Minor Changes Change Memories: Functional Magnetic Resonance Imaging and Behavioral Reflections of Episodic Prediction Errors

Sophie Siestrup[1,2] and Ricarda I. Schubotz[1,2]

Abstract

■ Episodic memories can be modified, a process that is potentially driven by mnemonic prediction errors. In the present study, we used modified cues to induce prediction errors of different episodic relevance. Participants encoded episodes in the form of short toy stories and then returned for an fMRI session on the subsequent day. Here, participants were presented either original episodes or slightly modified versions thereof. Modifications consisted of replacing a single object within the episode and either challenged the gist of an episode (gist modifications) or left it intact (surface modifications). On the next day, participants completed a post-fMRI memory test that probed memories for originally encoded episodes. Both types of modifications triggered brain activation in regions we previously found to be involved in the processing of content-based mnemonic prediction errors (i.e., the exchange of an object). Specifically, these were ventrolateral pFC, intraparietal cortex, and lateral occipitotemporal cortex. In addition, gist modifications triggered pronounced brain responses, whereas those for surface modification were only significant in the right inferior frontal sulcus. Processing of gist modifications also involved the posterior temporal cortex and the precuneus. Interestingly, our findings confirmed the posterior hippocampal role of detail processing in episodic memory, as evidenced by increased posterior hippocampal activity for surface modifications compared with gist modifications. In the post-fMRI memory test, previous experience with surface modified, but not gist-modified episodes, increased erroneous acceptance of the same modified versions as originally encoded. Whereas surface-level prediction errors might increase uncertainty and facilitate confusion of alternative episode representations, gist-level prediction errors seem to trigger the clear distinction of independent episodes. ■

INTRODUCTION

Episodic memories allow us to vividly remember events from our personal past (Tulving, 2002). These memories are not always exact but can be modified on the basis of new relevant information (Lee, Nader, & Schiller, 2017; Scully, Napper, & Hupbach, 2017; Nader, 2015; Nader & Einarsson, 2010). Such modifications are potentially fueled by mnemonic prediction errors (PEs), which occur when there is mismatch between what was expected based on memories and the true situational evidence (e.g., Siestrup et al., 2022; Sinclair & Barense, 2019). In that sense, PEs serve as learning signals to update internal predictive models, so that we can maintain valid predictions in the long run (Schubotz, 2015; Friston & Kiebel, 2009; Friston, 2005).

Memories include different kinds of details about experienced episodes that potentially differ in their relevance for our predictions. Although some perceptual and contextual details are not relevant to the overall storyline of an episode, there are central details that represent the episode's gist (Robin & Moscovitch, 2017). It has been suggested that both types of episode information, gist relevant and surface-level details, are encoded and support memory retrieval (e.g., Brainerd & Reyna, 2002). Although some argue that the gist and details of episodes are processed in two different traces (fuzzy trace theory; Brainerd & Reyna, 2002), a different view is that episodic memories can generally be assessed from different levels of specificity, which exist on a continuum (Greene & Naveh-Benjamin, 2021). The gist of an episode can be defined as "the central features of a particular episode (story line)" (Robin & Moscovitch, 2017). Noncentral details of an episode are much more likely to be forgotten (Sekeres et al., 2016) or influenced by misinformation, specifically if such misinformation is in line with the gist of the episode (Reyna, Corbin, Weldon, & Brainerd, 2016; Reyna & Lloyd, 1997). Furthermore, gist information is highly relevant for guiding judgment and decision-making (Corbin, Reyna, Weldon, & Brainerd, 2015). It has even been proposed that the episodic memory trace only represents the gist of an episode, whereas noncentral details are supplemented from semantic memory during retrieval (Cheng, Werning, & Suddendorf, 2016). Taken together, gist-relevant and

[1]University of Münster, Germany, [2]Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience, University of Münster, Germany

surface-level details likely differ in their predictive relevance in the context of episodic memory. Mnemonic PEs based on the two types of information should therefore elicit different neural and behavioral responses.

The aim of the present study was to investigate how unexpected episodic modifications that either do or do not change the gist of an episode influence brain activity and memory. We previously demonstrated that subtle changes during episodic cueing lead to specific brain responses for qualitatively different types of modifications, either affecting episode content or structure (Siestrup et al., 2022). Furthermore, encounters with both types of modified episodes induced false memories in a post-fMRI memory test (Siestrup, Jainta, Cheng, & Schubotz, 2023; Jainta et al., 2022; Siestrup et al., 2022). In the present work, we focused on content modifications and specifically manipulated the impact of those on the episodes' gist.

In a first experimental session, participants encoded different episodes consisting of short toy stories from video material. On the next day, participants returned for an fMRI session during which they were presented originally encoded episodes or slightly modified versions thereof. Episode videos were modified by the exchange of a single object (Siestrup et al., 2022). Although some of these modifications were irrelevant for the episode's storyline (in the following termed *surface modifications*), others intentionally changed it (in the following termed *gist modifications*). This intended impact of the content modifications was validated in a behavioral pilot study. On a third day, participants completed a post-fMRI memory test to probe memory for originally encoded episodes. To this end, participants rewatched original as well as modified episode videos and had to decide if these had been encoded during the first session. Lastly, participants went through a rating task to assess their subjective evaluations of episode modifications.

We hypothesized that both types of modifications would trigger activation in specific brain regions we previously identified to respond to content modifications in our paradigm (Siestrup et al., 2022), including occipitotemporal cortex (OTC), fusiform gyrus (FG), (posterior) intraparietal sulcus (IPS), and ventrolateral pFC, specifically inferior frontal gyrus (IFG) and inferior frontal sulcus (IFS) in the right hemisphere. These brain regions are known to be relevant for the processing of (new) object information in actions (El-Sourani, Trempler, Wurm, Fink, & Schubotz, 2019; Lingnau & Downing, 2015; Wiggett & Downing, 2011; Reber, Gitelman, Parrish, & Mesulam, 2005; Grill-Spector, Kourtzi, & Kanwisher, 2001). Modifications that change the gist of the story were expected to elicit more pronounced activation in said areas because of their overall higher relevance for episode content and therefore enhanced processing at both encoding and retrieval.

In addition, we expected that there would be differences in brain activation between the two types of content modifications, for example, in the hippocampal complex. When a modification leads to inconsistency with the known episode storyline, there could be increased activation in areas relevant to novelty processing. In a previous study, we found that the hippocampus (HPC) responded more strongly to novel compared with familiar episodes in our paradigm (Jainta et al., 2022). Therefore, it is possible that gist modifications elicit stronger hippocampal responses than surface modifications. Accordingly, parahippocampal areas were found to be included in the processing of scene gist (Schubotz & Von Cramon, 2009; Oliva & Torralba, 2006; Epstein, 2005). However, it may be that surface modifications trigger stronger activation in HPC than do gist modifications, as it has been suggested that the HPC is more relevant to details in episodes and that gist representations become independent of the HPC, but rather after longer delays between encoding and retrieval (e.g., Winocur & Moscovitch, 2011). It has also been demonstrated that episode details are predominantly represented in posterior HPC, whereas the anterior HPC processes episode gist (Sekeres, Winocur, & Moscovitch, 2018; Sekeres, Winocur, Moscovitch, Anderson et al., 2018; Robin & Moscovitch, 2017; Moscovitch, Cabeza, Winocur, & Nadel, 2016; Poppenk, Evensmoen, Moscovitch, & Nadel, 2013). Therefore, it is likely that both types of modifications trigger hippocampal activation, but the localization of activity within HPC was expected to differ between the two modification types. Specifically, activation in anterior HPC was expected to be larger for gist than surface modified cues, and vice versa in posterior HPC. Aside from that, there is evidence that episode gist is mediated by ACC (Sekeres, Winocur, & Moscovitch, 2018) and OFC (Schubotz & Von Cramon, 2009), so activation in these brain areas might be higher for gist than for surface modifications.

Concerning memory performance in the post-fMRI memory test, we expected to replicate our general finding that mnemonic PEs during the fMRI session trigger false memories afterward, putatively because of internal model updating (Schiffer, Ahlheim, Ulrichs, & Schubotz, 2013; Schiffer, Ahlheim, Wurm, & Schubotz, 2012). Therefore, one hypothesis was that prediction violations lead to increased acceptance of modified episode videos as originally encoded, potentially accompanied by a decreased acceptance for unmodified videos (Siestrup et al., 2023, 2022; Jainta et al., 2022). As mentioned above, memories for noncentral details of episodes are usually less accurate. Therefore, we reasoned that memory accuracy for originally encoded episodes, as assessed by false alarm and hit rates and, more generally, the area under the curve (AUC; Brady, Robinson, Williams, & Wixted, 2022), might be generally lower for surface than for gist-modified episodes. In addition, prediction violation through surface and gist modifications might influence subsequent memory differently, in line with previous reports that surface-level detail is more likely to be influenced by misinformation (Reyna et al., 2016; Reyna & Lloyd, 1997).

Thus, false memories might arise as a consequence of prediction violation because of surface, but not, or to a lesser extent, because of gist modifications.

METHODS

Participants

Forty-two volunteers (28 women, 14 men) participated in the study. All were native German speakers and right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). Participants had (corrected-to-) normal vision, had intact color perception, and reported no history of neurological or psychiatric disorders. Data from three participants were excluded from analyses because of increased movement during the fMRI session (two participants; ca. 5-mm and 7-mm movement) and a benign anatomical abnormality of the brain that impeded normalization (one participant). Two participants terminated the study preliminarily for personal reasons. The final sample consisted of 37 participants (25 women, 12 men) between the age of 18 and 29 years (M_{age} = 23 years, SD = 3.17 years). Similar sample sizes have previously yielded stable results (Jainta et al., 2022; Siestrup et al., 2022). All participants gave written informed consent to participate in this study and were reimbursed with course credits or money. The study was conducted in accordance with the Declaration of Helsinki and approved by the Local Ethics Committee of the University of Münster.

Stimuli

The video stimuli were very similar to those used in our previous studies (Jainta et al., 2022; Siestrup et al., 2022; available upon request at https://www.uni-muenster.de/IVV5PSY/AvicomSrv/). We worked with 80 short videos (duration = 8.8–14.76 sec, M = 12.31 sec), which showed an actress manipulating PLAYMOBIL toys. Only the toy objects and the hands and arms of the actress were visible and filmed from above with a digital single-lens reflex camera (Nikon D5300). For filming, the actress wore a black pullover and black rubber gloves. Matte white paper served as a base. Objects that were needed for a particular story were positioned on a table and were only moved into view in the moment at which they appeared in the story. Stories included six to nine action steps (M = 7.60 steps) and four to 11 separable objects (M = 6.23 objects). The exact same object appeared in only one of the stories.

Video material was edited with Adobe Premiere Pro CC (Adobe Systems Software, Version 12.1.2). Final video stimuli had a frame of size 1920 × 1080 pixels and a frame rate of 25 frames per second. They always started with seven frames showing only background (white base) and ended after seven frames showing the final object constellation. Throughout the study, videos were presented using the stimulus presentation software Presentation (Version 20.3 02.25.19, NeuroBehavioral Systems) at a visual angle of approximately 7.3° × 13°.

There were 24 stories that existed in three different versions each. The first version was the original version, which was used for encoding. In addition, we created two modified version. In those modified versions, one single object was exchanged compared with the original. The new object was never just the same object in a different color; the new object diverged from the old one in color and shape. In one modified version, the object exchange did not affect the storyline (i.e., gist) of the episode. In the following, this type of modification will be referred to as surface modification. In the other modified version, the new object changed the storyline, that is, affected the gist of the episode (gist modification). We validated our a priori classification of surface and gist modifications in a behavioral pilot study and confirmed that all modifications could be identified by participants. Modifications were not introduced in the first two action steps so that the beginning of a video served as a cue for prediction. No modifications were introduced in the last two action steps either. The exact time point of the modification in each video was determined by identifying the video frame that diverged from the original version. For an example of an episode and its modified versions, see Figure 1.

Two additional videos were used for practice trials for the encoding, fMRI, and rating tasks. One of those existed in only one version, and one additionally had two modified alternatives. Four more stories were first introduced in the fMRI experiment. They existed in one version only and will be referred to as novels in the following.

Procedure

Encoding

The encoding session was conducted in a computer laboratory at the University of Münster and lasted, on average, 2 hr 9 min. For encoding, participants sat at the setup where the stimulus material had been filmed and wore black gloves like those seen in the video. Episodes were encoded by watching each video 5 times. Afterward, participants had to give a detailed description to the experimenter, which had to include all objects (including their color) and the correct order of actions in the story. If the participant made a mistake, the experimenter interrupted and corrected them, and the participant had to start the description anew, until one completely correct description was given. On average, participants needed 1.54 description attempts per story (SD = 0.22). The context and interpretation of the story were discussed and clarified, to ensure that all participants encoded the story the same way. Lastly, the participant was asked to summarize the story in a short sentence. Each participant encoded all 24 episodes in a randomized order. The session started with a short practice phase (two videos) to familiarize the participant with the task. At the end of the encoding

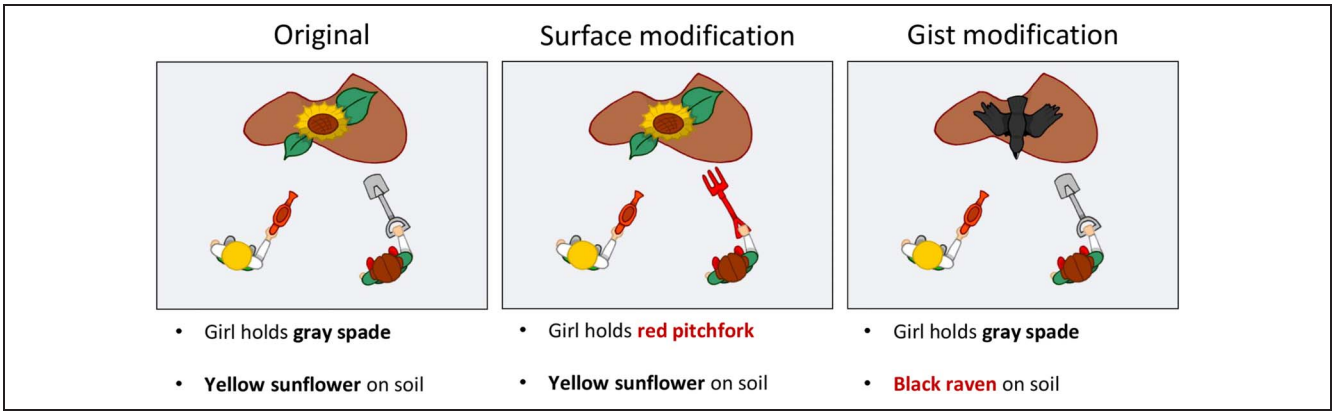

Figure 1. Example of episode modifications. Original (left) episodes were encoded by the participants. In one modified version, the object exchange did not affect the storyline of the episode (surface modification; here: Gray spade is exchanged for red pitchfork). In the other modified version, the new object changed the storyline (gist modification; here: Yellow sunflower is exchanged for black raven). Each participant only encountered one of the two modified version of an episode. We do not reproduce photos of our stimulus material because it is copyrighted material (PLAYMOBIL figures); instead, we provide schematic images.

session, participants briefly practiced the tasks they would do during the fMRI session.

