What happened when? Brain and behavioral responses to modified structure and content in episodic cueing

Sophie Siestrup[1,2*], Benjamin Jainta[1], Nadiya El-Sourani[1], Ima Trempler[1,2], Oliver T. Wolf[3], Sen Cheng[4], Ricarda I. Schubotz[1,2]

[1] Department of Psychology, University of Münster, Münster, Germany,

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

[3] Department of Cognitive Psychology, Faculty of Psychology, Ruhr University Bochum, Bochum, Germany

[4] Institute for Neural Computation, Ruhr University Bochum, Bochum, Germany

*** Correspondence:**
Sophie Siestrup
s.siestrup@uni-muenster.de

Number of words: 8027
Number of figures: 5
Number of tables: 1

Keywords: episodic memory, fMRI, memory modification, prediction error, memory trace

ABSTRACT

Episodic memories are not static but can be modified on the basis of new experiences, potentially allowing us to make valid predictions in the face of an ever-changing environment. Recent research has identified prediction errors during retrieval as a possible trigger for such changes. In the present study, we used modified episodic cues to investigate whether different types of mnemonic prediction errors modulate brain activity and subsequent memory performance. Participants encoded different episodes which consisted of short toy stories. During a subsequent functional magnetic resonance imaging (fMRI) session, participants were then presented videos showing the original episodes, or modified versions thereof. In modified videos, either the order of two subsequent action steps was changed or an object was exchanged for another. Content modifications recruited inferior frontal, parietal, and temporo-occipital areas reflecting the processing of the new object information, while brain responses to structure modifications in right dorsal premotor and inferior frontal areas remained subthreshold. In a post-fMRI memory test, the participants' tendency to accept modified episodes as originally encoded increased significantly when they had been presented modified versions already during the fMRI session. Our study provides valuable initial insights into the neural processing of different types of episodic prediction errors and their influence on subsequent memory.

1 INTRODUCTION

Episodic memories enable us to vividly re-live events that we experienced at some point in our personal life (Tulving, 2002). But there is evidence that they are not always veridical reconstructions of our past. On the contrary, memories are flexible and prone to change (Nader & Einarsson, 2010; Nader, 2015; Lee et al., 2017; Scully et al., 2017). In everyday life, situations we encounter are

usually not exactly the same as those we experienced before. So, there is always a certain discrepancy between our expectations, which we derive from our memories, and the events we then actually experience. According to the predictive coding framework, this discrepancy leads to a prediction error (Barto et al., 2013; Reichardt et al., 2020). Generally, prediction errors serve as bottom-up learning signals that allow us to adapt our internal predictive models to an ever-changing environment, so that we can maintain valid predictions in the long run (Friston, 2005; Friston & Kiebel, 2009; Schubotz, 2015). Accordingly, it is likely adaptive that memories are modified in favor of valid internal models informed and updated by later experiences (Exton-McGuinness et al., 2015; Fernández et al., 2016). Indeed, recent research has identified mnemonic prediction errors as potential drivers of memory change (Sinclair & Barense, 2019).

In the present study, we examined the neural and behavioral level of mnemonic prediction errors during episodic retrieval by targeting two basic types of episodic memory information: either their content ("what") or their structure ("when") (cf. Griffiths et al., 1999). To do so, we employed a new episodic cueing paradigm based on previous work (Schiffer et al., 2012; 2013). After encoding short episodes from videos, participants went through a functional magnetic resonance imaging (fMRI) session and were either presented original episode videos or slightly modified versions thereof. A subset of videos was manipulated with regard to the occurrence of an object (content modification) or the order of two consecutive action steps (structure modification) to elicit different types of mnemonic prediction errors. In a post-fMRI memory test, participants' memory for original and modified episodes was probed.

If new content and/or structure information elicited changes in the predictive model during fMRI, we expected that this would reduce memory performance in the post-fMRI memory test (Schiffer et al., 2012; 2013). This could be reflected in a weakening of the original episodic memory, i.e., false rejection of original videos in the memory test, and/or in the creation of alternative episode representations, i.e., false acceptance of modified videos in the memory test.

Concerning brain activations, we expected that some regions would respond to both types of episodic modifications. Particularly, the medial frontal cortex (MFC) is a candidate area that may serve more general control over consolidation and retrieval of long-term memories (Euston et al., 2012; Peters, et al., 2013). Furthermore, the hippocampus is regarded a core structure of episodic memory (Maguire et al., 2016; Horner & Doeller, 2017; Stachenfeld et al., 2017) and responds to mnemonic prediction errors (Long et al., 2016; Bein et al., 2020). Further, we expected increased activity in brain areas related to action observation (Caspers et al., 2010) since we previously found them attenuated when subjects adapted to episode modifications affecting both objects and sequential structure combined (Schiffer et al., 2012; 2013).

Apart from these common effects, structure and content episodic modifications were also expected to affect different regions. Structure modifications should elevate activity in premotor areas due to their central role in sequential order processing (Schubotz, 2004). Dorsal premotor and adjacent prefrontal sites along the superior frontal sulcus (PMd/SFS) were found for stepwise ordinal linking of individual action steps, as required in quite different, but especially predictive, tasks (Stadler et al., 2011; Schubotz et al., 2012; Tamber-Rosenau et al., 2011; Hrkać et al., 2014; Pomp et al., 2021). By contrast, content modifications were expected to engage areas related to object processing, including lateral occipitotemporal cortex (OTC) (Lingnau & Downing, 2015), anterior intraparietal sulcus (IPS) (Creem-Regehr, 2009; Schubotz et al., 2014), and fusiform gyrus (FG; Reber et al., 2005). Finally, based on earlier findings, we tested the hypothesis that within the inferior frontal gyrus (IFG), Brodmann area 44 (BA 44) may be especially involved in structure modification due to its

role in processing the sequential structure of actions (Fiebach & Schubotz, 2006; Clerget et al., 2009) whereas BA 45 may be more relevant for semantic processing, hence also content modification of episodes (Badre & Wagner, 2007; Heim et al., 2009).

2 MATERIALS AND METHODS

2.1 Participants

Forty female participants took part in the study. Only females participated in order to achieve a good match between participants' hands and the actress's hands in the stimulus material, as participants were supposed to believe they were shown in the videos themselves. Two participants did not return for the fMRI session after training. Additionally, data from one participant had to be excluded from analysis due to nausea and dizziness during the fMRI session. In another case, an unknown technical problem led to the incomplete acquisition of functional data. Thus, the final sample for fMRI analyses consisted of 36 participants, age 18 – 28 years (*M* = 22.67 years, *SD* = 2.40 years). Participants of the final sample were all right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). Handedness scores varied from +60 to +100 (*M* = 92.17, *SD* = 10.95). Data from one participant was not used for the analysis of the post-fMRI memory test due to misunderstood instructions. All participants had normal or corrected-to-normal vision, were native German speakers, and reported no history of neurological or psychiatric disorders or substance abuse. Participants signed a written informed consent and received monetary compensation or course credits for their participation. The study was conducted in accordance with the Declaration of Helsinki and approved by the Local Ethics Committee of the University of Münster.

2.2 Stimuli

Even though episodic memory can potentially never be completely independent of semantic knowledge (Cheng et al., 2016) we wanted to reduce influences of everyday experiences as far as possible in our study. To this end we chose to use abstract but complex toy-based stories so that we could ensure controlled encoding of unique episodes in the laboratory.

Stimuli were 78 short videos (duration = 8.80 – 17.88 s, *M* = 12.71 s) of stories that were played with PLAYMOBIL® toys, showing only the toys and hands and underarms of an actress. Stories comprised six to nine action steps (*M* = 7.4 steps) and four to 14 separable objects (*M* = 6.93 objects), such as characters, animals, vehicles and tools. The exact same object appeared in only one of the stories.

