/beta power decreases with successful associative recall and may index activation of a core recollection network involving posterior parietal and medial temporal brain regions (Martín-Buro et al., 2020), explaining the reduction in power for higher frequencies. We did not however observe the predicted theta power increase in the early part of the epoch, as sometimes found during recollection (Herweg et al., 2020). The similarity in initial EEG markers of memory retrieval across rehearsed and imagined conditions is consistent with prior fMRI evidence that imagination produces memories that are neurally similar to memories of experienced events (De Brigard & Parikh, 2019; Schacter & Addis, 2020).

Later in the epoch (after ~2 s post-cue) both ERPs and oscillations showed a different pattern. Late ERP slow drifts in memory paradigms have been associated with retrieval monitoring, sustained memory search and control processes (Bergström et al., 2013; Johansson et al.,

2007; Mecklinger et al., 2016). We found that such late ERP slow-drifts were more positive for both baseline and imagined conditions compared to the rehearsed condition, suggesting that both of the former conditions recruited such additional processing to a larger extent. The similarity of these late ERP slow drifts across imagination and baseline conditions suggest that they do not reflect a brain process that is specifically involved in resolving retrieval competition between true and imagined memories. Instead, the late slow drifts seem linked with generally low accuracy and low confidence in memory responses, possibly indexing the engagement of additional and/or temporally extended monitoring processes when a correct answer was not readily available (e.g., Hayama et al., 2008).

Around the same time as the ERP slow-drift effects, EEG oscillations showed a late reduction in alpha/beta power for the baseline condition compared to the rehearsed condition, and a weaker and non-significant numerical reduction when compared to the imagined condition. As with the ERPs, because this effect was largest for the baseline vs. rehearsed comparison, it does not seem to reflect inhibitory control processes that are engaged to resolve retrieval competition between multiple memories associated to the same cue (Jensen & Mazaheri, 2010; Waldhauser et al., 2012), since such an account would predict largest alpha/beta power increases in the imagined condition. Instead, late alpha/beta power reductions for baseline items could reflect a delayed recall success effect (Martín-Buro et al., 2020) in this condition, assuming that recall of actions for baseline cues would be slowest. However, this account is tentative since it does not explain why the late baseline effect had a more occipital scalp distribution than the earlier alpha/beta reduction for rehearsed and imagined conditions. It is also unclear why there were no EEG oscillation differences between imagined and rehearsed conditions, when such differences emerged in the ERPs. These measures are sensitive to complementary aspects of neural activity so inconsistencies are not uncommon (discussed in e.g., Hellerstedt et al., 2021); however, the statistical power of the analysis also differed across the two measures. Since oscillation analysis includes a third frequency dimension, correcting for the larger number of multiple comparisons resulted in a stricter statistical threshold when compared to the ERP analysis, which could explain why additional effects were detected in the ERPs. Nevertheless, the ERPs and oscillations provide novel evidence to suggest that participants engaged in similar late retrieval processes both when recalling weak memories and when trying to distinguish between competing memories of imagined and true versions of events.

However, understanding the exact functional role of these late retrieval processes requires further research.

In conclusion, our behavioral and neural findings show that vivid counterfactual imagination can impair true memories, even for strongly encoded and sensorimotor rich memories of interacting with the world. These results therefore demonstrate how our everyday reflections on how past events might have turned out differently can change our beliefs about what really happened, due to the malleable nature of memory.

Acknowledgments

We thank Georgios Kalimitridis and Juselle Aquino for help with data collection. Z.M.B. and R.H. were supported by a Leverhulme Research Project Grant (RPG-2016322).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Phot Dhammapeera  <http://orcid.org/0000-0003-2447-7893>
 Chloe Brunskill  <http://orcid.org/0000-0001-6928-9461>
 Robin Hellerstedt  <http://orcid.org/0000-0002-6433-4734>
 Zara M. Bergström  <http://orcid.org/0000-0001-8794-0182>