MRI Session

The fMRI session was conducted on the day after the encoding. Participants went through two experimental parts, which are described in the following.

Incomplete reminder. During the anatomical measurements, participants went through a so-called incomplete reminder to initiate memory destabilization, presumably facilitating later modification (Sinclair & Barense, 2018, 2019). Each video that had previously been encoded was presented, but interrupted during the step preceding the modification that would be presented during the fMRI experiment (e.g., while retracting the hand). Videos that would be shown in the original version during the fMRI experiment were shown in interrupted fashion as well. The time point of the interruption then matched the modified version that was later presented as an alternative in the memory test. It has been shown that allowing for complete retrieval following an incomplete cue can enhance memory (Antony, Ferreira, Norman, & Wimber, 2017). Therefore, participants had to solve a dummy task between incomplete video trials to prevent complete retrieval. This task was introduced to the participants as a "warm-up-task." After each video, a number between 1 and 4 was presented for 300 msec and participants had to indicate via a button press whether the number was odd (index finger) or even (middle finger).

fMRI experiment. The fMRI experiment closely followed our previously published protocol (Jainta et al., 2022; Siestrup et al., 2022). During the fMRI session, participants were presented with videos reminiscent of the previously encoded episodes. Videos were either displayed in the original version as during encoding (eight videos) or in either the surface or gist modified version (eight videos each). Which stories belonged to which conditions was balanced over participants. As reported previously, all episodes were presented repeatedly to simulate natural circumstances that potentially benefit memory modification (Siestrup et al., 2022; Schiffer, Ahlheim et al., 2013; Schiffer et al., 2012). In addition, four novel stories were shown in the fMRI session for the first time.

The experiment consisted of six blocks. Each video was presented once per block. In addition, each block contained three null events during which only a fixation cross was presented (duration: 7–10 sec).

Participants were instructed to attentively watch the presented videos. They were told that after some videos, they were required to solve a quick task about the previously presented video. Although the participants were not aware of this, the task was merely included to ensure the participants' constant attention (Jainta et al., 2022; Siestrup et al., 2022; El-Sourani et al., 2019) and was not of interest for our actual research question. These task trials were characterized by the presentation of two verbs left and right of the fixation cross. One verb was closely associated with the video shown before; one was not at all associated with the video (as validated by three experimenters). Importantly, verbs did not describe any action actually depicted in the video, and all associations remained intact in the case of modifications. The participants' task was to press the left button (index finger) if the left verb matched the video, and to press the right button (middle finger) if the right verb matched the video. Whether the associated verb was presented on the left or right was pseudorandom and balanced between conditions for each participant. The verbs were presented for a maximum of 3 sec or until participants responded. Upon response delivery, participants received a 1-sec written feedback whether they answered correctly, incorrectly, or too late, in case no response was given. Each video was followed by a task trial twice during the experiment,

resulting in a total number of 56 task trials in the experiment.

A fixation cross was presented between trials for 2 sec (1 sec after task trials) to serve as an interstimulus interval. In addition, a variable jitter of 0, .5, 1, or 1.5 sec of fixation was added before each trial to enhance the temporal resolution of the BOLD response (Figure 2). In total, the fMRI task had a duration of approximately 48 min. Transition probabilities were balanced between different trial types over the whole experiment.

Post-fMRI Memory Test

The post-fMRI memory test was conducted on the day after the MRI session and followed a similar protocol as previously reported (Jainta et al., 2022; Siestrup et al., 2022). Participants were not informed that their memory for encoded episodes would be tested at any time during the experiment.

Participants were seated in front of a computer and instructed to remember the encoding session. Then, they were presented all stories in two different versions. When modified videos had been presented during the fMRI experiment, these versions were presented again in the memory test and, additionally, each original counterpart was presented. When original episodes had been presented during the fMRI experiment, they were presented again and, additionally, each story was displayed either with a surface or a gist modification (in half of the cases each). Videos were presented in a pseudorandomized order, so that half of the stories (of each experimental condition) were first presented in their original version followed by a modified version and vice versa. Novel videos were displayed twice in the same version. Versions of the same episode were not displayed in direct succession but with minimally two other videos in between. In total, the memory test comprised 56 video trials.

After each video, participants had to rate how confident they were that they had encoded exactly this episode, using a Likert scale from 1 to 6 (with 1 = *100% no*, 2 = *80% no*, 3 = *60% no*, 4 = *60% yes*, 5 = *80% yes*, 6 = *100% yes*). Participants were instructed to respond quickly and intuitively. The duration of the task was approximately 15 min.

Rating Task

The rating task was conducted after the post-fMRI memory test on the third day of the study. Each original video was presented, followed by the same modified version that had been presented in the fMRI and/or memory test. Participants had to press a button as soon as they noticed the change in the modified video. The video was played completely and not interrupted by the response. Participants were instructed to carefully watch the full video. They then were explicitly asked if they noticed a change or not and had to respond via a button press. This question was included to allow participants who missed to indicate the modification during the video to still rate the modification if they had seen it. If participants answered "no," no further questions were displayed. If they had noticed the change, they were asked how unexpected they perceived the content modification on a Likert scale from 1 to 6 (with 1 = *0% unexpected*, 2 = *20% unexpected*, 3 = *40% unexpected*, 4 = *60% unexpected*, 5 = *80% unexpected*, 6 = *100% unexpected*; PE rating). Importantly, they were instructed to rate the unexpectedness in reference to the original story and to not consider that they

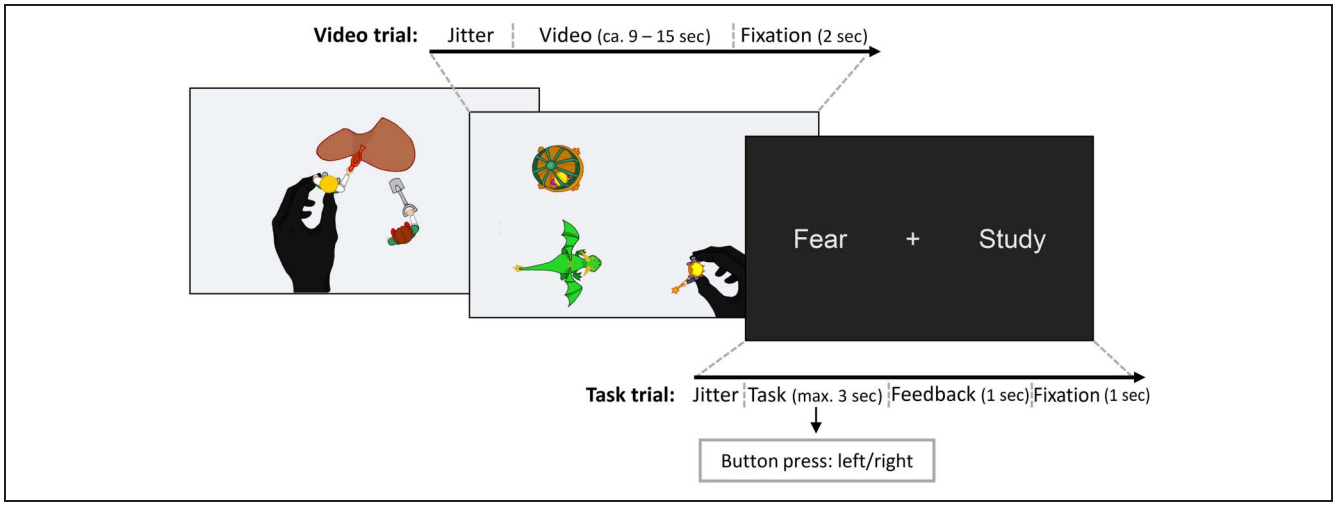

Figure 2. Schematic depiction of task during fMRI session. Video trials were composed of a variable jitter (0, 0.5, 1, or 1.5 sec of fixation), a video showing a toy story (ca. 9–15 sec), and a 2-sec interstimulus interval (fixation). Task trials included a variable jitter and the presentation of two verbs (maximally 3 sec long or terminated by response). The participants' task was to press the left or right button depending on which verb was more closely associated with the previously seen story. The 2-sec interstimulus interval after the task was divided into a 1-sec feedback ("correct," "incorrect," "too late") and a 1-sec fixation.

might have noticed the modification previously during the experiment. Next, they were asked how much the change affected the story of the episode on a Likert scale from 1 to 6 (with 1 = *0% different*, 2 = *20% different*, 3 = *40% different*, 4 = *60% different*, 5 = *80% different*, 6 = *100% different*; story-change rating). Lastly, the participant had to describe the modification orally to the experimenter. The task was self-paced so that participants always started a new video themselves via a button press. The duration of this task was approximately 20 min.

MRI Data Acquisition and Preprocessing

Magnetic resonance imaging was conducted with a 3-Tesla Siemens Magnetom Prisma MR tomograph (Siemens) using a 20-channel head coil. Participants lay supine on the scanner bed. Movements of head, arms, and hands were minimized by fixation with form-fitting cushions. The participants' right index and middle finger were positioned on the two appropriate buttons on a response box. To attenuate scanner noise, participants were provided with earplugs and headphones. Visual stimuli were presented via a screen that participants saw through a mirror mounted on the head coil.

Before functional imaging, high-resolution T1-weighted anatomical images were obtained with a 3-D magnetization prepared rapid gradient echo sequence (scanning parameters: 192 slices, slice thickness = 1 mm, repetition time = 2130 msec, echo time = 2.28 msec, flip angle = 8°, field of view = 256 × 256 mm^2). Functional images of the whole brain were acquired in interleaved order along the AC–PC plane using a gradient-echo EPI sequence to measure BOLD contrast (scanning parameters: 33 slices, slice thickness = 3 mm, repetition time = 2000 msec, echo time = 30 msec, flip angle = 90°, FoV = 192 × 192 mm^2).

Imaging data were preprocessed with SPM12 (Wellcome Trust) implemented in MATLAB (Version R2020b, MathWorks Inc.). Preprocessing included slice time correction to the middle slice, movement correction and realignment to the mean image, co-registration of the individual structural scans to the mean functional image, normalization of functional and structural images into the standard MNI space (Montreal Neurological Institute, Montreal, QC, Canada) on the basis of segmentation parameters, and spatial smoothing using a Gaussian kernel of FWHM of 8 mm. A 128-sec high-pass temporal filter was applied.

Statistical Data Analysis

fMRI Data Analysis

Design specifications. fMRI data were analyzed using general linear models (GLMs) for serially autocorrelated observations (Worsley & Friston, 1995; Friston et al., 1994) implemented in SPM12. Regressors were convolved with the canonical hemodynamic response function. We used the smoothed individual normalized gray matter image (8-mm FWHM), which was thresholded at .2 using ImCalc in SPM12 to create a binary gray matter mask that was applied at the first level of analyses.

The first model (GLM1) included 15 regressors. These were original videos (*original*), videos containing a surface modification (*surface*), and videos containing a gist modification (*gist*), each comprising 48 trials. For modified videos, onsets of events were time-locked to the point in the video at which the modification occurred (time of modification). For original trials, a hypothetical time of modification was calculated (mean of times that corresponded to points of surface and gist modifications in the nonmodified video) to serve as a comparable onset. As reported previously (Siestrup et al., 2022), these conditions were modeled as events to investigate phasic effect of the prediction violation at the precise moment of its occurrence. A parametric modulator was added to each of those regressors to model the repeated presentation of each video (descending coding). Please note that all parametric modulators were mean centered (Mumford, Poline, & Poldrack, 2015). The 24 novel videos were modeled as events, and onsets were timed to the middle of each video. In addition, regressors for the 18 null events and 56 task trials were included, with onsets time-locked to their respective onsets. Null events were modeled as epochs, including the full presentation time (7–10 sec). Task trials were modeled as events. The six subject-specific rigid-body transformations obtained from realignment were included as regressors of no interest.

The parametric modulation described above might not optimally account for differences between the first presentation of a modification (potentially the most surprising one) and all subsequent ones, as it assumes a linear decrease in activation. For this reason, we calculated a second model (GLM2) as an additional control for the repeated presentation of the same modification. The difference between GLM1 and GLM2 was that *original*, *surface*, and *gist* trials were separately modeled for the first presentation (eight trials per condition) and the remaining five presentations (40 trails per condition). No parametric modulators were included. We then calculated the same contrasts as for GLM1 (see below), but only including the second to sixth presentations of each condition. Qualitatively, results from whole-brain as well as ROI analyses were the same as for GLM1, with a slight decrease in power because of the reduced number of trials.

To control for participants' individual variation in the perception of surface and gist modification, as assessed by the rating task, we computed a third model (GLM3). Trials for original and novel videos, null events, and tasks were modeled as described for GLM1. Modified videos were not separated into surface and gist modifications, but included in the model as one regressor to which we added a parametric modulator composed of the

participants' story-change rating for each video obtained from the rating task (mean-centered). No further parametric modulators were added. Including movement parameters, GLM3 comprised 12 regressors. The parametric analysis of differences between surface and gist modifications gave rise to highly similar results as those we report for GLM1 with our a priori classification of modification types.

Whole-brain analysis. We calculated first-level *t*-contrasts for *surface* > *original* and *gist* > *original* as well as the direct contrasts *surface* > *gist* and *gist* > *surface* to analyze brain activity in response to the different modification types. In addition, we calculated the reverse contrasts, *original* > *surface*, and *original* > *gist* to further validate the specificity of modification responses. To investigate potential attenuation effects in brain responses because of repeated presentations of the same episodic modifications, we additionally computed the contrasts for the parametric modulators for surface and gist modifications against the implicit baseline. The corresponding results are reported in the appendix (Appendix A). Second-level group analyses were conducted with one-sample *t* tests across participants. Furthermore, we computed the first-level *t*-contrasts for each type of video versus novel videos (*original* > *novel*, *surface* > *novel*, *gist* > *novel*) and created the conjunction of these on the second level (Nichols, Brett, Andersson, Wager, & Poline, 2005) to replicate our previous findings of episodic retrieval in general (Jainta et al., 2022; Siestrup et al., 2022). We applied false discovery rate (FDR) correction with a threshold of $p < .05$ or higher (voxel level) to resulting *t* maps. For the conjunction contrast concerning episodic retrieval, we applied a more conservative threshold of $p < .001$ because due to the very strong activation, a threshold of $p < .05$ would not have been conclusive for relating activation to anatomical regions. For completeness of our analysis and with regard to our hypotheses, when no significant activation could be detected using this threshold, we applied a threshold of $p < .001$ (uncorrected). Generally, we only report clusters with a minimum number of 10 voxels. Brain activation was visualized with MRIcroGL (Version 1.2.20200331, McCausland Center for Brain Imaging, University of South Carolina).