Stories were filmed from above with a digital single-lens reflex camera (Nikon D5300) which was centrally mounted above the table and faced straight down. Matt white paper served as a base. A frame of 47.5 cm x 28 cm was taped on the paper, congruent with the section captured by the camera (in the following referred to as camera frame). Objects that were needed for a particular story were positioned next to the camera frame and were only moved into view in the moment at which they appeared in the story. During filming, the actress wore a black pullover and black rubber gloves. To facilitate future imitation from demo videos, the back of the right hand was marked with a yellow dot (Franz et al., 2007). Video material was edited using Adobe Premiere Pro CC (Adobe Systems Software, Dublin, Ireland, Version 12.1.2). All videos had a frame of size 1920 x 1080 pixels and a frame rate of 25 frames per second. Videos started with seven frames showing only background and ended after seven frames showing the final toy constellation. The original perspective of the videos was the third-person (observer) perspective (3pp) and they were rotated 180° to create their first-person (field) perspective (1pp) counterparts using Adobe Premiere Pro CC (Wurm & Schubotz,

2018). Throughout the experiment, videos were presented at a visual angle of approximately 7.3° x 13° using the stimulus presentation software Presentation (version 20.3 02.01.19, NeuroBehavioral Systems, Berkeley, CA, USA).

On the basis of two pilot studies, we chose 24 out of originally 30 stories for our stimulus set. Stories were excluded when they were particularly difficult to imitate or describe. One of the 30 stories was excluded due to low memorability as indicated by low performance in a signal detection task.

The 24 final stories existed in three different versions each: (1) an original version (*ori*) as encoded by the participants, (2) a version in which two adjacent action steps were switched (structure modification) (*str*) and (3) another variation of the original video in which one object was exchanged (content modification) (*con*). Story scripts were created by five experimenters who all had to agree that the original story as well as modifications thereof were semantically valid (within a toy world) and that modifications did not change the overall outcome of the story. For creating videos with modifications, the respective stories were played and filmed again exactly the way as for the original video. The only aspect that differed between original and modified versions was a single change of either the order of two action steps (i.e., one transition out of 7.33 transitions on average, for structure modifications) or one object (i.e., one object out of 6.95 objects on average, for content modifications). Modifications were never introduced in the first two action steps so that the beginning of a video served as a cue for prediction. Furthermore, 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 which diverged from the original version. Six other stories were used in one version only. Four of them were presented for the first time in the fMRI session, we refer to them as untrained episodes in the following. The two remaining videos were only used for practice and did not appear in the fMRI experiment.

For examples of different story versions, see Figure 1 and Supplementary Table 1. Videos are available upon reasonable request via the Action Video Corpus Muenster (https://www.uni-muenster.de/IVV5PSY/AvicomSrv/).

- insert **Figure 1** here -

2.3 Procedure

2.3.1 Training

Training sessions were conducted in a computer lab at the Department of Psychology at the University of Münster. The training consisted of two sessions that took place on two consecutive days and lasted about 2 h and 1.5 h, respectively. During each of the two sessions, participants encoded half of the episodes. We chose to split the training over two days to avoid fatigue or a decrease in motivation due to the relatively long duration of the task. For encoding episodes, participants imitated half of the 24 stories from demo videos (i.e., the original versions) and only observed the other half, equally distributed over both training days (factor AGENCY).

The 24 demo videos were organized in four subsets, containing six videos each, balanced for the number of action steps (*A1*, *A2*, *B1*, *B2*). Participants either imitated videos belonging to *A* subsets and observed videos belonging to *B* subsets or vice versa. In each training session, participants went through one imitation block and one observation block, each consisting of one subset of videos. This means that each participant trained each video either during session one or during session two, the same video was not trained on both days. The order of videos within these blocks was randomized

for each participant. All stories were imitated and only observed equally often over the course of the study. The first training session started with a short practice phase of the task (2 videos). Then, half of the participants started with imitation, the other half with observation. The same applied to the second training session, except that participants who had started with imitation on the first day now started with observation, and vice versa. The order of blocks and the order in which video subsets were trained was balanced between participants. Participants always went through each story eight times, including four times in 3pp and another four times in 1pp. If the story was to be imitated, participants first watched the story four times in 3pp, then once in 1pp and then performed the last three repetitions of the story themselves. If the story was only observed, the video was instead presented three more times in 1pp. This way, participants experienced each story equally often from each visual perspective, allowing us to establish the factor PERSPECTIVE in the fMRI session through the presentation of just one of both perspectives. The optimal number of video presentations for successful imitation had previously been determined in a pilot study.

During training, the participant sat at the same setup which had been used for filming the stimulus material and likewise wore a black pullover and gloves with a yellow dot on the right hand. The experimenter sat opposite of the participant, supervising the performance (Figure 2). For each story, the toys were positioned next to the camera frame, following the same arrangement as used while creating the stimulus material. After imitation or observation, participants had to deliver a detailed description of the story to ensure that they understood it correctly and had paid attention to all objects involved. If participants made a mistake during an imitation or description trial, they were immediately interrupted by the experimenter to avoid encoding of incorrect scripts. They would then start over with a new imitation/description attempt.

The two described experimental factors AGENCY (*imitation*, *observation*) and PERSPECTIVE (*1pp*, *3pp*) are not further addressed in the present paper, as they are central for a companion paper (Jainta et al., 2021) describing the influence of personal involvement on recall. Both factors were balanced with respect to the factors reported here, so we can exclude any confounding effects (fully-crossed design).

- insert **Figure 2** here -

2.3.2 fMRI session

The fMRI session took place approximately one week after the second training session (range = 7 - 12 days; *M* = 7.42 days, *SD* = 0.90 days). Participants were told that videos of themselves playing the stories would be presented in the fMRI session. Even though participants had actually been filmed during training, these videos were not used during the fMRI experiment. This was only a cover story to elevate personal identification with the videos to benefit episode reactivation. Participants were fully debriefed after completion of the study.

During the fMRI session, participants were presented with original and modified videos reminiscent of the previously encoded episodes. Importantly, each video was only shown in the original or one divergent version. Following a previously used paradigm (Schiffer et al., 2012; 2013), modified and original episodes were presented repeatedly to simulate the natural circumstances that potentially foster memory modification, i.e. updating of internal models due to increasing evidence for the validity of an alternative. Thus, eight videos were repeatedly presented in the original version, eight included a structure modification and eight a content modification. Which stories belonged to which conditions was counterbalanced between participants. Half of the videos were shown in 1pp, the

other half in 3pp, perfectly balanced over conditions. Additionally, four untrained stories were included in the fMRI session.

The fMRI experiment consisted of six blocks, each containing the 24 videos reminiscent of the previously encoded episodes. Consequently, each video was presented six times over the course of the session. Within blocks, videos were presented in pseudo-random order so that transition probabilities between conditions were balanced. Additionally, each block contained three null events during which only a fixation cross was presented (duration: 7 – 10 s). Furthermore, each untrained video was presented once per block. Therefore, the whole experiment contained 18 null events and 24 untrained video trials. Participants were not informed about the block-structure of the experiment.