References

- Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language*, 49(4), 415–445. <https://doi.org/10.1016/j.jml.2003.08.006>
- Anderson, M. C., & Hulbert, J. C. (2021). Active forgetting: Adaptation of memory by prefrontal control. *Annual Review of Psychology*, 72(1), 1–36. <https://doi.org/10.1146/annurev-psych-072720-094140>
- Bergström, Z. M., Henson, R. N., Taylor, J. R., & Simons, J. S. (2013). Multimodal imaging reveals the spatiotemporal dynamics of recollection. *Neuroimage: Reports*, 68, 141–153. <https://doi.org/10.1016/j.neuroimage.2012.11.030>
- Brandt, V. C., Bergström, Z. M., Buda, M., Henson, R. N., & Simons, J. S. (2014). Did I turn off the gas? Reality monitoring of everyday actions. *Cognitive, Affective, & Behavioral Neuroscience*, 14(1), 209–219. <https://doi.org/10.3758/s13415-013-0189-z>
- De Brigard, F., Hanna, E., St Jacques, P. L., & Schacter, D. L. (2019). How thinking about what could have been affects how we feel about what was. *Cognition and Emotion*, 33(4), 646–659. <https://doi.org/10.1080/02699931.2018.1478280>
- De Brigard, F., & Parikh, N. (2019). Episodic counterfactual thinking. *Current Directions in Psychological Science*, 28(1), 59–66. <https://doi.org/10.1177/0963721418806512>
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Dhammapeera, P., Hu, X., & Bergström, Z. M. (2020). Imagining a false alibi impairs concealed memory detection with the autobiographical implicit association test. *Journal of Experimental Psychology: Applied*, 26(2), 266–282. <https://doi.org/10.1037/xap0000250>
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1(2), 170–177. <https://doi.org/10.1037/1082-989X.1.2.170>
- Gerlach, K. D., Dornblaser, D. W., & Schacter, D. L. (2014). Adaptive constructive processes and memory accuracy: Consequences of counterfactual simulations in young and older adults. *Memory*, 22(1), 145–162. <https://doi.org/10.1080/09658211.2013.779381>
- Goff, L. M., & Roediger, H. L. (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recollections. *Memory & Cognition*, 26(1), 20–33. <https://doi.org/10.3758/bf03211367>
- Griffiths, B. J., Mayhew, S. D., Mullinger, K. J., Jorge, J., Charest, I., Wimber, M., & Hanslmayr, S. (2019). Alpha/Beta power decreases track the fidelity of stimulus-specific information. *eLife*, 8, e49562. <https://doi.org/10.7554/eLife.49562>
- Hayama, H. R., Johnson, J. D., & Rugg, M. D. (2008). The relationship between the right frontal old/new ERP effect and post-retrieval monitoring: Specific or non-specific? *Neuropsychologia*, 46(5), 1211–1223. <https://doi.org/10.1016/j.neuropsychologia.2007.11.021>
- Hellerstedt, R., Johansson, M. (2016). Competitive semantic memory retrieval: Temporal dynamics revealed by event-related potentials. *Public Library of Science One*, 11(2), e0150091. <https://doi.org/10.1371/journal.pone.0150091>
- Hellerstedt, R., Moccia, A., Brunskill, C. M., Bowman, H., & Bergström, Z. M. (2021). Aging reduces EEG markers of recognition despite intact performance: Implications for forensic memory detection. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 140, 80–97. <https://doi.org/10.1016/j.cortex.2021.03.015>
- Herweg, N. A., Solomon, E. A., & Kahana, M. J. (2020). Theta oscillations in human memory. *Trends in Cognitive Sciences*, 24(3), 208–227. <https://doi.org/10.1016/j.tics.2019.12.006>
- Hicks, J. L., & Starns, J. J. (2004). Retrieval-induced forgetting occurs in tests of item recognition. *Psychonomic Bulletin & Review*, 11(1), 125–130. <https://doi.org/10.3758/BF03206471>
- Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Frontiers in Human Neuroscience*, 4, 4. <https://doi.org/10.3389/fnhum.2010.00186>
- Johansson, M., Aslan, A., Bäuml, K.-H., Gäbel, A., & Mecklinger, A. (2007). When remembering causes forgetting: Electrophysiological correlates of retrieval-induced forgetting. *Cerebral Cortex*, 17(6), 1335–1341. <https://doi.org/10.1093/cercor/bhl044>

- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114(1), 3–28. <https://doi.org/10.1037/0033-2909.114.1.3>
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- Martin-Buro, M. C., Wimber, M., Henson, R. N., & Staresina, B. P. (2020). Alpha rhythms reveal when and where item and associative memories are retrieved. *The Journal of Neuroscience*, 40(12), 2510–2518. <https://doi.org/10.1523/JNEUROSCI.1982-19.2020>
- Mecklinger, A., Rosburg, T., & Johansson, M. (2016). Reconstructing the past: The late posterior negativity (LPN) in episodic memory studies. *Neuroscience and Biobehavioral Reviews*, 68, 621–638. <https://doi.org/10.1016/j.neubiorev.2016.06.024>
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, 156869. <https://doi.org/10.1155/2011/156869>
- Roberts, B. R. T., MacLeod, C. M., & Fernandes, M. A. (2022). The enactment effect: A systematic review and meta-analysis of behavioral, neuroimaging, and patient studies. *Psychological Bulletin*, 148(5–6), 397–434. <https://doi.org/10.1037/bul0000360>
- Roese, N. J., & Epstude, K. (2017). The functional theory of counterfactual thinking: New evidence, new challenges, new insights. In *Advances in experimental social psychology* (pp. 1–79). Elsevier Academic Press. <https://doi.org/10.1016/bs.aesp.2017.02.001>
- Rotello, C. M., & Heit, E. (2000). Associative recognition: A case of recall-to-reject processing. *Memory & Cognition*, 28(6), 907–922. <https://doi.org/10.3758/BF03209339>
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, 11(6), 251–257. <https://doi.org/10.1016/j.tics.2007.04.004>
- Schacter, D. L., & Addis, D. R. (2020). Memory and imagination: Perspectives on constructive episodic simulation. In A. Abraham (Ed.), *The Cambridge handbook of the imagination* (pp. 111–131). Cambridge University Press. <https://doi.org/10.1017/9781108580298.008>
- Simons, J. S., Garrison, J. R., & Johnson, M. K. (2017). Brain mechanisms of reality monitoring. *Trends in Cognitive Sciences*, 21(6), 462–473. <https://doi.org/10.1016/j.tics.2017.03.012>
- Staresina, B. P., & Wimber, M. (2019). A neural chronometry of memory recall. *Trends in Cognitive Sciences*, 23(12), 1071–1085. <https://doi.org/10.1016/j.tics.2019.09.011>
- Thomas, A. K., Bulevich, J. B., & Loftus, E. F. (2003). Exploring the role of repetition and sensory elaboration in the imagination inflation effect. *Memory & Cognition*, 31(4), 630–640. <https://doi.org/10.3758/BF03196103>
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning & Verbal Behavior*, 5(4), 381–391. [https://doi.org/10.1016/S0022-5371\(66\)80048-8](https://doi.org/10.1016/S0022-5371(66)80048-8)
- Waldhauser, G. T., Johansson, M., & Hanslmayr, S. (2012). Alpha/Beta oscillations indicate inhibition of interfering visual memories. *The Journal of Neuroscience*, 32(6), 1953–1961. <https://doi.org/10.1523/JNEUROSCI.4201-11.2012>