ROI analysis. With regard to our hypotheses, we performed ROI analyses to more specifically investigate which brain areas are involved in the processing of surface and gist modifications. We hypothesized that both, surface and gist modifications, activate brain regions we identified to be involved in the processing of content modifications in general in our paradigm. We used functional ROIs of these regions, namely, right IFS, bilateral OTC and IPS (Siestrup et al., 2022), and extracted mean contrast estimate values from the *surface* > *original* and *gist* > *original* contrasts using the MarsBar Toolbox (Brett, Anton, Valabregue, & Poline, 2002). Functional ROIs were created from the peak voxel coordinates we identified in our previous study (Siestrup et al., 2022; right IFS: $x = 42, y = 8, z = 32$; left OTC: $x = -42, y = -58, z = -7$; right OTC: $x = 51, y = -52, z = -10$; left IPS: $x = -27, y = -61, z = 50$; right IPS: $x = 33, y = -67, z = 35$), which served as central points for spheres with a radius of 6 mm.

Furthermore, we investigated the influence of surface and gist modifications on activation in anterior HPC and posterior HPC. For creating hippocampal ROIs, we used the automated anatomical labeling atlas (Tzourio-Mazoyer et al., 2002) from the Wake Forest University Pickatlas toolbox (Maldjian, Laurienti, Kraft, & Burdette, 2003) in SPM12 to extract maps of the left and right HPC. To divide the HPC into anterior and posterior sections, we used an anterior–posterior border at $y = -21$ (Poppenk et al., 2013). To avoid contamination between the two sections, a 2-mm coronal gap was introduced from this border in anterior and posterior directions (Guo et al., 2020; Li, Li, Wang, Li, & Li, 2018). Contrast estimates were extracted for *surface* > *original* and *gist* > *original* from anterior HPC and posterior HPC, aggregated over both hemispheres. We conducted an additional explorative ROI analysis using a more fine-grained hippocampal parcellation. Results are reported in Appendix B.

Contrast estimates for *surface* > *original* and *gist* > *original* were subjected to one-tailed, one-sample *t* tests to check for significant activation within ROIs. Furthermore, we used two-tailed paired *t* test to compare contrast estimates from *surface* > *original* and *gist* > *original* in the functional ROIs (right IFS, OTC, IPS). For hippocampal ROIs, we conducted a 2 × 2 repeated-measures analysis of variance (rmANOVA) with the factors ROI (*anterior HPC, posterior HPC*) and Modification (*surface, gist*). Pairwise comparisons between *surface* > *original* and *gist* > *original* in hippocampal ROIs were conducted using one-tailed *t* tests, according to our hypotheses. Briefly, we expected that in anterior HPC, activation would be larger for *gist* > *original* compared with *surface* > *original*, whereas the opposite was expected for posterior HPC (e.g., Poppenk et al., 2013). *p* Values obtained from one-sample and pairwise *t* tests were Bonferroni-corrected for multiple comparisons within each ROI/analysis.

Behavioral Data Analysis

The analysis of behavioral data was conducted with RStudio (R Core Team, 2020; Version 1.3.1073).

To test the hypothesis that repeated presentations of modified videos during the fMRI decrease memory accuracy in the post-fMRI memory test, we computed the AUC (Brady et al., 2022), as well as false alarm and hit rates (Ratings 1, 2 and 3 aggregated as "rejection," and 4, 5, and 6 aggregated as "acceptance"). In addition, we considered RTs from the post-fMRI memory test, as longer RTs in memory tasks are believed to be indicative of increased retrieval difficulty because of higher cognitive processing demands (Larsen & Plunkett, 1987; Noppeney

& Price, 2004). For the analysis of AUC, hit rates, false alarm rates, and RTs for modified videos in the post-fMRI memory test ($\text{modified}_{\text{MT}}$), a 2 × 2 within-subject factorial design with the factors $\text{MODIFICATION}_{\text{fMRI}}$ (*yes, no*) and $\text{VERSION}_{\text{MT}}$ (*surface, gist*) was applied. For analyzing RTs for original videos in the post-fMRI memory test ($\text{original}_{\text{MT}}$), we applied a within-subject design with the factor $\text{VERSION}_{\text{fMRI}}$ (*original*, *surface*, *gist*).

To again confirm our classification of surface and gist modifications, we analyzed participants' story-change ratings from the rating task. Furthermore, we considered the RTs to notice changes in modified videos. The same 2 × 2 within-subject factorial design as described above was applied. In individual cases, participants' story-change ratings were inconsistent with our a priori classification. This means that participants sometimes chose higher story-change ratings (Ratings 4, 5, 6) when we had previously rated the change to be less impactful (corresponding to Ratings 1, 2, 3) or vice versa. On average, this applied to 1.70 out of 24 videos per participant. To address this individual variation, we repeated all behavioral analyses with individually rearranged surface and gist modification conditions, which consistently gave rise to the same results we report on the basis of our a priori classification.

Furthermore, we analyzed the behavioral data from the cover task in the fMRI session by calculating the error rate and mean RT according to the within-subject factor $\text{VERSION}_{\text{fMRI}}$ (*original*, *surface*, *gist*, *novel*) per participant. In only 0.77% of all question trials, no response was given so that these trials were not further considered in the analysis.

For all RT analyses, only trials with correct answers were considered. Furthermore, extreme outliers in RT (as defined as values above Quartile 3 + 3 * interquartile range or lower than Quartile 1–3 * interquartile range) were removed from each participant's data. RTs were averaged over all trials of the same factorial combination for each participant. For choosing statistical tests, data were inspected for normal distribution using the Shapiro–Wilk test and checked for extreme outliers. When data were normally distributed or could be transformed to fit normal distribution (RTs from memory test and rating task; logarithmic transformation) and included no extreme outliers, we used conventional rmANOVA. When the prerequisites for parametric analysis were not met, we applied nonparametric rmANOVA based on aligned rank-transformed data (package ARTool; Wobbrock, Findlater, Gergle, & Higgins, 2011). Post hoc pairwise comparisons were conducted with paired *t* tests or Wilcoxon signed-ranks tests (one-tailed with respect to the hypotheses, always two-tailed for explorative analysis of fMRI task).

We report mean values and standard errors of the mean. A significance level of $\alpha = .05$ was applied, and we report Bonferroni-corrected *p* values for pairwise comparisons (Bonferroni, 1936). If the assumption of sphericity was violated (Mauchly's test of sphericity), we report Greenhouse–Geisser-corrected degrees of freedom and *p* values.

RESULTS

Behavioral Pilot Study

We conducted a behavioral pilot study in an independent sample with 18 participants ($M_{\text{age}} = 23.839$ years, $SD = 4.730$, age range = 18–32 years, 15 women, three men) to validate the suitability of our stimulus material and procedure. Participants encoded episodes the same way as participants of the fMRI study did. On the next day, participants returned to complete a memory test and a rating task. Both were conducted as described for the main study.

In total, memorability of all videos was comparable, as we did not detect any extreme outliers in AUC per video ($M = .950 \pm .010$). As expected, participants showed a reduced memory performance in the surface modification condition compared with the gist modification condition ($M_{\text{surface}} = .907 \pm .013$; $M_{\text{gist}} = .995 \pm .003$; $Z = -3.88$, $p < .001$ [one-tailed]) in the memory test. Although we expected that all modifications, no matter if surface or gist modification, could be visually identified by the participants, it might take longer to become subjectively aware of surface modifications because of the smaller impact on the storyline. Indeed, participants did not miss significantly more surface than gist modifications ($M_{\text{surface}} = 0.278 \pm 0.135$, $M_{\text{gist}} = 0.0 \pm 0.0$; $Z = -1.70$, $p = .089$ [two-tailed]). Over the whole pilot study, five out of 24 surface modifications were missed only once each; gist modifications were never missed. When analyzing the time it took participants to notice the modifications in the rating task, we did not detect a significant difference between conditions ($M_{\text{surface}} = 907.682 \pm 75.694$ msec, $M_{\text{gist}} = 822.913 \pm 64.316$ msec; $t(17) = -1.44$, $p = .084$ [one-tailed]). The rating for the change of storyline was significantly higher for gist than for surface modifications ($M_{\text{surface}} = 1.304 \pm 0.059$, $M_{\text{gist}} = 4.894 \pm 0.110$; $t(17) = 38.88$, $p < .001$ [one-tailed]), as expected. Therefore, we validated that surface and gist modifications were perceived according to our a priori categorization.

fMRI Results

Behavioral Performance during fMRI Session

During the fMRI session, participants solved a cover task during which two verbs were presented after some videos and participants had to indicate which verb was more closely associated with that video. A nonparametric rmANOVA on error rates with the factor $\text{VERSION}_{\text{fMRI}}$ (*original*, *surface*, *gist*, *novel*) did not reveal significant differences, $F(3, 108) = 0.22$, $p = .883$, $\eta p^2 = .01$. For all factor levels, error rates were generally low ($M_{\text{original}} = .061 \pm .009$; $M_{\text{surface}} = .056 \pm .011$; $M_{\text{gist}} = .058 \pm .009$;

Table 1. Whole-brain Activation for Contrasts of Modified with Original Episodes at FDR $p < .05$/Uncorrected at $p < .001$ (Voxel Level)

			MNI Coordinates			
Localization	*H*	*Cluster Extent*	*x*	*y*	*z*	*t Value*
Gist > original (FDR-corrected at p < .05)						
Posterior IPS	R	37	36	−55	47	4.83
	L	24	−27	−55	44	4.02
Inferior frontal junction	L	74	−42	−1	35	5.72
	R	205	39	5	32	6.47
Inferior frontal sulcus BA 45	R	l.m.	45	29	20	6.16
	L	10	−39	29	17	3.83
Angular sulcus	R	24	42	−64	29	4.11
pSTS	R	37	48	−46	23	4.26
pMTG	R	l.m.	45	−49	11	3.75
Inferior temporal sulcus	R	107	48	−46	−10	6.27
FG	R	l.m.	36	−43	−16	5.04
MTG	R	l.m.	63	−46	−7	3.88
FG	L	76	−42	−46	−13	5.49
Original > gist (FDR-corrected at p < .05)						
SFS	R	10	27	41	41	4.21
MCC	L	22	−3	−10	38	4.29
Middle occipital gyrus	L	39	−18	−97	5	5.40
pACC	L	252	−6	35	−1	5.79
	R	l.m.	9	41	−1	4.99
Surface > original (uncorrected at p < .001)						
Posterior IPS	R	14	33	−55	47	3.81
Medial SFG (BA 8)	R	20	3	29	44	4.01
IFJ	R	20	39	5	35	4.43
Caudate nucleus	R	10	12	2	23	4.59
Anterior insula	R	19	33	26	−7	4.04
FG	L	14	−30	−49	−16	3.95
Original > surface (uncorrected at p < .001)						
Postcentral sulcus	L	30	−48	−22	38	4.60
pMTG	L	181	−54	−61	11	5.03
pSTS	L	l.m.	−51	−46	11	4.55
Middle occipital gyrus	L	l.m.	−48	−79	8	3.60
pSTS	R	34	48	−37	11	4.04

Only clusters with a minimum extent of 10 voxels are reported. H = hemisphere; MNI = Montreal Neurological Institute; L = left; R = right; BA = Brodmann's area; l.m. = local maximum.

M_{novel} = .100 ± .024). VERSION$_{fMRI}$ had a significant effect on RTs in the fMRI task, $F(3, 108) = 7.89, p < .001, \eta p^2 = .18$. Post hoc paired t tests revealed that participants took significantly longer to respond in the task after novel videos than after any other version, *original* versus *novel:* $t(36) = -4.08$, $p = .001$; *surface* versus *novel:* $t(36) = -3.35$, $p = .012$; *gist* versus *novel:* $t(36) = -3.48$, $p = .008$; $M_{original}$ = 1267.910 ± 35.721 msec; $M_{surface}$ = 1275.592 ± 32.560 msec; M_{gist} = 1277.184 ± 33.558 msec; M_{novel} = 1328.929 ± 35.486 msec (all tests two-tailed), whereas other pairwise comparisons did not reach significance, $t(36) > -0.80$, $p = 1$ (all tests two-tailed).

Whole-brain Analysis

Neural responses to modified episodic cueing. To investigate which brain regions respond to surface or gist modification in episodic cueing, we calculated the contrasts *gist* > *original* and *surface* > *original*. To better understand differences between different types of episodes, we also computed the reverse contrasts, *original* > *gist* and *original* > *surface*.

In contrast to original episodes, gist-modified episodes elicited activation in bilateral IPS, inferior frontal junction (IFJ), IFS (BA 45), and FG. In addition, we found significant activation in right angular sulcus, inferior temporal sulcus, (posterior) middle temporal gyrus (pMTG), and posterior superior temporal sulcus (pSTS; Table 1, Figure 3A). The reverse contrast (*original* > *gist*) revealed reduced activation of several brain regions in gist-modified episodes. These were bilateral pregenual ACC (pACC), right superior frontal sulcus (SFS), and left midcingulate cortex (MCC; Table 1, Figure 3C). Interestingly, this deactivation pattern was reminiscent of the one we previously reported for novel episodes compared with reactivated ones (Jainta et al., 2022; Siestrup et al., 2022).