Participants were instructed to attentively watch the presented videos. They were told that after some videos a short description would be presented (e.g. “Rescuing princess”) that either matched or did not match the story shown in the video (question trials). The task was to either accept or reject the description by pressing one of two buttons on a response box with the right index or middle finger, respectively. This type of task has been used successfully before to focus participants’ attention on complex video stimuli (El-Sourani et al., 2019). The descriptions represented the stories on a gist-level, so specific details were irrelevant to the correct answer (Figure 3). Throughout the entire experiment, each story was once followed by a matching description and once by a non-matching description, resulting in a total number of 56 question trials in the experiment. Each block contained nine to ten question trials and per block, approximately 50 % of descriptions were to be accepted, 50 % were to be rejected. The question was presented for a maximum of 3 s or until participants responded. Upon response delivery, participants received 1 s written feedback whether they answered correctly, incorrectly, or too late, in case no response was given. Participants were naïve with regard to this distribution of question trials.

Between trials, a fixation cross was presented for a duration of 2 s (1 s after question trials) to serve as an interstimulus interval. Before each trial, a variable jitter of 0, 0.5, 1 or 1.5 s of fixation was added for enhancement of the temporal resolution of the BOLD response (Figure 3).

In total, the fMRI task had a duration of approximately 48 min. The task had previously been practiced at the end of the second training session (four video trials, four question trials, one null event). Practice videos were shown twice each and not re-used in the fMRI session.

- insert **Figure 3** here -

2.3.3 Post-fMRI memory test

Immediately after the fMRI session, participants completed an explicit memory test. Importantly, encoding occurred incidentally, as participants were not informed beforehand that their memory for episodes would be tested.

Participants were seated in a separate room in front of a laptop and instructed to remember their training sessions one week prior. They were presented all stories that they had seen in the fMRI session in two different versions. More precisely, when modified videos had been presented during the fMRI session, these modified videos were presented again during the memory test and *additionally* each story was shown in the original version. When original episode videos had been presented during the fMRI scan, these original videos were presented again in the memory test and *additionally*, each story was shown in a modified version, *either* containing a structure modification in half of the cases *or* a content modification.

The participants' task was to rate after each video whether they knew this exact episode from the training, using a Likert scale including 1 (yes), 2 (rather yes), 3 (rather no) and 4 (no), by pressing one out of four marked keys on the laptop's keyboard. Response time was not restricted but participants were instructed to respond quickly and intuitively. Videos were presented in a pseudo-randomized 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. Untrained videos were shown twice in the same version, so that, in total, the memory test comprised 56 video trials. The completion of the task took approximately 15 min.

2.4 MRI data acquisition and preprocessing

MRI scans were conducted with a 3-Tesla Siemens Magnetom Prisma MR tomograph (Siemens, Erlangen, Germany) using a 20-channel head coil. Participants lay supine on the scanner bed with their right index and middle finger positioned on the two appropriate buttons on a response box. Head, arm and hand movements were minimized by tight fixation with form-fitting cushions. Participants were provided with earplugs and headphones to attenuate scanner noise. Stimuli were projected on a screen that the participants saw via an individually adjusted mirror which was mounted on the head coil.

High resolution T1 weighted anatomical images were obtained with a 3D-multiplanar rapidly acquired gradient-echo (MPRAGE) sequence prior to functional imaging. 192 slices with a thickness of 1 mm were acquired, using a repetition time (TR) of 2130 ms, an echo time (TE) of 2.28 ms, a flip angle of 8° and a field of view (FoV) of 256 x 256 mm². Functional images of the whole brain were acquired in interleaved order along the AC–PC plane using a gradient-echo echoplanar imaging (EPI) sequence to measure blood-oxygen-level-dependent (BOLD) contrast. Thirty-three axial slices with a thickness of 3 mm were obtained in an interleaved order, using a TR of 2000 ms, a TE of 30 ms, a FoV of 192 x 192 mm² and a flip angle of 90°.

Processing of imaging data was conducted with SPM12 (Wellcome Trust, London, England) implemented in MATLAB (version R2018b, The MathWorks Inc., Natick, MA, USA). Data were preprocessed by slice time correction to the middle slice, movement correction and realignment to the mean image, co-registration of the functional data to individual structural scans, 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 full-width at half maximum (FWHM) of 8 mm. Further, a 128 s high-pass temporal filter was applied.

2.5 Statistical data analysis

2.5.1 Behavioral data

The behavioral data analysis was conducted using RStudio (R Core Team, 2020; version 1.3.1073).

To check our hypothesis that repeated presentation of modified videos in the fMRI session leads to a decrease in memory performance in the memory test, we considered the discrimination index P_r (Snodgrass & Corwin, 1988). In addition, we evaluated ratings for modified and unmodified videos separately to assess the extent of erroneous acceptance of modified videos as originally encoded and/or erroneous rejection of original episodes. Further, we examined reaction times in the memory test, which can serve as an indicator of how long it takes to retrieve information (correctly) from memory (Collins & Quillian, 1969). Longer reaction times indicate increased difficulty of retrieval

due to higher cognitive processing demands (Larsen & Plunkett, 1987; Noppeney & Price, 2004), which may also occur when competing versions of an episode must be processed.

For the analysis of P_r values (unmodified videos = targets, modified videos = distractors) as well as ratings and reaction times for modified videos in the memory test ($\text{modified}_{\text{MT}}$), we applied a 2 x 2 within-subject factorial design with the factors $\text{MODIFICATION}_{\text{FMRI}}$ (*yes*, *no*) and $\text{VERSION}_{\text{MT}}$ (*str*, *con*). For analyzing responses to original videos in the memory test ($\text{original}_{\text{MT}}$), we applied a within-subject design with the factor $\text{VERSION}_{\text{FMRI}}$ (*ori*, *str*, *con*).

For each factorial combination, we calculated the discrimination index P_r (hit rate minus false alarm rate; ratings "yes" and "rather yes" were grouped as "acceptance", and "no" and "rather no" as "rejection"), the average rating, and the average reaction time for correct responses. We report additional analyses on hit rates and false alarm rates in the Supplementary Material. Finally, we compared ratings for untrained video to those for $\text{original}_{\text{MT}}$ and $\text{modified}_{\text{MT}}$ videos. Please note that low ratings indicate acceptance, while high ratings indicate rejection.

We also conducted an explorative analysis on behavioral data from the fMRI session. We calculated the error rate and mean reaction time according to the within-subject factor $\text{VESRION}_{\text{FMRI}}$ (*ori*, *str*, *con*) per participant. No response was given in only 0.7 % of all question trials, and these trials were not further considered in the analysis.

For the choice of statistical tests, data were inspected for normal distribution using the Shapiro Wilk Test. Further, data were checked for extreme outliers as defined as values above quartile 3 + 3* interquartile range (*IQR*) or lower than quartile 1 – 3 * *IQR*. When data were normally distributed or could be transformed to fit normal distribution (reaction times for original videos, reaction times during fMRI task; logarithmic transformation) and showed no extreme outliers, we used repeated measures analyses of variance (rmANOVA). Post-hoc pair-wise comparisons were conducted with paired *t*-tests (one-tailed when comparing *ori* and *str* and *ori* and *con*, two-tailed when comparing *str* and *con*). When the prerequisites for parametric analysis were not met, we used a non-parametric rmANOVA based on aligned rank data (package *ARTool*; Wobbrock et al., 2011; P_r values, ratings, error rates). Several participants did not give any correct answers (i.e. rejection) in response to $\text{modified}_{\text{MT}}$ videos for one or more factorial combinations. To avoid the listwise exclusion of many datapoints for reaction times when using a rmANOVA approach, we decided to use one-tailed Wilcoxon signed rank tests when comparing reaction time between levels of the factor $\text{MODIFICATION}_{\text{FMRI}}$, separately for each modification type. This way, we could include 16 datapoints for structure and 29 datapoints for content modified videos. We additionally excluded another datapoint for content modifications due to a very long mean reaction time in the *yes* condition (8.85 s) which likely does not represent purely mnemonic processes. Please note that this did not influence the overall outcome of the analysis.

When parametric statistical approaches were used, we report (untransformed) mean values and standard errors of the mean as descriptive statistics, when non-parametric approaches were applied, we report median values and *IQR*. For all behavioral analyses, we applied a significance level of α = 0.05. *P*-values were adjusted according to the Bonferroni correction for multiple comparisons (Bonferroni, 1936).

2.5.2 fMRI design specifications

Statistical analyses of the fMRI data were conducted with SPM12. We used a general linear model (GLM) for serially autocorrelated observations (Friston et al., 1994; Worsley & Friston, 1995) and convolved regressors with the canonical hemodynamic response function. Regressors were original videos (*ori*), videos containing a structure modification (*str*) and videos containing a content modification (*con*), each comprising 48 trials. For *str* and *con* trials, the onsets of events were time-locked to the point in the video at which the modification occurred (time of modification). For *ori* trials, we calculated a hypothetical time of modification (mean of times that corresponded to points of structure and content modification in the non-modified video) to serve as a comparable onset. These conditions were modeled as events as we were interested in the phasic effect of the prediction violation at the precise moment it occurred. To each of those regressors, we added a parametric modulator to model the repeated presentation of each video. The 24 untrained videos were modelled as events as well, with onsets timed to the middle of the video. Two additional regressors modeled the 18 null events and the 56 question trials. The modeled activation of null events and questions was time-locked to their respective onsets. Null events were modeled as epochs, containing their full presentation time (7 – 10 s), while questions were modeled as events. The six subject-specific rigid-body transformations obtained from realignment were included as regressors of no interest. Therefore, the GLM comprised 15 regressors in total.

We calculated first-level-*t*-contrasts for *str* > *ori* and *con* > *ori* to analyze brain activity in response to the specific modification types. Grey matter masking was applied on the first level of the analysis. For masking, we used the smoothed individual normalized grey matter image (8 mm FWHM) which was thresholded at 0.2 using ImCalc in SPM12 to create a binary mask. Second-level group analyses were performed by using one-sample *t*-tests across participants. A conjunction of *str* > *ori* and *con* > *ori* contrasts was calculated to detect shared effects of both modifications (Nichols et al., 2005). We applied false discovery rate (FDR) correction and a threshold of $p < 0.05$ (peak level). In case no significant activation was detected using this threshold, we applied no correction for multiple comparison with $p < 0.001$. Brain activation patterns were visualized with the software MRIcroGL (version 1.2.20200331, McCausland Center for Brain Imaging, University of South Carolina, USA).

To test our hypothesis that BA 44 (IFG pars opercularis) is more involved in detecting structure modification than content changes in episodes, while the reverse pattern may hold for BA 45 (IFG pars triangularis), separate structural ROIs were generated for the two areas in each hemisphere using the automated anatomical labeling atlas (Tzourio-Mazoyer et al., 2002) with the WFU PickAtlas toolbox (Maldjian et al., 2003) in SPM12. Mean beta values for the *str* > *ori* and *con* > *ori* contrasts were extracted from each ROI using the MarsBaR toolbox (Brett et al., 2002). The conditions for parametric analysis were confirmed as described above for behavioral data. We performed a rmANOVA with the factors HEMISPHERE (*left*, *right*), ROI (*BA 44*, *BA 45*), and CONTRAST (*str* > *ori*, *con* > *ori*). Post-hoc comparisons were conducted with one-tailed paired *t*-tests and we report adjusted *p*-values (Bonferroni, 1936).

3 RESULTS

3.1 Post-fMRI memory test

3.1.1 Discrimination index P_r

First, we investigated the general memory performance in the memory test, using the discrimination index P_r. A non-parametric rmANOVA with the factors $\text{MODIFICATION}_{\text{FMRI}}$ (*yes*, *no*) and

$VERSION_{MT}$ (*str, con*) revealed a significant main effect of $MODIFICATION_{FMRI}$ ($F_{(1,34)} = 20.220$, $p < 0.001$, $\eta p^2 = 0.373$) which was driven by higher P_r values for *no* ($Median_{no} = 0.50$, $IQR_{no} = 0.375$) compared to *yes* ($Median_{yes} = 0.250$, $IQR_{yes} = 0.313$), indicating a better memory performance when no modifications had been presented during the fMRI session. There was also a significant main effect of $VERSION_{MT}$ ($F_{(1,34)} = 85.828$, $p < 0.001$, $\eta p^2 = 0.716$), which was explained by higher P_r values for *con* ($Median_{con} = 0.688$, $IQR_{con} = 0.313$) than for *str* ($Median_{str} = 0.125$, $IQR_{str} = 0.250$). There was no significant interaction of $MODIFICATION_{FMRI}$ and $VERSION_{MT}$ ($F_{(1,34)} = 2.074$, $p = 0.159$, $\eta p^2 = 0.058$; Figure 4A).

3.1.2 Ratings and reaction times for $modified_{MT}$ videos

Next, we investigated whether the repeated presentation of modified episodes leads to the false acceptance of the same as originally encoded in the memory test. We computed a non-parametric rmANOVA with the factors $MODIFICATION_{FMRI}$ (*yes, no*) and $VERSION_{MT}$ (*str, con*), which revealed a significant main effect of $MODIFICATION_{FMRI}$ ($F_{(1,34)} = 25.554$, $p < 0.001$, $\eta p^2 = 0.429$). The effect was driven by more rejective ratings for *no* ($Median_{no} = 2.625$, $IQR_{no} = 0.938$) than for *yes* ($Median_{yes} = 2.25$, $IQR_{yes} = 1.094$). Thus, when modified episodes had been presented during the fMRI session, participants were more likely to accept the same modified versions in the memory test than when original episodes had been presented. Additionally, there was a significant main effect of $VERSION_{MT}$ ($F_{(1,34)} = 122.647$, $p < 0.001$, $\eta p^2 = 0.783$). More rejective ratings were detected for *con* ($Median_{con} = 3.250$, $IQR_{con} = 1.406$) than for *str* ($Median_{con} = 1.563$, $IQR_{con} = 0.719$), which means that participants were generally more prone to accept structurally modified videos than those with modified content. There was no significant interaction of the two factors $MODIFICATION_{FMRI}$ and $VERSION_{MT}$ ($F_{(1,34)} = 2.010$, $p = 0.165$, $\eta p^2 = 0.056$; Figure 4B). The same results were obtained when analyzing false alarm rates, which were significantly increased after the presentation of modified episodes in the fMRI session and, in addition, for structurally modified videos in general (cf. Supplementary Results for details).

We found a near-significant trend that reaction times were longer for structurally modified videos when this version had already been presented in the fMRI session ($Z = -1.80$, $p = 0.072$; Figure 4C). No significant difference was found concerning reaction times for videos with content modification ($Z = -1.089$, $p = 0.276$; Figure 4D).

3.1.3 Ratings and reaction times for $original_{MT}$ videos

To see whether the experience during the fMRI session influenced acceptance of original videos in the memory test, we calculated a non-parametric rmANOVA for the factor $VERSION_{FMRI}$ (*ori, str, con*). Descriptively, ratings indicated a slightly lower acceptance of $original_{MT}$ videos when modified videos had been presented during the fMRI session ($Median_{ori} = 1.0$, $IQR_{ori} = 0.313$; $Median_{str} = 1.250$, $IQR_{str} = 0.375$; $Median_{con} = 1.125$, $IQR_{con} = 0.50$), but this effect was not significant ($F_{(2,68)} = 1.5822.010$, $p = 0.213$, $\eta p^2 = 0.044$; Figure 4E). This finding was also reflected by ceiling hit rates (cf. Supplementary Results for details).