For surface-modified episodes compared with original ones, we did not detect significant activation after correction for multiple comparisons. For completeness, we present results without correction for multiple comparisons at $p < .001$. Such subthreshold activation (uncorrected, $p < .001$) was detected in hypothesized brain regions, which were the right IPS and an inferior frontal area (IFJ), as well as left FG. In addition, activation clusters were located in the right medial superior frontal gyrus (SFG; BA 8), caudate nucleus, and anterior insula (Table 1, Figure 3B). The reverse contrast (*original* > *surface*) did not yield significant activation with correction for multiple comparisons. Subthreshold activation clusters were found in bilateral pSTS, as well as in left

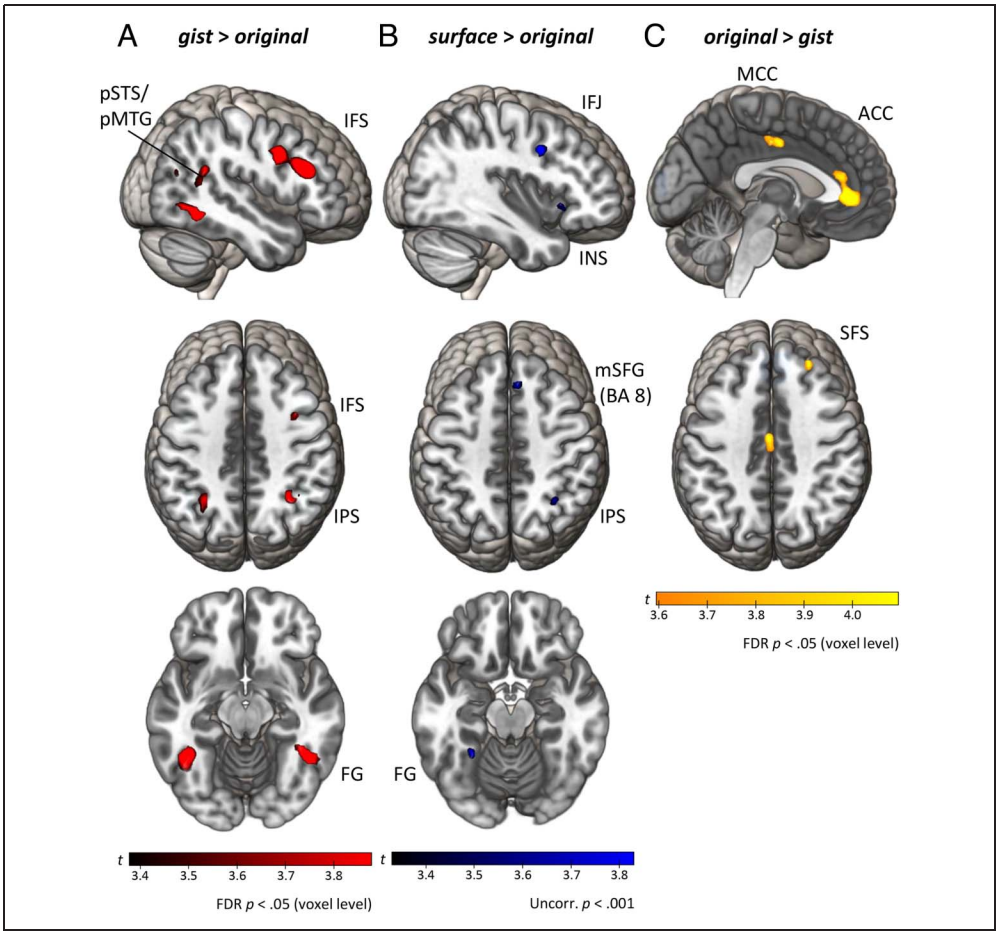

Figure 3. Whole-brain activation for episodic modifications compared with original episodes. (A) FDR-corrected ($p < .05$) t map for *gist* > *original* contrast. (B) Uncorrected ($p < .001$) t map for *surface* > *original* contrast. Please note that in this contrast, no significant activation was found with correction for multiple comparisons. For completeness and with regard to our hypotheses, we therefore show this contrast without correction at $p < .001$. (C) FDR-corrected ($p < .05$) t map for *original* > *gist* contrast. INS = insula; mSFG = medial superior frontal gyrus; BA = Brodmann's area.

postcentral sulcus, posterior MTG, and middle occipital gyrus (Table 1).

In addition, we calculated the direct contrasts between both modification types (*gist* > *surface*, *surface* > *gist*). Compared with surface modifications, gist modifications elicited significantly higher activation in bilateral IFJ/IFS (BA 44, 45), pMTG, pSTS, and FG. In addition, activity was increased in left IPS and right precuneus and postcentral sulcus (Table 2, Figure 4A). Surface modifications yielded higher brain activation than gist modifications in bilateral pACC, MCC, and left cuneus, as well as right SFS (Table 2, Figure 4B).

Neural effects of episodic reactivation. Aside from our main research question, we aimed to replicate our previous findings concerning the neural effects of episodic recall compared with novel episodes. To this end, we calculated the conjunction of the three contrasts *original* > *novel*, *surface* > *novel*, and *gist* > *novel*. Because of the very strong activation we found for this contrast, we used a more conservative threshold of $p < .001$ (FDR) to allow for a better attribution of activation to anatomical regions. In fact, we detected activation in previously identified brain regions. These were bilateral pACC, lingual gyrus, cuneus, precuneus, posterior cingulate cortex (PCC), angular gyrus (AG), middle frontal gyrus, and left insula (Jainta et al., 2022; Siestrup et al., 2022). In addition, we found activation in bilateral SFS and MTG, as well as left STS and putamen (Table 3, Figure 5).

ROI Analyses

To more specifically investigate which brain areas are in involved in the processing of surface and gist modifications, we performed ROI analyses. We used functional ROIs (right IFS, OTC, IPS) of regions that responded to

Table 2. Whole-brain Activation for Direct Contrasts of Different Episodic Modifications at FDR $p < .05$ (Voxel Level)

			MNI Coordinates			
Localization	*H*	*Cluster Extent*	*x*	*y*	*z*	*t Value*
Gist > surface (FDR-corrected at p < .05)						
Posterior IPS	L	18	−33	−58	41	3.56
Precuneus	R	56	6	−55	41	3.66
Postcentral sulcus	R	29	57	−13	35	3.95
IFJ	L	121	−39	5	29	6.30
IFS (BA 44)	L	l.m.	−42	20	26	5.39
IFJ	R	105	39	14	26	5.22
IFS (BA 45)	R	l.m.	48	29	20	4.65
pMTG	R	540	60	−61	8	7.17
pSTS	R	l.m.	48	−43	17	5.07
pMTG	L	382	−45	−64	14	5.25
pSTS	L	l.m.	−54	−49	14	4.75
FG	L	36	−42	−46	−13	5.70
	R	29	42	−49	−16	4.88
Surface > gist (FDR-corrected at p < .05)						
SFS	R	88	18	53	35	5.32
MCC	L + R	18	0	−16	32	3.80
pACC	R	584	6	35	11	8.88
	L	l.m.	−6	35	2	6.73
Cuneus/calcarine sulcus	L	12	−6	−97	5	3.76

Only clusters with a minimum extent of 10 voxels are reported. H = hemisphere; MNI = Montreal Neurological Institute; L = left; R = right; BA = Brodmann's area; l.m. = local maximum.

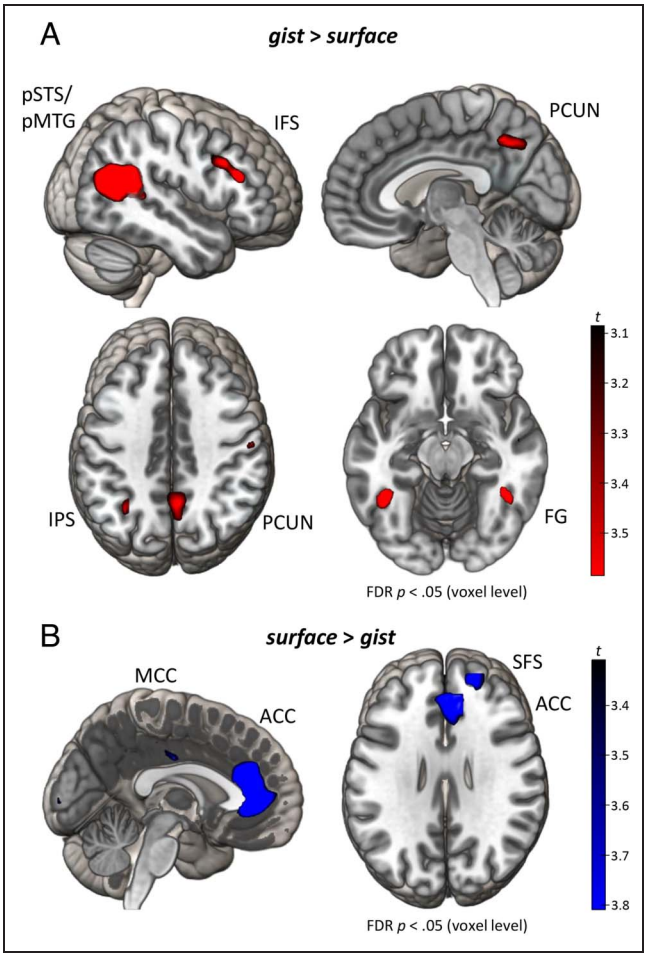

Figure 4. Whole-brain activation for direct contrasts of different episodic modifications. (A) FDR-corrected ($p < .05$) t map for *gist > surface* contrast. (B) FDR-corrected ($p < .05$) t map for *surface > gist* contrast. PCUN = precuneus.

content modification, that is, object exchanges, in our previous study (Siestrup et al., 2022) to test our hypotheses that (1) both types of modifications elicit activation in these areas and (2) that activation in these regions might be stronger for gist than for surface modifications. Further ROIs were the anterior and posterior HPC to test whether gist and surface modifications differ regarding long-axis hippocampal activation. Specifically, we expected that in anterior HPC, activation would be larger for *gist > original* compared with *surface > original*, whereas the opposite was expected for posterior HPC (e.g., Poppenk et al., 2013). Mean contrast estimates were extracted from *surface > original* and *gist > original* contrasts.

In right IFS, there was significant activation for both, surface and gist modifications in contrast to originals, surface: $M = 0.403 \pm 0.133$, $t(36) = 3.03$, $p = .005$ (one-tailed); gist: $M = 0.829 \pm 0.148$, $t(36) = 5.62$, $p < .001$ (one-tailed). A paired t test revealed that activation was greater for gist than for surface modifications, $t(36) = -3.27$, $p = .002$ (two-tailed). In the OTC ROI, gist modifications showed significant activation, $M = 0.383 \pm 0.080$, $t(36) = 4.79$, $p < .001$ (one-tailed), whereas there was a nonsignificant trend for activation in surface modifications, $M = 0.147 \pm 0.078$, $t(36) = 1.90$, $p = .066$ (one-tailed). Gist modifications elicited significantly higher activation than surface modifications, $t(36) = -2.35$, $p = .025$ (two-tailed). Likewise, gist modifications led to significant activation in the IPS ROI, $M = 0.430 \pm 0.132$, $t(36) = 3.25$, $p = .002$ (one-tailed), whereas this was not the case for surface modifications, $M = 0.138 \pm 0.111$, $t(36) = 1.25$, $p = .219$ (one-tailed). The difference between the two modification types was significant, $t(36) = -2.30$, $p = .028$ (two-tailed).

In anterior and posterior HPC, neither type of modification elicited significant activation, anterior HPC: surface: $M = 0.026 \pm 0.033$, $t(36) = 0.78$, $p = .876$ (one-tailed); gist: $M = 0.033 \pm 0.044$, $t(36) = 0.75$, $p = .915$ (one-tailed); posterior HPC: $M = 0.066 \pm 0.036$, $t(36) = 1.84$, $p = .147$ (one-tailed); $M = -0.021 \pm 0.041$, $t(36) = -0.51$, $p = 1$ (one-tailed). A rmANOVA with the factors ROI (*anterior HPC, posterior HPC*) and Modification (*surface, gist*) revealed no significant effects of ROI, $F(1, 36) = 0.07$, $p = .794$, $\eta p^2 = .00$, or Modification, $F(1, 36) = 0.93$, $p = .342$, $\eta p^2 = .03$, but a significant interaction of both factors, $F(1, 36) = 10.18$, $p = .003$, $\eta p^2 = .22$. Pairwise t tests revealed that, as expected, activation was significantly stronger for surface modifications than for gist modifications, $t(36) = 2.04$, $p = .049$ (one-tailed), in posterior HPC. However, activation for modification types did not differ in anterior HPC, $t(36) = -0.15$, $p = .884$ (one-tailed; Figure 6).

Post-fMRI Memory Test

In the post-fMRI memory test, participants were presented modified episodes from the fMRI session and, additionally, their originally encoded counterparts. Episodes that had appeared in the original version during scanning were now presented in the original and, additionally, one modified version. After each video, participants had to rate how confident they were that they had originally encoded exactly this episode.

We calculated a nonparametric rmANOVA with the factors $\text{Modification}_{\text{fMRI}}$ (*yes, no*) and $\text{Version}_{\text{MT}}$ (*surface, gist*) to analyze AUC values as a measure of memory performance in the post-fMRI memory test. There was a significant effect of $\text{Modification}_{\text{fMRI}}$, $F(1, 36) = 18.94$, $p < .001$, $\eta p^2 = .34$, as lower AUC values were achieved when participants had previously encountered modified episodes during the fMRI session ($M_{\text{yes}} = .906 \pm .010$; $M_{\text{no}} = .936 \pm .010$). In addition, there was a significant effect of $\text{Version}_{\text{MT}}$, $F(1, 36) = 239.97$, $p < .001$, $\eta p^2 = .87$, driven by higher AUC values for gist than for surface-modified versions ($M_{\text{surface}} = .853 \pm .016$; $M_{\text{gist}} = .988 \pm .004$). Furthermore, the interaction between $\text{Modification}_{\text{fMRI}}$ and $\text{Version}_{\text{MT}}$ was significant, $F(1, 36) = 12.35$, $p = .001$, $\eta p^2 = .26$. Post hoc Wilcoxon tests revealed that AUC values were only reduced significantly

Table 3. Whole-brain Activation for Episodic Recall

Localization	*H*	*Cluster Extent*	*MNI Coordinates*			*t Value*
			x	*y*	*z*	
(Original > novel) ∩ (surface > novel) ∩ (gist > novel) (FDR-corrected at p < .001)						
SFG	L	10	−9	26	59	4.23
	L	41	−9	44	44	4.60
Middle frontal gyrus	L	38	−36	23	50	5.09
	R	38	36	23	41	5.23
AG	L	160	−51	−64	44	6.89
	R	114	51	−64	44	6.83
pACC	L	1379	−6	50	2	7.00
	R	l.m.	9	50	−1	6.21
SFS	R	l.m.	21	47	32	6.60
Lingual gyrus	R (+ L)	3210	9	−82	−1	12.21
Cuneus	L + R	l.m.	0	−88	20	11.10
Calcarine sulcus, extending into posterior hippocampus (L)	L (+ R)	l.m.	−6	−79	11	10.97
Precuneus, extending into PCC	L + R	l.m.	0	−76	38	9.95
Superior temporal sulcus	L	72	−66	−31	−4	5.55
MTG	L	l.m.	−63	−28	−13	4.33
	R	44	63	−22	−16	5.47
Anterior superior temporal sulcus	L	26	−51	8	−25	4.51
Putamen	L	24	−24	5	−4	4.28
Insula	L	l.m.	−36	11	−7	4.12
Anterior middle frontal gyrus	R	11	45	53	−4	4.28

Please note that for a more conclusive attribution of activation to anatomical regions, we report activation at FDR $p < .001$ (voxel level). Only clusters with a minimum extent of 10 voxels are reported. H = hemisphere; MNI = Montreal Neurological Institute; L = left; R = right; BA = Brodmann's area; l.m. = local maximum.

by previous experience with modified versions for surface ($z = -2.97$, $p = .006$ [one-tailed]) but not for gist modifications ($z = -1.00$, $p = .634$ [one-tailed]; Figure 7A).