Further, a rmANOVA revealed a significant main effect of $VERSION_{FMRI}$ on reaction times for $original_{MT}$ videos ($F_{(2,68)} = 3.725$, $p = 0.029$, $\eta p^2 = 0.099$). Pairwise comparisons revealed that reaction times were significantly longer when structurally modified videos had been presented in the scanner compared to original videos ($t_{(34)} = 3.287$, $p = 0.004$). No significant differences were detected between the levels *con* and *ori* ($t_{(34)} = 1.779$, $p = 0.126$) and *str* and *con* ($t_{(34)} = 0.725$, $p = 1$; Figure 4F).

- insert **Figure 4** here -

3.1.4 Ratings for untrained$_{MT}$ videos

To control for a general acceptance bias, we considered ratings for untrained$_{MT}$ videos as well. A non-parametric rmANOVA on ratings for original$_{MT}$, modified$_{MT}$ and untrained$_{MT}$ videos revealed a significant difference between the three types of videos ($F_{(2,68)} = 418.16$, $p < 0.001$, $\eta p^2 = 0.925$), as ratings for untrained$_{MT}$ videos ($Median_{untr} = 4$, $IQR_{untr} = 0$) were significantly more rejective than for original$_{MT}$ ($Median_{ori} = 1.208$, $IQR_{or\ i} = 0.250$; $Z = -5.283$, $p < 0.001$) and modified$_{MT}$ videos ($Median_{mod} = 2.469$, $IQR_{mod} = 1.031$; $Z = -5.280$, $p < 0.001$). This finding indicates that participants correctly remembered that untrained videos had not been part of the training, even though they had been repeatedly presented during the fMRI session.

3.2 Behavioral performance during fMRI session

We calculated a non-parametric rmANOVA on error rates for the fMRI task with the factor VESRION$_{FMRI}$ (*ori*, *str*, *con*). Descriptively, error rates were generally very low for all factor levels ($Median_{ori} = 0$, $IQR_{ori} = 0$; $Median_{con} = 0$, $IQR_{con} = 0.02$; $Median_{str} = 0$, $IQR_{str} = 0$) and did not differ significantly ($F_{(2,70)} = 0.113$, $p = 0.894$, $\eta p^2 = 0.003$). We also did not find a significant main effect of VESRION$_{FMRI}$ on reaction times during the fMRI session ($F_{(2,70)} = 0.983$, $p = 0.380$, $\eta p^2 = 0.027$). These findings indicate that the participants' attention to (and understanding of) the stories was very good and not deteriorated by the switch of single action steps or objects. Importantly, this shows that any (neural) differences between these conditions are not simply based on confusion or generally elevated cognitive effort in respect to modifications.

3.3 fMRI results

3.3.1 Whole brain analysis

First, we tested whether structure and content modifications elicit common brain activation patterns compared to original episodes. The conjunction of the whole brain contrasts *str* > *ori* and *con* > *ori* did not yield significant activation after correction for multiple comparisons. Without correction and at a threshold of $p < 0.001$, there was significant activation in the opercular part of the right IFG which corresponded to BA 44 ($x = 51$, $y = 17$, $z = 23$; 48 voxels).

To identify which brain regions were active for structure modifications, we calculated the whole brain contrast *str* > *ori*. This contrast did not yield significant activation after correction for multiple comparisons. Without correction and at a threshold of $p < 0.001$, activation clusters were detected in the right hemisphere in SFS, PMd and the opercular part of the IFG (BA 44) (Figure 5 and Table 1).

To detect brain regions which respond to content modifications, we calculated the whole brain contrast *con* > *ori*. Significant bilateral activation was found in pre-SMA and the middle IPS. Additionally, there were activation clusters in left and right inferior frontal sulcus (IFS) comprising the triangular part of the IFG (BA 45) and the inferior frontal junction (IFJ). Further, there was significant bilateral activation in the anterior insula and in the OTC, including bilateral FG and extending into parahippocampal gyrus (PHG) in the right hemisphere (Figure 5 and Table 1).

- insert **Figure 5** here –

- insert **Table 1** here -

3.3.2 ROI analysis of BA 44 and BA 45

The rmANOVA with the factors HEMISPHERE (*left*, *right*), ROI (*BA 44*, *BA 45*), and CONTRAST (*str* > *ori*, *con* > *ori*) revealed a significant main effect of HEMISPHERE ($F_{(1,35)} = 18.458$, $p < 0.001$, $\eta p^2 = 0.345$), driven by higher beta values in the right ($M_{right} = 0.359$, $SE_{right} = 0.092$) than in the left hemisphere ($M_{left} = 0.096$, $SE_{left} = 0.060$). Further, we found a significant main effect of CONTRAST ($F_{(1,35)} = 11.317$, $p = 0.002$, $\eta p^2 = 0.244$), as beta values extracted from the *con* > *ori* contrast ($M_{con} = 0.331$, $SE_{con} = 0.082$) were generally higher as those extracted from the *str* > *ori* contrast ($M_{str} = 0.124$, $SE_{str} = 0.073$). While we found no significant effect of ROI ($F_{(1,35)} = 0.027$, $p = 0.871$, $\eta p^2 = 0.001$), there was a significant interaction of the factors ROI and CONTRAST ($F_{(1,35)} = 50.405$, $p < 0.001$, $\eta p^2 = 0.590$). Post-hoc paired *t*-tests revealed that beta values were significantly lower in BA 45 for structure ($M_{str_45} = 0.074$, $SE_{str_45} = 0.075$) than for content modifications ($M_{con_45} = 0.387$, $SE_{con_45} = 0.084$; $t_{(35)} = -4.963$, $p < 0.001$). Further, we found that beta values for structure modifications were significantly higher in BA 44 ($M_{str_44} = 0.173$, $SE_{str_44} = 0.078$) than in BA 45 ($t_{(35)} = 2.349$, $p = 0.049$), while the opposite was the case for content modifications ($M_{con_44} = 0.275$, $SE_{con_44} = 0.086$; $t_{(35)} = -2.572$, $p = 0.029$).

4 DISCUSSION

Episodic memories allow us to mentally relive our personal past. Against what we might believe, they are not exact imprints of previous experiences, as they are prone to later modification, potentially triggered by prediction errors. In the present study, we investigated brain and behavioral responses to retrieval cue induced prediction errors based on two properties which characterize episodic memories, namely structure and content information. While brain responses were largely as expected for content prediction errors, only subtle activation was detected following structure prediction errors. A post-fMRI memory test revealed an increased tendency to accept modified episodes when the same had been presented during the fMRI scan. Together, findings provide first evidence that different types of mnemonic prediction errors are processed differently by the brain and may contribute to memory changes.

4.1 Influence of episodic modifications on post-fMRI memory performance

As expected, episodes that had been presented in a modified version during the fMRI experiment were generally remembered worse, as reflected by lower discrimination indexes (P_r). More specifically, modified videos were more often perceived as original stories thereafter, as evidenced by significantly more acceptive ratings and increased false alarm rates. Others have reported that mnemonic prediction errors lead to an intrusion of new information into an established memory repertoire (Long et al., 2016; Sinclair & Barense, 2018).

Previous studies report that the inclusion of new information into an old memory is not necessarily accompanied by a deterioration of the original memory (St. Jacques et al., 2013; Sinclair & Barense, 2018). Also in our study, the acceptance of original videos was only descriptively reduced after encounters of alternative versions, while hit rates indicating detection of original episodes were at ceiling. In contrast, Kim et al., showed that prediction violations led to decreased recognition of original memory content. In their study, original memories were encoded by showing the to-be-predicted sequence only once (Kim et al., 2014), whereas participants in our study had experienced and actively encoded each story multiple times. Thus, it seems plausible that memory traces were particularly stable in the current paradigm, such that a greater number of repeated alternative evidences would have been required to significantly reduce memory performance for the initially encoded episodes.

Participants took longer to correctly classify unmodified videos when they had previously experienced viewed structurally modified videos. Longer response times in cued-memory paradigms are interpreted as indicative of increased difficulty of retrieval due to higher cognitive processing demands (Larsen & Plunkett, 1987; Noppeney & Price, 2004). Thus, it is likely that it became more difficult to differentiate between alternative competing versions of episodes when versions diverging from the original experience had been already encountered in the fMRI session.

Taken together, the behavioral findings suggest that structure as well as content modifications during cueing of episodic retrieval influenced subsequent memory for these episodes. Our findings corroborate the observation that mnemonic prediction errors can trigger episodic memory modification (Sinclair & Barense, 2019). While changes in memory may sometimes reflect the updating of the original memory trace (Sinclair & Barense, 2018; 2019), our findings suggest that original memory traces remained largely intact as original episodes were still correctly recognized. Our results rather indicate that unexpected new information led to the storing of an additional, alternative version of the original episode.

What remains unclear is how exactly memory traces were influenced by our intervention. For example, it has been discussed that memory modification can result from an interference of old and new memory traces (Sederberg et al., 2011; Klingmüller et al., 2017; Sinclair & Barense, 2018; 2019) or from source confusion (Hekkanen & McEvoy, 2002), which could both explain our findings. On the other hand, participants remembered correctly that untrained videos had not been part of the original episode repertoire even though they had repeatedly been presented during the fMRI experiment as well. This finding suggests that our post-fMRI memory test results are probably not a product of source confusion in its simplest form.

Another interesting behavioral finding was that participants generally had a strong tendency to accept structurally modified versions as originals in the memory test. Although both types of modifications resulted from a single change in the story (cf. Supplementary Table 1), it is likely that structural modifications were generally less salient than content modifications. This would be matched by the fact that after the fMRI session, two-thirds of the participants reported noticing at least one object swap, while less than a quarter of them had noticed a change in the sequence of action steps. Recently, it was reported that memory performance following prediction errors differed depending on whether changes were detected (and remembered) by participants or not (Wahlheim & Zacks, 2019). While undetected changes led to reduced memory performance, detected changes had the opposite effect. Depending on contextual factors, prediction errors can even improve subsequent memory (Smith et al., 2013; Greve et al., 2017). Although our results suggest that structural changes were rarely noticed, more frequent detection in the case of content modification did not lead to enhanced, but on the contrary, also to worse memory performance. Thus, it appears that the effect of prediction violation on memory is subject to complex dynamics. Specifically, it has been put forward that strong prediction errors lead to better memory by promoting learning of new associations (Greve et al., 2017). Since weak prediction errors were also found to enhance memory performance (Greve et al., 2019), it may turn out that the strength of prediction errors need to overcome a certain threshold to initiate memory modification (Exton-McGuinness et al., 2015) or modulates memory even in a U-shaped fashion (Greve et al., 2019). The effect of the quality of prediction errors in episodic retrieval and the effect of their quantity need to be investigated simultaneously in future studies.

4.2 Neural responses to different episodic modifications

New content and structure information of the episodic cue was expected to draw on distinct brain areas, but also to share some common activation in MFC, areas involved in action observation, and the hippocampal formation. However, only without correction for multiple comparisons, common activation was detected in right IFG, which reflects our previous finding that IFG increases for inconsistent or highly informative detail in observed actions (Wurm & Schubotz, 2012; Hrkać et al., 2015; El-Sourani et al., 2019). Here, it is not possible to distinguish whether nonsignificance was due to a large functional disparity between content and structural information in episodic memories, or simply to an overly subtle brain response to violated structural predictions.

On the one hand, we had expected that structure modifications lead to activation in brain regions involved in the temporal organization of episodes. We did not find significant activation patterns after correction for multiple comparisons. Without such correction, we detected activation in a hypothesized area, right PMd/SFS. The activity of a region comprising PMd/SFS is related to linking successive action steps (Stadler et al., 2011; Schubotz et al., 2012; Hrkać et al., 2014; Pomp et al., 2021), and could contribute to updating the current event or action model with respect to each next segment (Kurby & Zacks, 2008; Tamber-Rosenau et al., 2011; Schubotz et al., 2012; Pomp et al., 2021).

On the other hand, we found content modifications to specifically recruit pIPS and OTC, including FG, which were hypothesized on the basis of their role for processing of object properties in the context of actions (Grill-Spector et al., 2001; Reber et al., 2005; Wiggett & Downing, 2011; Lingnau & Downing, 2015; El-Sourani et al., 2019). More specifically, pIPS encodes basic visual features of graspable objects (Creem-Regehr, 2009; Mruczek et al., 2013), reflecting the interaction with toy objects in our paradigm. Additionally, content modifications elicited activity in the hippocampal formation (right PHG). In general, hippocampus and PHG are important in learning contexts (Aguirre et al., 1996; O'Reilly & Rudy, 2000; Davachi & Wagner, 2002; Köhler et al., 2002) and there is evidence that the PHG is involved in processing of competing memories (Kuhl et al., 2012). Moreover, the hippocampal formation is believed to generate mismatch signals when predictions do not fit perceptual inputs (Kumaran & Maguire, 2007; Duncan et al., 2009; Long et al., 2016). Because our post-fMRI memory test data imply that content changes were more salient than structural changes, one could speculate that the discrepancy between prediction and perception was strong enough to be reflected in the activation of the hippocampal formation for content modifications. This would also fit with our finding that content modifications triggered activity in the anterior insula, which is relevant for error detection (Klein et al., 2013) and salience processing (Menon & Uddin, 2010).

Consistent with our conjunction analysis, the ROI analysis showed stronger activity in the right hemisphere in different IFG regions. Activation in BA 45 was stronger for content changes than for structure changes. These results confirm the assumption that BA 45 is more responsive to content than to structural prediction errors (Badre & Wagner, 2007; Heim et al., 2009). We take this as a first indication that within IFG, different subregions might process different types of episodic prediction errors. However, our study cannot rule out the possibility that this effect is merely due to quantitative (rather than qualitative) differences between structural and content prediction errors, as structure modifications might have elicited weaker prediction errors in our paradigm.

4.3 Limitations and implications for future research

One factor that may limit the generalizability of our findings is that, for practical reasons, only women participated in the study. However, because the processing of episodic memory in the brain

seems to be broadly similar between women and men (Nyberg et al., 2000), we are confident that our findings are applicable to a more general population. In the future, our paradigm could be adapted to circumvent such practical limitations, for example, by applying virtual reality techniques so that encoding could be detached from the true physical appearance of participants' hands.

Second, we used new content and structure information which contrasted details of the encoded episodes to elicit prediction errors. Even though these interventions can also be interpreted as contextual and associative novelty, respectively, the unexpected new input within the familiar context will give rise to mnemonic prediction errors according to the predictive coding framework (see Reichardt et al., 2020 for a review). It would be valuable to find a way to keep novelty constant and have participants make active predictions which are then either violated or not.

Finally, another limitation of the current study is that brain responses to structure modifications were much weaker than those to content modifications, which may be explained by lower salience. This issue will be testable if stronger structural breaches of expectation are employed in episodic retrieval.