Notably, AUC scores obtained from the pilot study and the main study when only originals were considered from the fMRI session, that is, when participants did not have any previous experience with modified episodes before the

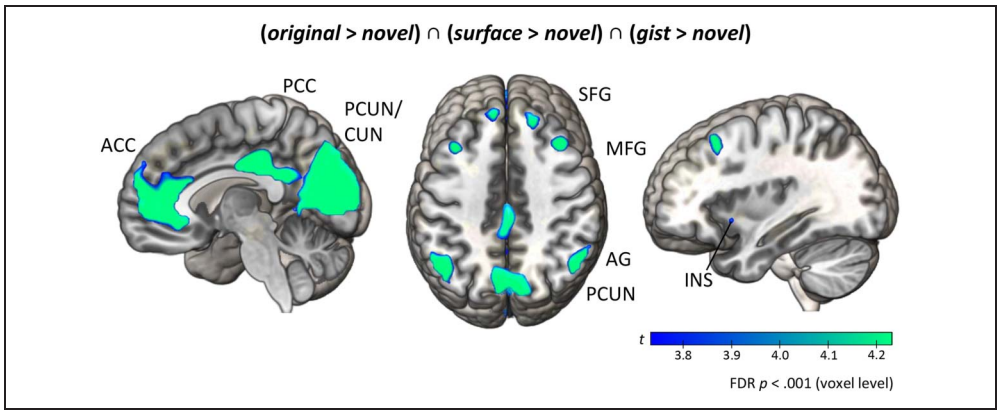

Figure 5. Whole-brain activation for episodic recall compared with novel videos. FDR-corrected *t* map for conjunction of *original* > *novel*, *surface* > *novel*, and *gist* > *novel* contrasts. For a smoother visualization of the conjunction, the *t* map was resampled to a resolution of 1 mm^3 voxels. Please note that for a more conclusive attribution of activation to anatomical regions, we report activation at FDR $p < .001$ (voxel level). PCUN = precuneus; CUN = cuneus; MFG = middle frontal gyrus; INS = insula.

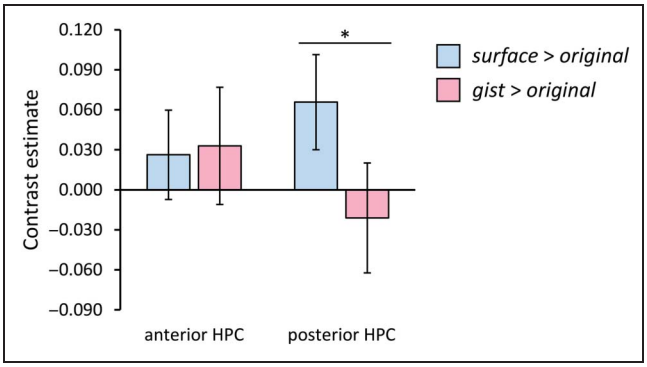

Figure 6. ROI analysis of bilateral anterior and posterior HPC. Contrast estimates were extracted from contrasts *surface* > *originals* and *gist* > *originals*. Statistics: rmANOVA, one-sample *t* tests and paired *t* tests (one-tailed). *$p < .05$.

memory test, were highly similar (pilot study: $M_{surface} = .907 \pm .013$; $M_{gist} = .995 \pm .003$; main study: $M_{no\text{-}surface} = .882 \pm .019$; $M_{no\text{-}gist} = .990 \pm .006$).

To more specifically investigate the influence of modified episodic cueing on memory, we additionally analyzed the false alarm and hit rates. Using a nonparametric rmANOVA, we found a significant effect of the factor $\text{Modification}_{fMRI}$ on false alarm rates, $F(1, 36) = 41.59$, $p < .001$, $\eta p^2 = .54$, which were higher when modified episodes had been experienced already during the fMRI session ($M_{yes} = .188 \pm .017$; $M_{no} = .118 \pm .018$). In addition, there was a significant influence of Version_{MT}, $F(1, 36) = 473.63$, $p < .001$, $\eta p^2 = .93$, as participants had a higher tendency to falsely accept surface-modified versions compared with gist-modified versions as originals ($M_{surface} = .291 \pm .028$; $M_{gist} = .015 \pm .006$). In addition, there was a significant interaction between both factors, $F(1, 36) = 18.16$, $p < .001$, $\eta p^2 = .34$. Post hoc Wilcoxon tests showed that there was only an increase in false alarms for surface modifications ($z = -3.29$, $p = .003$ [one-tailed]), but not for gist modifications ($z = -1.47$, $p = .284$ [one-tailed]) when participants had previously seen modified versions during scanning (Figure 7B). Hit rates were close to ceiling for all experimental conditions ($M_{yes\text{-}surface} = .929 \pm .020$; $M_{yes\text{-}gist} = .949 \pm .014$; $M_{no\text{-}surface} = .932 \pm .023$; $M_{no\text{-}gist} = .966 \pm .014$) and were neither significantly affected by the factor $\text{Modification}_{fMRI}$, $F(1, 36) = 1.84$, $p = .184$, $\eta p^2 = .05$; Version_{MT}, $F(1, 36) = 0.02$, $p = .899$, $\eta p^2 = .00$; nor an interaction of both, $F(1, 36) = 0.23$, $p = .631$, $\eta p^2 = .01$ (Figure 7C).

For modified_{MT} videos, there was a significant effect of $\text{Modification}_{fMRI}$ on RTs, $F(1, 36) = 4.43$, $p = .042$, $\eta p^2 = .11$, as participants responded faster when they had not experienced the modification during the fMRI session ($M_{yes} = 471.383 \pm 23.250$ msec; $M_{no} = 439.958 \pm 20.392$ msec). In addition, there was a nonsignificant trend for an effect of Version_{MT}, $F(1, 36) = 3.57$, $p = .067$, $\eta p^2 = .09$, driven by the participants' tendency to respond faster for gist than for surface-modified videos ($M_{surface} = 475.620 \pm 25.157$ msec; $M_{gist} = 435.721 \pm 17.603$ msec). There was no significant interaction, $F(1, 36) = 2.00$, $p = .166$, $\eta p^2 = .05$ (Figure 7D). Concerning RTs for original_{MT} videos, we found no significant effect of the factor Version_{fMRI}, $F(1.59, 57.23) = 0.25$, $p = .727$, $\eta p^2 = .01$ (Figure 7E).

For novel videos, we found very low false alarm rates ($M = .020 \pm .017$) and a mean RT of 527.041 ± 27.941 msec in the post-fMRI memory test.

Rating Task

In the rating task, participants had to rate how much the storyline of modified episodes deviated from the originally encoded version (from 1 = *0% different* to 6 = *100% different*). A nonparametric rmANOVA with the factors $\text{Modification}_{fMRI}$ (*yes, no*) and Version_{MT} (*surface, gist*) confirmed a significant influence of Version_{MT} on story-change ratings in the post-fMRI rating task, $F(1, 36) = 347.29$, $p < .001$, $\eta p^2 = .91$. As expected, surface-modified versions received lower ratings than gist-modified versions ($M_{surface} = 1.363 \pm 0.058$; $M_{gist} = 4.897 \pm 0.093$), once again confirming our classification of modifications. Interestingly, there was also a significant effect of $\text{Modification}_{fMRI}$, $F(1, 36) = 4.87$, $p = .034$, $\eta p^2 = .12$, as modified videos that had already been presented during the fMRI session received lower story-change ratings than those that had been presented in the original version before ($M_{yes} = 3.083 \pm 0.066$; $M_{no} = 3.177 \pm 0.069$). There was no significant interaction of both factors, $F(1, 36) = 1.791$, $p = .189$, $\eta p^2 = .05$.

A rmANOVA on RTs in the rating task, that is, how long it took participants to notice the change in modified videos, revealed a significant effect of Version_{MT}, $F(1, 36) = 9.72$, $p = .004$, $\eta p^2 = .21$. Participants took longer to notice the modification in surface-modified versions than in gist-modified versions ($M_{surface} = 882.181 \pm 42.644$ msec; $M_{gist} = 798.305 \pm 32.695$ msec). Furthermore, there was a nonsignificant trend for an effect of $\text{Modification}_{fMRI}$, $F(1, 36) = 3.94$, $p = .055$, $\eta p^2 = .10$, as participants tended to notice the modification sooner when they had already seen the same modified version during the fMRI session ($M_{yes} = 793.114 \pm 26.710$ msec; $M_{no} = 887.373 \pm 54.061$ msec). No significant interaction was found, $F(1, 36) = 0.88$, $p = .356$, $\eta p^2 = .02$.

DISCUSSION

In the present study, we examined the impact of PEs on episodic memories. To this end, we used modified cues to induce PEs of different episodic relevance, altering the episodic gist or only surface properties. Surface modifications led to significant activation in right IFS. In contrast, modifications that also changed the gist of the episode led to overall more pronounced activation in the same area and additional activity in OTC, IPS, and pMTG. Posterior HPC responded to surface changes more strongly than to gist changes. Interestingly, there was a

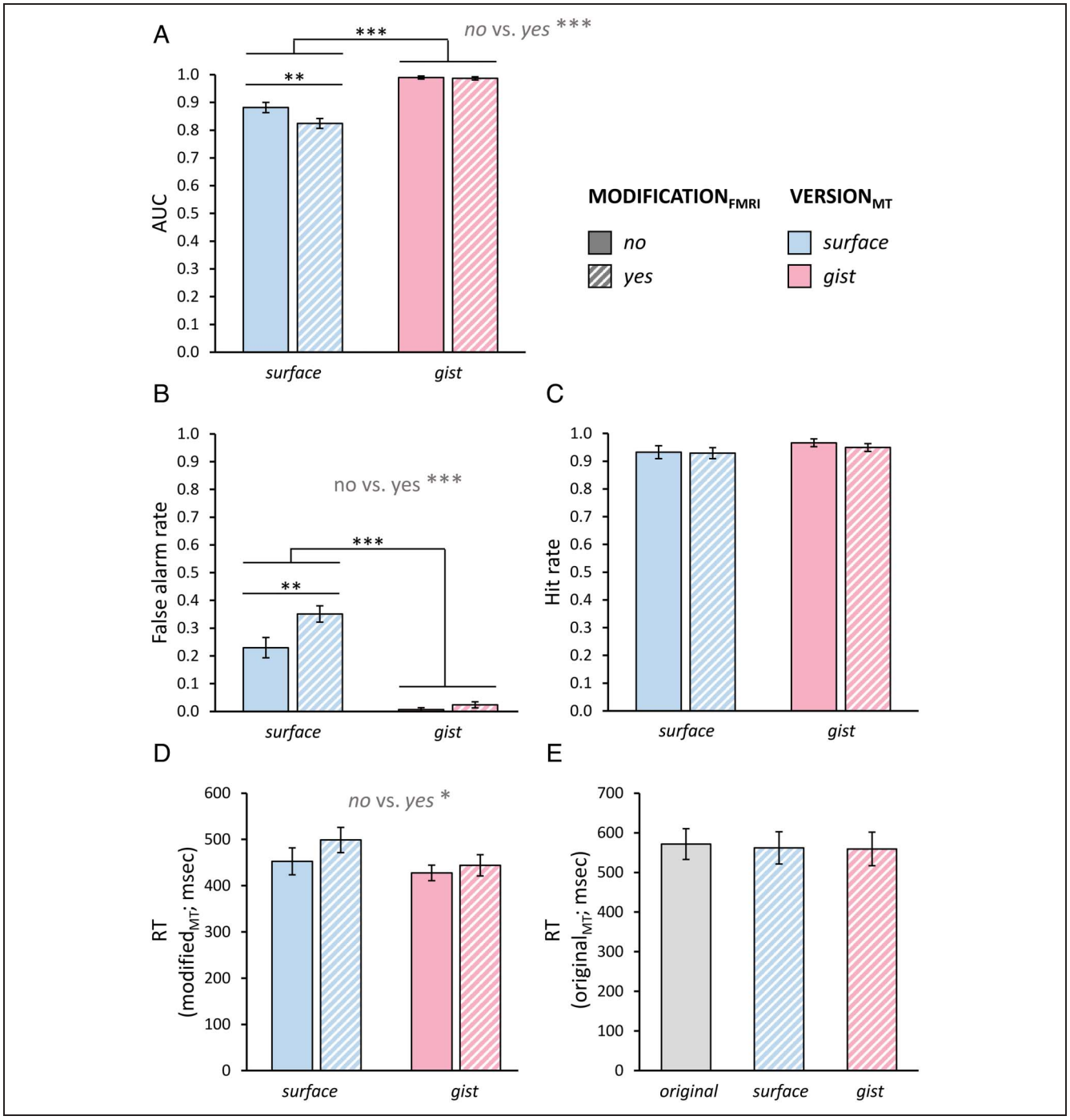

Figure 7. Behavioral results from post-fMRI memory test. Participants rated modified$_{MT}$ and original$_{MT}$ videos to decide whether they showed originally encoded episodes or not. (A) AUC. Statistics: nonparametric rmANOVA and Wilcoxon signed-ranks tests (one-tailed). (B) False alarm rates for modified$_{MT}$ videos. Statistics: nonparametric rmANOVA and Wilcoxon signed-ranks tests (one-tailed). (C) Hit rates for original$_{MT}$ videos. Statistics: nonparametric rmANOVA. (D) RTs for modified$_{MT}$ videos. Statistics: rmANOVA. (E) RTs for original$_{MT}$ videos. Statistics: rmANOVA. Bar plots show means and standard errors. *$p < .05$, **$p < .01$, ***$p < .001$.

significant reduction of cingulate activity for gist modifications as compared with original or surface-modified cueing. The day after the fMRI session, we found that a history with surface-modified, but not with gist-modified episodic cues increased the erroneous acceptance of modified stories as originals. Our findings demonstrate that the relevance of mnemonic PEs for the storyline of episodes is crucial for brain responses and influences on memory.

Neural Responses to Episodic Modifications

We expected that both types of episodic modifications, characterized by the exchange of only one object, elicit

brain activation in areas that we had previously identified for content-based PEs with intermediate episodic relevance (Siestrup et al., 2022). Indeed, gist modifications elicited brain activation in those hypothesized regions, namely, OTC, including FG, posterior IPS, and IFS. This was revealed by whole-brain and ROI analyses. For surface modifications, we found subthreshold whole-brain activation in right IFS, left FG, and right posterior IPS. The ROI analysis revealed significant activation in the right IFS, as well as a trend for activation in OTC for surface modifications. As previously discussed (Siestrup et al., 2022), these brain areas are involved in the processing of object information in actions (El-Sourani et al., 2019; Lingnau & Downing, 2015; Wiggett & Downing, 2011; Reber et al., 2005; Grill-Spector et al., 2001). Furthermore, lateral OTC, which was primarily activated by gist modifications, codes for conceptual action representations (Wurm & Caramazza, 2022). Activation in right IFS, OTC, and IPS ROIs was significantly greater for gist than for surface modifications, as expected because of the gist modifications' comparably high relevance for episode content. In summary, we confirmed our general hypotheses concerning activation typical for PEs elicited by content modifications in an episodic cueing setting.

In addition, we found subthreshold activation clusters for surface modifications in the caudate nucleus, anterior insula, and medial BA 8. These three areas, along with TPJ and IPS, were found to be co-activated when predictions about complex sequences of movements were violated (Schiffer & Schubotz, 2011). Although the activations found in the present study did not remain significant after correction for multiple comparisons, the similarity to this previously identified network is striking. Caudate nucleus is known to respond to breaches of expectation (Schiffer et al., 2012; Schiffer & Schubotz, 2011; Delgado, Li, Schiller, & Phelps, 2008; Haruno & Kawato, 2006), whereas BA 8 and anterior insula activations were shown to reflect unexpectedness and uncertainty (Schiffer, Krause, & Schubotz, 2013; Sarinopoulos et al., 2010; Zaretsky, Mendelsohn, Mintz, & Hendler, 2010; Volz, Schubotz, & Von Cramon, 2003, 2005). In summary, the high degree of similarity between surface-modified episodes and originals might have contributed to increased uncertainty, whereas gist-modified episodes were more clearly distinguished from original ones.

As hypothesized, we found significant qualitative differences in brain activation between gist and surface modifications. For gist compared with surface modifications (and original episodes), there was increased activation in pSTS and pMTG. This finding is especially interesting as we have previously observed these areas to be activated specifically for structural episodic modifications, that is, the exchange of adjacent action steps, but not for the exchange of an object with intermediate episodic relevance (Siestrup et al., 2022), which will be further discussed below. Posterior temporal cortex, as part of the OTC (Wurm & Caramazza, 2022), is crucial for conceptual processing (Martin, 2007) and is thought to represent conceptual object properties (Fairhall & Caramazza, 2013). In the context of such concepts, especially left pMTG is suggested to be important for retrieval of semantic information and the comprehension and recognition of actions (Davey et al., 2016), together with other network components that were identified for gist modifications, namely, STS, IPS, and lateral OTC (James, VanDerKlok, Stevenson, & James, 2011). Davey and colleagues (2016) found that pMTG was functionally connected with IFG and IPS, and proposed that pMTG, together with these regions, is involved in adjusting semantic retrieval to different contexts. Interestingly, pMTG showed increased activation when goal-based expectations were violated (Jastorff, Clavagnier, Gergely, & Orban, 2011), which can also be transferred to prediction violations in the present study. Furthermore, pMTG in conjunction with IFG has been suggested to have a role in the resolution of interfering episodic representations via semantic elaboration (Han, O'Connor, Eslick, & Dobbins, 2012). In that sense, pMTG could subserve the integration of semantic and episodic information, which might be particularly required in the case of gist modifications.

Furthermore, gist modifications yielded higher brain activation in precuneus compared with surface modifications. This area is involved in episodic memory retrieval (Cavanna & Trimble, 2006). Increased activity for gist modifications here might reflect a strong dissociation between the self-experienced episode and the presented version of the episode. Interestingly, we previously identified this area for structure but not for content modifications with intermediate episodic relevance (Siestrup et al., 2022). Aside from its role in the sequential organization of memories, precuneus has been shown to be involved in their contextual reconstruction (Foudil, Kwok, & Macaluso, 2020). Therefore, activation in this area for gist modifications is a hint that the exchange of a gist-relevant object seems to be meaningful not only for episode content but also for structural and contextual (re)evaluations of that memory, specifically with reference to our previous work (Siestrup et al., 2022).

In summary, we saw that gist modifications elicited activation we identified to be specific for content, but also structure-based PEs in episodic memory in our previous work (Siestrup et al., 2022). Schiffer, Ahlheim, and colleagues (2013) showed that adaptation of internal models following PEs occurred in exactly those areas that responded to gist modifications in the present study. In their experiment, episodes were modified by introducing a breach-point from which the action unfolded with new content and structural aspects. Thus, brain responses to Schiffer's gist modifications reflected more than mere object information processing, as the modified aspect was relevant for the integrity of an episode as a whole. Similarly, in the present study, integration of new episodic gist likely requires revision of object information and structure and context of events within an episode as the new storyline unfolds.

We had hypothesized to find increased ACC activation for gist modifications because it has been reported that more gist-like memories depend strongly on ACC (e.g., Sekeres, Winocur, Moscovitch, Anderson et al., 2018). Surprisingly, we observed exactly the opposite: Compared with original and surface-modified episodes, pACC activation, as well as MCC activation, was reduced in gist-modified episodes. Bonasia and colleagues (2018) found a similar response pattern in their study. Participants encoded episodes with a gist that was either congruent or incongruent with previous experiences. With increasing congruence, enhanced encoding-related activation in pACC, PCC, superior temporal gyrus, and AG was observed. This is consistent with our findings for gist-modified episodes and also for novel episodes we presented. For the latter, activation was decreased compared with previously experienced episodes in all four areas mentioned by the researchers. The authors argued that ACC and PCC were likely engaged because of their role during the activation of prior knowledge and semantic processing (Bonasia et al., 2018), thus linking new experience to prior ones. It has been suggested that ACC is responsible for gist processing by providing a template that allows for prediction and interpretation of events (Sekeres, Winocur, & Moscovitch, 2018). In the case of gist modifications in the present study, the gist of the encoded episode was no longer intact, so predictions could potentially no longer rely on the established template. In addition, it is known that ACC is involved in conflict processing (Vassena, Holroyd, & Alexander, 2017; Botvinick, Cohen, & Carter, 2004). In gist-modified episodes, conflict arising because of the mismatch with encoded episodes might be rapidly resolved, as the new gist defines a new episode. This also matches our behavioral findings in the post-fMRI memory test as well as the fact that brain responses to gist modifications strongly decreased over repeated encounters with the same modified episode. Similarly, Webb, Turney, and Dennis (2016) reported that retrieval of non-gist details was characterized by higher (dorsal) ACC activation than retrieval of gist details in a recognition memory task and argued that this finding was because of elevated cognitive efforts and conflict monitoring processes, as mediated by this brain area. In our previous study, pACC activation in response to episodic modifications was increased in the condition that was characterized by strong memory modification effects in the post-fMRI memory test and pACC activation increased over time for later false memories (Siestrup et al., 2023). In the present study, pACC activation was reduced in gist-modified episodes, which did not trigger memory modification in the post-fMRI memory test. Therefore, we suggest that pACC might be involved in memory modification in response to mnemonic PEs. In summary, the relevance of PE within the episode is probably responsible for the impact on memory: When the overall storyline remains intact, a memory might be more likely to be (slightly) modified through PE, whereas a new, distinct episode is established when PE is gist relevant. This idea is similar to a computational approach of memory modification formulated by Gershman, Monfils, Norman, and Niv (2017). They proposed that memories are modified when old and new information are inferred to share the same latent cause, whereas new memories are formed when sensory input is ascribed to a new latent cause.

Surface modifications elicited higher levels of activation in posterior HPC than gist modifications. This is consistent with the suggestion that detailed memory representations depend more on the posterior HPC than general ones do (Sekeres, Winocur, & Moscovitch, 2018; Sekeres, Winocur, Moscovitch, Anderson et al., 2018; Moscovitch et al., 2016; Poppenk et al., 2013). Accordingly, surface modifications likely elicited increased processing of more fine-grained information within modified episodes (Robin & Moscovitch, 2017; Moscovitch et al., 2016). Posterior HPC activation has been linked to episodic retrieval (e.g., Poppenk et al., 2013), which lends further support to our suggestion that surface modifications are processed with close reference to existing memories, whereas gist modifications are rather interpreted as new episodes because of the clear drift in content compared with the original. However, we could not confirm that gist modifications, compared with surface modifications, lead to increased activation in anterior HPC, neither on the whole brain level nor on the basis of ROI analyses. It can be speculated that this null finding might be attributed to the signal-to-noise ratio in this region (Brunec et al., 2018), which might have complicated the detection of condition differences.

Despite the significant interaction between HPC ROI and type of modification, the overall BOLD response in the HPC to episodic PE was unexpectedly low (did not significantly differ from random). This is surprising, because several previous studies showed that HPC responds to mnemonic PEs (e.g., Bein, Duncan, & Davachi, 2020; Long, Lee, & Kuhl, 2016; Duncan, Ketz, Inati, & Davachi, 2012). It is not clear why we could not replicate these findings, but it is possible that univariate analysis approaches cannot fully capture the complexity of hippocampal responses to mnemonic PEs (e.g., Sinclair, Manalili, Brunec, Adcock, & Barense, 2021) or that hippocampal mismatch responses were strongest on the first trial with a modified cue and decreased afterward. Furthermore, it has been shown that the magnitude of hippocampal responses to PE is positively related to the number of changes in episodes (Duncan et al., 2012). Our findings provide the first evidence that the HPC is sensitive to not only quantitative differences in PE strength but also qualitative ones, as implemented through surface and gist modifications. Furthermore, future work could focus on the influence of surface and gist modification on representations in posterior and anterior HPC, respectively, to shed more light on processes of pattern separation and integration in episodic memory (Brunec, Robin, Olsen, Moscovitch, & Barense, 2020).

Influence of Episodic Modifications on Memory Accuracy in Post-fMRI Memory Test

We replicated our finding that the repeated experience of modified episodes during the fMRI session influences memory performance in a post-fMRI memory test. Importantly, in the current study, we found this effect only for surface- but not for gist-modified episodes. Specifically, previous encounters with surface-modified videos increased the participants' tendency to erroneously accept modified versions as truly encoded, as reflected by reduced AUC values and increased false alarm rates. Hit rates were not influenced and remained at a high level even after the repeated exposure to modified episodes, similar to what we showed before (Siestrup et al., 2023, 2022; Jainta et al., 2022). This finding is in line with reports that the inclusion of new information into an old memory is not necessarily accompanied by a deterioration of the original memory (Sinclair & Barense, 2018; St. Jacques, Olm, & Schacter, 2013). Together, the pattern of results in the post-fMRI memory test suggests that memory modification occurred because of the encoding of an additional memory trace that interfered with the original one, but not a replacement of original memory content (Siestrup et al., 2023). We found the memory modification effect only for surface but not for gist modifications, which was expected because of their differential relevance for the episodes. In line with that, participants were also faster to detect gist than surface modification in the rating task, that is, when explicitly asked to rate how much the storyline of modified episodes deviated from the originally encoded version. It has previously been demonstrated that the experience of episode modifications can also improve memory, but only when such modifications are detected and remembered by participants (Wahlheim & Zacks, 2019). Furthermore, it is known that distinct items, that is, those that violate the current context, are remembered better and that such distinctiveness reduced false memories (Sommer & Sander, 2022; Hunt, 2013). The fact that gist modifications effectively constituted new episodes likely contributed to distinctiveness between original and modified versions, explaining the lack of confusion between them and, ultimately, the low rate of false alarms. This distinctiveness is probably also relevant for another finding from the post-fMRI memory test: Memory performance was generally worse for episodes with a surface modification compared with those with a gist modification, irrespective of any previous experience with modified videos during scanning. This was evidenced by reduced AUC values and increased false alarm rates in the surface modification condition. Therefore, it was more difficult for participants to correctly differentiate modified from original versions when the modification was not gist relevant. This observation is in agreement with reports that noncentral details are less likely to be remembered (Sekeres et al., 2016) and can be easily confused by misinformation (Reyna et al., 2016; Reyna & Lloyd, 1997).

In addition, when modified episodes had been presented in the scanner, participants were slower to correctly judge the same as not part of the original episode repertoire in the post-fMRI memory test. Longer RTs in memory tasks can be indicative of higher cognitive processing demands during retrieval (Noppeney & Price, 2004; Larsen & Plunkett, 1987). Potentially, it became more difficult for participants to correctly reject modified versions because of the encoding of alternative episode versions.

In a previous study, we found an increase in activity in the episodic memory network for episodes that later lead to false memories (Siestrup et al., 2023). In the present study, it was not possibly to model brain responses as a function of subsequent behavioral outcomes in the post-fMRI test. This was because the total number of false alarms was too low, and, as expected, they were correlated with the surface type of cue manipulation. Thus, low statistical power and expected confound prevented us from testing a replication of our earlier findings.

Overall, behavioral evidence we collected so far suggests that based on mnemonic PEs, alternative episode representations are encoded, which then stand in conflict to the originally encoded ones. In the case of surface-level PEs, these alternative representations interfered, leading to false memories, whereas in the case of gist level PEs, the two alternatives were clearly separated.

Conclusion

Our results indicate that the relevance of mnemonic PEs within episodes shapes neural responses and memory performance. Although both gist and surface modifications of episodic cues were characterized by the exchange of only one object within an episode, their implications were highly specific: Surface-level PE increased uncertainty and conflict between alternative episode representations, whereas gist-level PE induced the encoding of a new, distinct episode not conflicting with the original one. An important implication of our findings is that not all types of mnemonic PE influence memory in the same way, which needs to be considered in future research.