5 CONCLUSION

When recalling episodes, our memory can change, for instance due to prediction errors in the re-activation process. Our results suggest that structural and content prediction errors in episode retrieval differ in their neural processing. The tendency to falsely classify modified episodes as originally experienced ones increased after experiencing repeated structural as well as content-related prediction errors. Accordingly, different types of prediction errors can confuse episodic memory and possibly lead to the emergence of alternative versions of one and the same memory trace. Our results and the new paradigm introduced here may provide a fruitful starting point for further research on the mutability of our episodic memory.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

SS, BJ, NE and RS contributed to conception and design of the study. SS and BJ collected the data. SS performed the statistical analysis. IT provided substantial help with the data analysis. SS wrote the first draft of the manuscript. RS wrote sections of the manuscript. SS created the figures. SC, OW and RS contributed with scientific support. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft) – project numbers 419037023, 419039274, 419037518. The funders had no role in study design, data collection, analysis and interpretation, decision to publish, or writing of the report.

ACKNOWLEDGMENTS

The authors thank Monika Mertens for her help during data collection. Furthermore, we thank Annika Garlichs, Christin Schwarzer, Helena Sydlik and Yuyi Xu for their assistance during the creation of stimulus material, data collection and data management. Lastly, we thank Jennifer Pomp, Lena Schliephake, Falko Mecklenbrauck and Nina Heins for advice regarding data analysis and the members of research unit FOR 2812 for valuable discussions. A preprint of this article has been posted on bioRxiv (Siestrup et al., 2021).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: https://osf.io/34knw/?view_only=3194f4ad2fa54a0fa47f4189b2aa1f77.

ETHICS STATEMENTS

The studies involving human participants were reviewed and approved by the Ethics Committee of the Department of Psychology, University of Münster. The participants provided their written informed consent to participate in this study.

REFERENCES

Aguirre, G. K., Detre, J. A., Alsop, D. C., and D'Esposito, M. (1996). The Parahippocampus Subserves Topographical Learning in Man. *Cereb. Cortex* 6, 823–829. doi: 10.1093/cercor/6.6.823

Badre, D., and Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia* 45, 2883–2901. doi: 10.1016/j.neuropsychologia.2007.06.015

Barto, A., Mirolli, M., and Baldassarre, G. (2013). Novelty or Surprise? *Front. Psychol.* 4, 1–15. doi: 10.3389/fpsyg.2013.00907

Bein, O., Duncan, K., and Davachi, L. (2020). Mnemonic prediction errors bias hippocampal states. *Nat. Commun.* 11. doi: 10.1038/s41467-020-17287-1

Bonferroni, C. E. (1936). *Teoria Statistica Delle Classi e Calcolo Delle Probabilità*, Vol. 8. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze, 3–62.

Brett, M., Anton, J. L., Valabregue, R., and Poline, J.-B. (2002). Region of interest analysis using an SPM toolbox [Abstract]. Presented at the 8th International Conferance on Functional Mapping of the Human Brain, June 2-6, 2002, Sendai, Japan. Available on CD-ROM in *Neuroimage* 16, Abstract 497.

Caspers, S., Zilles, K., Laird, A. R., and Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage* 50, 1148–1167. doi: 10.1016/j.neuroimage.2009.12.112

Cheng, S., Werning, M., and Suddendorf, T. (2016). Dissociating memory traces and scenario construction in mental time travel. *Neurosci. Biobehav. Rev*. 60, 82–89. doi: 10.1016/j.neubiorev.2015.11.011

Clerget, E., Winderickx, A., Fadiga, L., and Olivier, E. (2009). Role of Broca's area in encoding sequential human actions: A virtual lesion study. *Neuroreport* 20, 1496–1499. doi: 10.1097/WNR.0b013e3283329be8

Collins, A. M., and Quillian, M. R. (1969). Retrieval Time from Semantic Memory. *J. Verbal Learning Verbal Behav.* 8, 240–247. doi: 10.1016/S0022-5371(69)80069-1

Creem-Regehr, S. H. (2009). Sensory-motor and cognitive functions of the human posterior parietal cortex involved in manual actions. *Neurobiol. Learn. Mem.* 91, 166–171. doi: 10.1016/j.nlm.2008.10.004

Davachi, L., and Wagner, A. D. (2002). Hippocampal Contributions to Episodic Encoding: Insights From Relational and Item-Based Learning. *J. Neurophysiol*. 88, 982–990. doi: 10.1152/jn.2002.88.2.982

Duncan, K., Curtis, C., and Davachi, L. (2009). Distinct Memory Signatures in the Hippocampus: Intentional States Distinguish Match and Mismatch Enhancement Signals. *J. Neurosci.* 29, 131–139. doi: 10.1523/JNEUROSCI.2998-08.2009

El-Sourani, N., Trempler, I., Wurm, M. F., Fink, G. R., and Schubotz, R. I. (2019). Predictive Impact of Contextual Objects during Action Observation: Evidence from Functional Magnetic Resonance Imaging. *J. Cogn. Neurosci.* 32, 326–337. doi: 10.1162/jocn_a_01480

Euston, D. R., Gruber, A. J., and McNaughton, B. L. (2012). The Role of Medial Prefrontal Cortex in Memory and Decision Making. *Neuron* 76, 1057–1070. doi: 10.1016/j.neuron.2012.12.002

Exton-McGuinness, M. T. J., Lee, J. L. C., and Reichelt, A. C. (2015). Updating memories-The role of prediction errors in memory reconsolidation. *Behav. Brain Res.* 278, 375–384. doi: 10.1016/j.bbr.2014.10.011

Fernández, R. S., Boccia, M. M., and Pedreira, M. E. (2016). The fate of memory: Reconsolidation and the case of Prediction Error. *Neurosci. Biobehav. Rev.* 68, 423–441. doi: 10.1016/j.neubiorev.2016.06.004

Fiebach, C. J., and Schubotz, R. I. (2006). Dynamic anticipatory processing of hierarchical sequential events: A common role for Broca's area and ventral premotor cortex across domains? *Cortex* 42, 499–502. doi: 10.1016/S0010-9452(08)70386-1

Franz, E. A., Ford, S., and Werner, S. (2007). Brain and cognitive processes of imitation in bimanual situations: Making inferences about mirror neuron systems. *Brain Res.* 1145, 138–149. doi: 10.1016/j.brainres.2007.01.136

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.□P., Frith, C. D., and Frackowiak, R. S. J. (1994). Statistical Parametric Maps in Functional Imaging: A General Linear Approach. *Hum. Brain Mapp.* 2, 189–210. doi: 10.1002/hbm.460020402

Friston, K., and Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1211–1221. doi: 10.1098/rstb.2008.0300

Greve, A., Cooper, E., Kaula, A., Anderson, M. C., and Henson, R. (2017). Does prediction error drive one-shot declarative learning? *J. Mem. Lang.* 94, 149–165. doi: 10.1016/j.jml.2016.11.001

Greve, A., Cooper, E., Tibon, R., and Henson, R. N. (2019). Knowledge Is Power: Prior Knowledge Aids Memory for Both Congruent and Incongruent Events, but in Different Ways. *J. Exp. Psychol. Gen.* 148, 325–341. doi: 10.1037/xge0000498.supp

Griffiths, D., Dickinson, A., and Clayton, N. (1999). Episodic memory: What can animals remember about their past? *Trends Cogn. Sci.* 3, 74–80. doi: 10.1016/S1364-6613(98)01272-8

Grill-Spector, K., Kourtzi, Z., and Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Res.* 41, 1409–1422. doi: 10.1016/S0042-6989(01)00073-6

Heim, S., Eickhoff, S. B., and Amunts, K. (2009). Different roles of cytoarchitectonic BA 44 and BA 45 in phonological and semantic verbal fluency as revealed by dynamic causal modelling. *Neuroimage*, 48, 616–624. doi: 10.1016/j.neuroimage.2009.06.044