APPENDIX A

To investigate whether brain responses to episodic modification attenuate with repeated presentation of the same modified episodes, we additionally investigated the parametric contrasts for surface and gist modifications. For gist modifications, we found significant attenuation in several areas that responded to modifications (Table A1). These included bilateral IPS, IFS/IFJ, FG, and inferior temporal regions. Furthermore, attenuation was found in bilateral parahippocampal gyrus (PHG), retrosplenial cortex, and right IFG (Table A1). For surface modifications, we did not detect significant attenuation effects with correction for multiple comparisons. For completeness, we

Table A1. Whole-brain Activation for Parametric Decrease of Activation in Response to Episodic Modifications at FDR $p < .05$/ Uncorrected at $p < .001$ (Voxel Level)

			MNI Coordinates			
Localization	*H*	*Cluster extent*	*x*	*y*	*z*	*t Value*
Gist, parametric decrease (FDR-corrected at p < .05)						
Posterior intraparietal sulcus	R	105	33	−52	53	4.81
Superior precentral sulcus	L	229	−39	−1	38	5.06
Inferior frontal sulcus	L	l.m.	−48	23	26	4.04
Inferior frontal junction	R	262	39	14	29	5.23
Superior precentral sulcus	R	l.m.	39	2	41	4.48
Retrosplenial cortex	R	118	21	−55	23	4.67
	L	l.m.	−6	−55	14	3.20
Inferior frontal gyrus (BA 47)	R	21	36	32	−10	4.16
Fusiform gyrus, extending into parahippocampal gyrus	L	874	−33	−52	−13	5.84
Inferior temporal gyrus	L	l.m.	−36	−37	−16	5.38
Middle occipital gyrus, extending into posterior intraparietal sulcus	L	l.m.	−33	−73	35	5.32
AG	L	l.m.	−45	−67	23	4.45
Fusiform gyrus, extending into parahippocampal gyrus	R	472	36	−40	−16	6.05
Inferior temporal gyrus	R	l.m.	42	−61	−7	5.28
AG	R	l.m.	42	−64	17	4.91
Middle occipital gyrus, extending into posterior intraparietal sulcus	R	l.m.	36	−70	26	4.63
Surface, parametric decrease (uncorrected at p < .001)						
Posterior intraparietal sulcus	R	17	30	−55	44	3.89
	R	10	30	−67	38	3.74
Retrosplenial cortex	R	34	15	−55	26	5.30
Fusiform gyrus	R	52	36	−40	−16	5.02
Parahippocampal gyrus	R	l.m.	27	−34	−19	4.04
Inferior temporal gyrus	R	l.m.	39	−58	−7	3.39
Fusiform gyrus, extending into parahippocampal gyrus	L	64	−36	−40	−16	4.85

Only clusters with a minimum extent of 10 voxels are reported. H = hemisphere; MNI = Montreal Neurological Institute; L = left; R = right; BA = Brodmann's area; l.m. = local maximum.

investigated the parametric contrast at a subthreshold level (uncorrected, $p < .001$) and found attenuation in the right IPS and retrosplenial cortex and bilateral FG and PHG (Table A1).

APPENDIX B

In addition to our ROI analysis in anterior and posterior HPC, we conducted an explorative analysis using a more fine-grained parcellation of HPC. To this end, anatomical ROIs were created from probabilistic maps from the Julich-Brain Cytoarchitectonic Atlas (Amunts, Mohlberg, Bludau, & Zilles, 2020). We used maps of hippocampal subfields CA1 (Palomero-Gallagher, Kedo, Mohlberg, Zilles, & Amunts, 2020a, 2020b), CA2 (Palomero-Gallagher et al., 2020b; Palomero-Gallagher, Kedo, Mohlberg, Zilles, & Amunts, 2020c), and CA3 (Palomero-Gallagher et al., 2020b; Palomero-Gallagher, Kedo, Mohlberg, Zilles, & Amunts, 2020d). A threshold of 0.3 was implemented in ImCalc in SPM12, and maps were divided into anterior and posterior portions as described above. Contrast estimates were extracted as described above from *surface* > *original* and *gist* > *original* contrasts. A nonparametric rmANOVA (chosen because of missing normality and

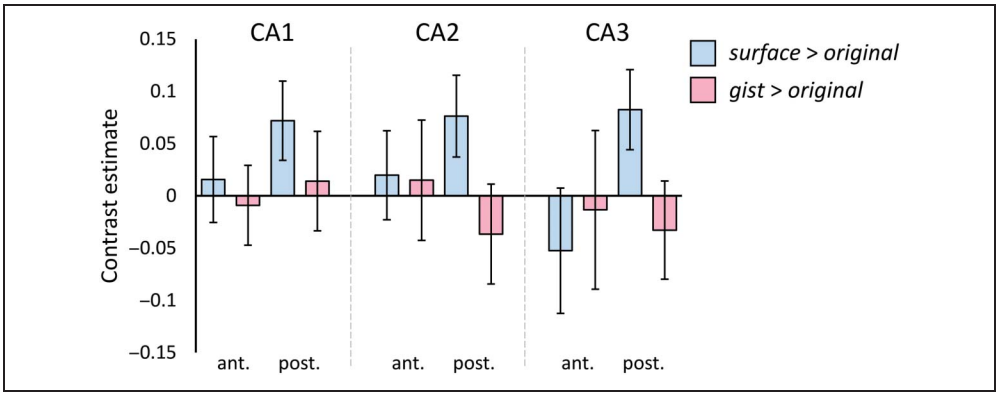

Figure B1. Explorative ROI analysis in hippocampal subfields CA1, CA2, and CA3. ROIs were created from probabilistic maps from the Julich-Brain Cytoarchitectonic Atlas (Amunts et al., 2020) and thresholded at 0.3. Contrast estimates were extracted for *surface > original* and *gist > original* contrasts. Statistics: nonparametric rmANOVA and Wilcoxon signed-ranks tests (one-tailed).

extreme outliers) with the factors Modification (*surface, gist*), ROI (*CA1*, *CA2*, *CA3*), and Location (*anterior*, *posterior*) revealed no significant main effects, Modification: $F(1, 36) = 2.04$, $p = .162$, $\eta p^2 = .05$; ROI: $F(2, 72) = 1.04$, $p = .357$, $\eta p^2 = .03$; Location: $F(1, 36) = 1.82$, $p = .186$, $\eta p^2 = .05$. There was a nonsignificant trend for an interaction of Modification and ROI, $F(2, 72) = 2.48$, $p = .091$, $\eta p^2 = .06$. Furthermore, we found a significant interaction of Modification and Location, $F(1, 36) = 4.68$, $p = .037$, $\eta p^2 = .12$. Paired Wilcoxon tests according to our hypotheses revealed that in posterior locations, surface modifications elicited significantly more activation than gist modifications ($z = -2.56$, $p = .02$ [one-tailed]; $M_{\text{surface-posterior}} = 0.077 \pm 0.035$; $M_{\text{gist-posterior}} = -0.018 \pm 0.045$). There was no difference between modification types in anterior locations ($z = -0.53$, $p = 1$ [one-tailed]; $M_{\text{surface-anterior}} = -0.006 \pm 0.043$; $M_{\text{gist-anterior}} = -0.002 \pm 0.055$). There were no further interaction effects (Figure B1). Please note that no further HPC subfields (dentate gyrus, subiculum) were analyzed because of signal loss in the anterior portion of these ROIs.

Acknowledgments

The authors thank Monika Mertens, Alea Bexten, and Lana Steuernagel for their help during data collection. We thank Benjamin Jainta for helpful comments on an earlier draft of this article and members of Research Unit FOR 2812 for valuable discussions.

Reprint requests should be sent to Sophie Siestrup, University of Münster, Fliednerstraße 21, 48149 Münster, Germany, or via e-mail: s.siestrup@uni-muenster.de.

Data Availability Statement

The data sets generated for this study are available on request to the corresponding author.

Author Contributions

Sophie Siestrup: Conceptualization; Formal analysis; Investigation; Methodology; Visualization; Writing—Original draft; Writing—Review & editing. Ricarda I. Schubotz: Conceptualization; Funding acquisition; Methodology; Resources; Supervision; Writing—Original draft; Writing—Review & editing.

Funding Information

This work was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft; https://dx.doi.org/10.13039/501100001659), grant number: 419037023. The funder had no role in study design, data collection, analysis and interpretation, decision to publish, or writing of the report.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience* (*JoCN*) during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance.

REFERENCES

Amunts, K., Mohlberg, H., Bludau, S., & Zilles, K. (2020). Julich-brain: A 3D probabilistic atlas of the human brain's cytoarchitecture. *Science*, *36*, 988–992. https://doi.org/10.1126/science.abb4588, PubMed: 32732281

Antony, J. W., Ferreira, C. S., Norman, K. A., & Wimber, M. (2017). Retrieval as a fast route to memory consolidation. *Trends in Cognitive Sciences*, *21*, 573–576. https://doi.org/10.1016/j.tics.2017.05.001, PubMed: 28583416

Bein, O., Duncan, K., & Davachi, L. (2020). Mnemonic prediction errors bias hippocampal states. *Nature Communications*, *11*, 3451. https://doi.org/10.1038/s41467-020-17287-1, PubMed: 32651370

Bonasia, K., Sekeres, M. J., Gilboa, A., Grady, C. L., Winocur, G., & Moscovitch, M. (2018). Prior knowledge modulates the neural substrates of encoding and retrieving naturalistic events at short and long delays. *Neurobiology of Learning and Memory*, *153*, 26–39. https://doi.org/10.1016/j.nlm.2018.02.017, PubMed: 29474955

Bonferroni, C. E. (1936). *Teoria statistica delle classi e calcolo delle probabilità* (Vol. 8, pp. 3–62). Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, *8*, 539–546. https://doi.org/10.1016/j.tics.2004.10.003, PubMed: 15556023

Brady, T. F., Robinson, M. M., Williams, J. R., & Wixted, J. T. (2022). Measuring memory is harder than you think: How to avoid problematic measurement practices in memory research. *Psychonomic Bulletin and Review*, *30*, 421–449. https://doi.org/10.3758/s13423-022-02179-w, PubMed: 36260270

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, *11*, 164–169. https://doi.org/10.1111/1467-8721.00192

Brett, M., Anton, J. L., Valabregue, R., & Poline, J.-B. (2002). Region of interest analysis using an SPM toolbox. *Neuroimage*, *16*, 769–1198. https://doi.org/10.1201/b14650-28

Brunec, I. K., Bellana, B., Ozubko, J. D., Man, V., Robin, J., Liu, Z. X., et al. (2018). Multiple scales of representation along the hippocampal anteroposterior axis in humans. *Current Biology*, *28*, 2129–2135. https://doi.org/10.1016/j.cub.2018.05.016, PubMed: 29937352

Brunec, I. K., Robin, J., Olsen, R. K., Moscovitch, M., & Barense, M. D. (2020). Integration and differentiation of hippocampal memory traces. *Neuroscience and Biobehavioral Reviews*, *118*, 196–208. https://doi.org/10.1016/j.neubiorev.2020.07.024, PubMed: 32712280

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain*, *129*, 564–583. https://doi.org/10.1093/brain/awl004, PubMed: 16399806

Cheng, S., Werning, M., & Suddendorf, T. (2016). Dissociating memory traces and scenario construction in mental time travel. *Neuroscience and Biobehavioral Reviews*, *60*, 82–89. https://doi.org/10.1016/j.neubiorev.2015.11.011, PubMed: 26627866

Corbin, J. C., Reyna, V. F., Weldon, R. B., & Brainerd, C. J. (2015). How reasoning, judgment, and decision making are colored by gist-based intuition: A fuzzy-trace theory approach. *Journal of Applied Research in Memory and Cognition*, *4*, 344–355. https://doi.org/10.1016/j.jarmac.2015.09.001, PubMed: 26664820

Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., et al. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *Neuroimage*, *137*, 165–177. https://doi.org/10.1016/j.neuroimage.2016.05.051, PubMed: 27236083

Delgado, M. R., Li, J., Schiller, D., & Phelps, E. A. (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *363*, 3787–3800. https://doi.org/10.1098/rstb.2008.0161, PubMed: 18829426

Duncan, K., Ketz, N., Inati, S. J., & Davachi, L. (2012). Evidence for area CA1 as a match/mismatch detector: A high-resolution fMRI study of the human hippocampus. *Hippocampus*, *22*, 389–398. https://doi.org/10.1002/hipo.20933, PubMed: 21484934

El-Sourani, N., Trempler, I., Wurm, M. F., Fink, G. R., & Schubotz, R. I. (2019). Predictive impact of contextual objects during action observation: Evidence from functional magnetic resonance imaging. *Journal of Cognitive Neuroscience*, *32*, 326–337. https://doi.org/10.1162/jocn_a_01480, PubMed: 31617822

Epstein, R. A. (2005). The cortical basis of visual scene processing. *Visual Cognition*, *12*, 954–978. https://doi.org/10.1080/13506280444000607

Fairhall, S. L., & Caramazza, A. (2013). Brain regions that represent amodal conceptual knowledge. *Journal of Neuroscience*, *33*, 10552–10558. https://doi.org/10.1523/JNEUROSCI.0051-13.2013, PubMed: 23785167

Foudil, S. A., Kwok, S. C., & Macaluso, E. (2020). Context-dependent coding of temporal distance between cinematic events in the human precuneus. *Journal of Neuroscience*, *40*, 2129–2138. https://doi.org/10.1523/JNEUROSCI.2296-19.2020, PubMed: 31996453

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *360*, 815–836. https://doi.org/10.1098/rstb.2005.1622, PubMed: 15937014

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*, 189–210. https://doi.org/10.1002/hbm.460020402

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *364*, 1211–1221. https://doi.org/10.1098/rstb.2008.0300, PubMed: 19528002

Gershman, S. J., Monfils, M. H., Norman, K. A., & Niv, Y. (2017). The computational nature of memory modification. *eLife*, *6*, e23763. https://doi.org/10.7554/eLife.23763, PubMed: 28294944

Greene, N. R., & Naveh-Benjamin, M. (2021). The effects of divided attention at encoding on item and associative memory. *Memory and Cognition*, *31*, 1021–1035. https://doi.org/10.3758/bf03196123, PubMed: 14704017

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*, 1409–1422. https://doi.org/10.1016/s0042-6989(01)00073-6, PubMed: 11322983

Guo, P., Li, Q., Wang, X., Li, X., Wang, S., Xie, Y., et al. (2020). Structural covariance changes of anterior and posterior hippocampus during musical training in young adults. *Frontiers in Neuroanatomy*, *14*, 20. https://doi.org/10.3389/fnana.2020.00020, PubMed: 32508600

Han, S., O'Connor, A. R., Eslick, A. N., & Dobbins, I. G. (2012). The role of left ventrolateral prefrontal cortex during episodic decisions: Semantic elaboration or resolution of episodic interference? *Journal of Cognitive Neuroscience*, *24*, 223–234. https://doi.org/10.1162/jocn_a_00133, PubMed: 21916561

Haruno, M., & Kawato, M. (2006). Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus–action–reward association learning. *Journal of Neurophysiology*, *95*, 948–959. https://doi.org/10.1152/jn.00382.2005, PubMed: 16192338

Hunt, R. R. (2013). Precision in memory through distinctive processing. *Current Directions in Psychological Science*, *22*, 10–15. https://doi.org/10.1177/0963721412463228

Jainta, B., Siestrup, S., El-Sourani, N., Trempler, I., Wurm, M. F., Werning, M., et al. (2022). Seeing what I did (not): Cerebral

and behavioral effects of agency and perspective on episodic memory re-activation. *Frontiers in Behavioral Neuroscience*, *15*, 793115. https://doi.org/10.3389/fnbeh.2021.793115, PubMed: 35069141

James, T. W., VanDerKlok, R. M., Stevenson, R. A., & James, K. H. (2011). Multisensory perception of action in posterior temporal and parietal cortices. *Neuropsychologia*, *49*, 108–114. https://doi.org/10.1016/j.neuropsychologia.2010.10.030, PubMed: 21036183

Jastorff, J., Clavagnier, S., Gergely, G., & Orban, G. A. (2011). Neural mechanisms of understanding rational actions: Middle temporal gyrus activation by contextual violation. *Cerebral Cortex*, *21*, 318–329. https://doi.org/10.1093/cercor/bhq098, PubMed: 20513657

Larsen, S. F., & Plunkett, K. (1987). Remembering experienced and reported events. *Applied Cognitive Psychology*, *1*, 15–26. https://doi.org/10.1002/acp.2350010104

Lee, J. L. C., Nader, K., & Schiller, D. (2017). An update on memory reconsolidation updating. *Trends in Cognitive Sciences*, *21*, 531–545. https://doi.org/10.1016/j.tics.2017.04.006, PubMed: 28495311

Li, X., Li, Q., Wang, X., Li, D., & Li, S. (2018). Differential age-related changes in structural covariance networks of human anterior and posterior hippocampus. *Frontiers in Physiology*, *9*, 518. https://doi.org/10.3389/fphys.2018.00518, PubMed: 29867561

Lingnau, A., & Downing, P. E. (2015). The lateral occipitotemporal cortex in action. *Trends in Cognitive Sciences*, *19*, 268–277. https://doi.org/10.1016/j.tics.2015.03.006, PubMed: 25843544

Long, N. M., Lee, H., & Kuhl, B. A. (2016). Hippocampal mismatch signals are modulated by the strength of neural predictions and their similarity to outcomes. *Journal of Neuroscience*, *36*, 12677–12687. https://doi.org/10.1523/JNEUROSCI.1850-16.2016, PubMed: 27821577

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage*, *19*, 1233–1239. https://doi.org/10.1016/S1053-8119(03)00169-1, PubMed: 12880848

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, *58*, 25–45. https://doi.org/10.1146/annurev.psych.57.102904.190143, PubMed: 16968210

Moscovitch, M., Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic memory and beyond: The hippocampus and neocortex in transformation. *Annual Review of Psychology*, *67*, 105–134. https://doi.org/10.1146/annurev-psych-113011-143733, PubMed: 26726963

Mumford, J. A., Poline, J. B., & Poldrack, R. A. (2015). Orthogonalization of regressors in fMRI models. *PLoS One*, *10*, e0126255. https://doi.org/10.1371/journal.pone.0126255, PubMed: 25919488

Nader, K. (2015). Reconsolidation and the dynamic nature of memory. *Cold Spring Harbor Perspectives in Biology*, *7*, a021782. https://doi.org/10.1101/cshperspect.a021782, PubMed: 26354895

Nader, K., & Einarsson, E. Ö. (2010). Memory reconsolidation: An update. *Annals of the New York Academy of Sciences*, *1191*, 27–41. https://doi.org/10.1111/j.1749-6632.2010.05443.x, PubMed: 20392274

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage*, *25*, 653–660. https://doi.org/10.1016/j.neuroimage.2004.12.005, PubMed: 15808966

Noppeney, U., & Price, C. J. (2004). Retrieval of abstract semantics. *Neuroimage*, *22*, 164–170. https://doi.org/10.1016/j.neuroimage.2003.12.010, PubMed: 15110006

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113. https://doi.org/10.1016/0028-3932(71)90067-4, PubMed: 5146491

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, *155*, 23–36. https://doi.org/10.1016/S0079-6123(06)55002-2, PubMed: 17027377

Palomero-Gallagher, N., Kedo, O., Mohlberg, H., Zilles, K., & Amunts, K. (2020a). *Probabilistic cytoarchitectonic map of CA1 (hippocampus)* (v11.1) [Data set]. EBRAINS. https://doi.org/10.25493/4A6X-6F0

Palomero-Gallagher, N., Kedo, O., Mohlberg, H., Zilles, K., & Amunts, K. (2020b). Multimodal mapping and analysis of the cyto- and receptorarchitecture of the human hippocampus. *Brain Structure and Function*, *225*, 881–907. https://doi.org/10.1007/s00429-019-02022-4, PubMed: 31955294

Palomero-Gallagher, N., Kedo, O., Mohlberg, H., Zilles, K., & Amunts, K. (2020c). *Probabilistic cytoarchitectonic map of CA2 (hippocampus)* (v11.1) [Data set]. EBRAINS. https://doi.org/10.25493/46G6-T33

Palomero-Gallagher, N., Kedo, O., Mohlberg, H., Zilles, K., & Amunts, K. (2020d). *Probabilistic cytoarchitectonic map of CA3 (hippocampus)* (v11.1) [Data set]. EBRAINS. https://doi.org/10.25493/C546-GS0

Poppenk, J., Evensmoen, H. R., Moscovitch, M., & Nadel, L. (2013). Long-axis specialization of the human hippocampus. *Trends in Cognitive Sciences*, *17*, 230–240. https://doi.org/10.1016/j.tics.2013.03.005, PubMed: 23597720

R Core Team. (2020). *R: A language and environment for statistical computing*. https://www.r-project.org

Reber, P. J., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2005). Priming effects in the fusiform gyrus: Changes in neural activity beyond the second presentation. *Cerebral Cortex*, *15*, 787–795. https://doi.org/10.1093/cercor/bhh179, PubMed: 15371295

Reyna, V. F., Corbin, J. C., Weldon, R. B., & Brainerd, C. J. (2016). How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives. *Journal of Applied Research in Memory and Cognition*, *5*, 1–9. https://doi.org/10.1016/j.jarmac.2015.12.003, PubMed: 27042402

Reyna, V. F., & Lloyd, F. (1997). Theories of false memory in children and adults. *Learning and Individual Differences*, *9*, 95–123. https://doi.org/10.1016/S1041-6080(97)90002-9

Robin, J., & Moscovitch, M. (2017). Details, gist and schema: Hippocampal–neocortical interactions underlying recent and remote episodic and spatial memory. *Current Opinion in Behavioral Sciences*, *17*, 114–123. https://doi.org/10.1016/j.cobeha.2017.07.016

Sarinopoulos, I., Grupe, D. W., Mackiewicz, K. L., Herrington, J. D., Lor, M., Steege, E. E., et al. (2010). Uncertainty during anticipation modulates neural responses to aversion in human insula and amygdala. *Cerebral Cortex*, *20*, 929–940. https://doi.org/10.1093/cercor/bhp155, PubMed: 19679543

Schiffer, A. M., Ahlheim, C., Ulrichs, K., & Schubotz, R. I. (2013). Neural changes when actions change: Adaptation of strong and weak expectations. *Human Brain Mapping*, *34*, 1713–1727. https://doi.org/10.1002/hbm.22023, PubMed: 22422724

Schiffer, A. M., Ahlheim, C., Wurm, M. F., & Schubotz, R. I. (2012). Surprised at all the entropy: Hippocampal, caudate and midbrain contributions to learning from prediction errors. *PLoS One*, *7*, e36445. https://doi.org/10.1371/journal.pone.0036445, PubMed: 22570715

Schiffer, A. M., Krause, K. H., & Schubotz, R. I. (2013). Surprisingly correct: Unexpectedness of observed actions activates the medial prefrontal cortex. *Human Brain*

Mapping, *35*, 1615–1629. https://doi.org/10.1002/hbm.22277, PubMed: 23670963
Schiffer, A. M., & Schubotz, R. I. (2011). Caudate nucleus signals for breaches of expectation in a movement observation paradigm. *Frontiers in Human Neuroscience*, *5*, 38. https://doi.org/10.3389/fnhum.2011.00038, PubMed: 21519392
Schubotz, R. I. (2015). Prediction and expectation. In A. W. Toga (Ed.), *Brain mapping: An encyclopedic reference* (Vol. 3, pp. 295–302). Academic Press/Elsevier. https://doi.org/10.1016/B978-0-12-397025-1.00205-0
Schubotz, R. I., & Von Cramon, D. Y. (2009). The case of pretense: Observing actions and inferring goals. *Journal of Cognitive Neuroscience*, *21*, 642–653. https://doi.org/10.1162/jocn.2009.21049, PubMed: 18564039
Scully, I. D., Napper, L. E., & Hupbach, A. (2017). Does reactivation trigger episodic memory change? A meta-analysis. *Neurobiology of Learning and Memory*, *142*, 99–107. https://doi.org/10.1016/j.nlm.2016.12.012, PubMed: 28025069
Sekeres, M. J., Bonasia, K., St-Laurent, M., Pishdadian, S., Winocur, G., Grady, C., et al. (2016). Recovering and preventing loss of detailed memory: Differential rates of forgetting for detail types in episodic memory. *Learning and Memory*, *23*, 72–82. https://doi.org/10.1101/lm.039057.115, PubMed: 26773100
Sekeres, M. J., Winocur, G., & Moscovitch, M. (2018). The hippocampus and related neocortical structures in memory transformation. *Neuroscience Letters*, *680*, 39–53. https://doi.org/10.1016/j.neulet.2018.05.006, PubMed: 29733974
Sekeres, M. J., Winocur, G., Moscovitch, M., Anderson, J. A. E., Pishdadian, S., Martin Wojtowicz, J., et al. (2018). Changes in patterns of neural activity underlie a time-dependent transformation of memory in rats and humans. *Hippocampus*, *28*, 745–764. https://doi.org/10.1002/hipo.23009, PubMed: 29989271
Siestrup, S., Jainta, B., Cheng, S., & Schubotz, R. I. (2023). Solidity meets surprise: Cerebral and behavioral effects of learning from episodic prediction errors. *Journal of Cognitive Neuroscience*, *35*, 291–313. https://doi.org/10.1162/jocn_a_01948, PubMed: 36473102
Siestrup, S., Jainta, B., El-Sourani, N., Trempler, I., Wurm, M. F., Wolf, O. T., et al. (2022). What happened when? Cerebral processing of modified structure and content in episodic cueing. *Journal of Cognitive Neuroscience*, *34*, 1287–1305. https://doi.org/10.1162/JOCN_A_01862, PubMed: 35552744
Sinclair, A. H., & Barense, M. D. (2018). Surprise and destabilize: Prediction error triggers episodic memory updating. *Learning & Memory*, *25*, 369–381. https://doi.org/10.1101/lm.046912.117.25, PubMed: 30012882
Sinclair, A. H., & Barense, M. D. (2019). Prediction error and memory reactivation: How incomplete reminders drive reconsolidation. *Trends in Neurosciences*, *42*, 727–739. https://doi.org/10.1016/j.tins.2019.08.007, PubMed: 31506189
Sinclair, A. H., Manalili, G. M., Brunec, I. K., Adcock, R. A., & Barense, M. D. (2021). Prediction errors disrupt hippocampal representations and update episodic memories. *Proceedings of the National Academy of Sciences, U.S.A.*, *118*, e2117625118. https://doi.org/10.1073/pnas.2117625118, PubMed: 34911768
Sommer, V. R., & Sander, M. C. (2022). Contributions of representational distinctiveness and stability to memory performance and age differences. *Aging, Neuropsychology, and Cognition*, *29*, 443–462. https://doi.org/10.1080/13825585.2021.2019184, PubMed: 34939904
St. Jacques, P. L., Olm, C., & Schacter, D. L. (2013). Neural mechanisms of reactivation-induced updating that enhance and distort memory. *Proceedings of the National Academy of Sciences, U.S.A.*, *110*, 19671–19678. https://doi.org/10.1073/pnas.1319630110, PubMed: 24191059
Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*, 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114, PubMed: 11752477
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, *15*, 273–289. https://doi.org/10.1006/nimg.2001.0978, PubMed: 11771995
Vassena, E., Holroyd, C. B., & Alexander, W. H. (2017). Computational models of anterior cingulate cortex: At the crossroads between prediction and effort. *Frontiers in Neuroscience*, *11*, 316. https://doi.org/10.3389/fnins.2017.00316, PubMed: 28634438
Volz, K. G., Schubotz, R. I., & Von Cramon, D. Y. (2003). Predicting events of varying probability: Uncertainty investigated by fMRI. *Neuroimage*, *19*, 271–280. https://doi.org/10.1016/S1053-8119(03)00122-8, PubMed: 12814578
Volz, K. G., Schubotz, R. I., & Von Cramon, D. Y. (2005). Variants of uncertainty in decision-making and their neural correlates. *Brain Research Bulletin*, *67*, 403–412. https://doi.org/10.1016/j.brainresbull.2005.06.011, PubMed: 16216687
Wahlheim, C. N., & Zacks, J. M. (2019). Memory guides the processing of event changes for older and younger adults. *Journal of Experimental Psychology: General*, *148*, 30–50. https://doi.org/10.1037/xge0000458, PubMed: 29985021
Webb, C. E., Turney, I. C., & Dennis, N. A. (2016). What's the gist? The influence of schemas on the neural correlates underlying true and false memories. *Neuropsychologia*, *93*, 61–75. https://doi.org/10.1016/j.neuropsychologia.2016.09.023, PubMed: 27697593
Wiggett, A. J., & Downing, P. E. (2011). Representation of action in occipito-temporal cortex. *Journal of Cognitive Neuroscience*, *23*, 1765–1780. https://doi.org/10.1162/jocn.2010.21552, PubMed: 20807060
Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society*, *17*, 766–780. https://doi.org/10.1017/S1355617711000683, PubMed: 21729403
Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 143–146). New York: Association for Computing Machinery. https://doi.org/10.1145/1978942.1978963
Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited—Again. *Neuroimage*, *2*, 173–181. https://doi.org/10.1006/nimg.1995.1023, PubMed: 9343600
Wurm, M. F., & Caramazza, A. (2022). Two 'what' pathways for action and object recognition. *Trends in Cognitive Sciences*, *26*, 103–116. https://doi.org/10.1016/j.tics.2021.10.003, PubMed: 34702661
Zaretsky, M., Mendelsohn, A., Mintz, M., & Hendler, T. (2010). In the eye of the beholder: Internally driven uncertainty of danger recruits the amygdala and dorsomedial prefrontal cortex. *Journal of Cognitive Neuroscience*, *22*, 2263–2275. https://doi.org/10.1162/jocn.2009.21402, PubMed: 19929759