Hekkanen, S. T., and McEvoy, C. (2002). False memories and source-monitoring problems: Criterion differences. *Appl. Cogn. Psychol.* 16, 73–85. doi: 10.1002/acp.753

Horner, A. J., and Doeller, C. F. (2017). Plasticity of hippocampal memories in humans. *Curr. Opin. Neurobiol.* 43, 102–109. doi: 10.1016/j.conb.2017.02.004

Hrkać, M., Wurm, M. F., and Schubotz, R. I. (2014). Action Observers Implicitly Expect Actors to Act Goal-Coherently, Even if they Do Not: An fMRI Study. *Hum. Brain Mapp.* 35, 2178–2190. doi: 10.1002/hbm.22319

Hrkać, M., Wurm, M. F., Kühn, A. B., and Schubotz, R. I. (2015). Objects Mediate Goal Integration in Ventrolateral Prefrontal Cortex during Action Observation. *PLoS One* 10, 1–12. doi: 10.1371/journal.pone.0134316

Jainta, B., Siestrup, S., El-Sourani, N., Trempler I., Wurm, M., Werning, M., Cheng, S., Schubotz, R. I. (2021). Seeing what I did (not): Cerebral and behavioral effects of agency and perspective on episodic memory re-activation. *OSF Preprints* [Preprint]. doi: 10.31219/osf.io/rm648

Kim, G., Lewis-Peacock, J. A., Norman, K. A., and Turk-Browne, N. B. (2014). Pruning of memories by context-based prediction error. *Proc. Natl. Acad. Sci. U. S. A.* 111, 8997–9002. doi: 10.1073/pnas.1319438111

Klein, T. A., Ullsperger, M., and Danielmeier, C. (2013). Error awareness and the insula: links to neurological and psychiatric diseases. *Front. Hum. Neurosci.* 7, 1–14. doi: 10.3389/fnhum.2013.00014

Klingmüller, A., Caplan, J. B., and Sommer, T. (2017). Intrusions in episodic memory: reconsolidation or interference. *Learn. Mem.* 24, 216–224. doi: 10.1101/lm.045047.117

Köhler, S., Crane, J., and Milner, B. (2002). Differential Contributions of the Parahippocampal Place Area and the Anterior Hippocampus to Human Memory for Scenes. *Hippocampus* 12, 718–723. doi: 10.1002/hipo.10077

Kuhl, B. A., Bainbridge, W. A., and Chun, M. M. (2012). Neural Reactivation Reveals Mechanisms for Updating Memory. *J. Neurosci.* 32, 3453–3461. doi: 10.1523/JNEUROSCI.5846-11.2012

Kumaran, D., and Maguire, E. A. (2007). Match–Mismatch Processes Underlie Human Hippocampal Responses to Associative Novelty. *J. Neurosci.* 27, 8517–8524. doi: 10.1523/JNEUROSCI.1677-07.2007

Kurby, C. A., and Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends Cogn. Sci.* 12, 72–79. doi: 10.1016/j.tics.2007.11.004

Larsen, S. F., and Plunkett, K. (1987). Remembering experienced and reported events. *Appl. Cogn. Psychol.* 1, 15–26. doi: 10.1002/acp.2350010104

Lee, J. L. C., Nader, K., and Schiller, D. (2017). An Update on Memory Reconsolidation Updating. *Trends Cogn. Sci.* 21, 531–545. doi: 10.1016/j.tics.2017.04.006

Lingnau, A., and Downing, P. E. (2015). The lateral occipitotemporal cortex in action. *Trends Cogn. Sci.* 19, 268–277. doi: 10.1016/j.tics.2015.03.006

Long, N. M., Lee, H., and Kuhl, B. A. (2016). Hippocampal Mismatch Signals Are Modulated by the Strength of Neural Predictions and Their Similarity to Outcomes. *J. Neurosci.* 36, 12677–12687. doi: 10.1523/JNEUROSCI.1850-16.2016

Maguire, E. A., Intraub, H., and Mullally, S. L. (2016). Scenes, Spaces, and Memory Traces: What Does the Hippocampus Do? *Neuroscientist*, 22, 432–439. doi: 10.1177/1073858415600389

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

Menon, V., and Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214, 655–667. doi: 10.1007/s00429-010-0262-0

Mruczek, R. E. B., von Loga, I. S., and Kastner, S. (2013). The representation of tool and non-tool object information in the human intraparietal sulcus. *J. Neurophysiol.* 109, 2883–2896. doi: 10.1152/jn.00658.2012

Nader, K. (2015). Reconsolidation and the Dynamic Nature of Memory. *Nov. Mech. Mem.* 7, 1–20. doi: 10.1007/978-3-319-24364-1_1

Nader, K., and Einarsson, E. Ö. (2010). Memory reconsolidation: An update. *Ann. N. Y. Acad. Sci.* 1191, 27–41. doi: 10.1111/j.1749-6632.2010.05443.x

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

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

Nyberg, L., Habib, R., and Herlitz, A. (2000). Brain activation during episodic memory retrieval: Sex differences. Acta Psychol. (Amst). 105, 181–194. doi: 10.1016/s0001-6918(00)00060-3

O'Reilly, R. C., and Rudy, J. W. (2000). Computational Principles of Learning in the Neocortex and Hippocampus. *Hippocampus*, 10, 389–397. doi: 10.1002/1098-1063(2000)10:4<389::AID-HIPO5>3.0.CO;2-P

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

Peters, G. J., David, C. N., Marcus, M. D., and Smith, D. M. (2013). The medial prefrontal cortex is critical for memory retrieval and resolving interference. *Learn. Mem.* 20, 201–209. doi: 10.1101/lm.029249.112

Pomp, J., Heins, N., Trempler, I., Kulvicius, T., Tamosiunaite, M., Mecklenbrauck, F., et al. (2021). Touching events predict human action segmentation in brain and behavior. *Neuroimage* 243, 1–15. doi: 10.1016/j.neuroimage.2021.118534

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reber, P. J., Gitelman, D. R., Parrish, T. B., and Mesulam, M. M. (2005). Priming Effects in the Fusiform Gyrus: Changes in Neural Activity beyond the Second Presentation. *Cereb. Cortex* 15, 787–795. doi: 10.1093/cercor/bhh179

Reichardt, R., Polner, B., and Simor, P. (2020). Novelty Manipulations, Memory Performance, and Predictive Coding: the Role of Unexpectedness. *Front. Hum. Neurosci.* 14, 1–11. doi: 10.3389/fnhum.2020.00152

Schiffer, A. M., Ahlheim, C., Ulrichs, K., and Schubotz, R. I. (2013). Neural Changes when Actions Change: Adaptation of Strong and Weak Expectations. *Hum. Brain Mapp.* 34, 1713–1727. doi: 10.1002/hbm.22023

Schiffer, A. M., Ahlheim, C., Wurm, M. F., and Schubotz, R. I. (2012). Surprised at all the entropy: Hippocampal, caudate and midbrain contributions to learning from prediction errors. *PLoS One* 7, 1–11. doi: 10.1371/journal.pone.0036445

Schubotz, R. I. (2004). Human premotor cortex: Beyond motor performance. MPI Series in Human Cognitive and Brain Sciences, vol. 50. Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig.

Schubotz, R. I. (2015). “Prediction and expectation,” in Brain Mapping: An Encyclopedic Reference (Vol. 3), ed. A. W. Toga (Cambridge, MA: Academic Press: Elsevier), 295–302. doi: 10.1016/B978-0-12-397025-1.00205-0

Schubotz, R. I., Korb, F. M., Schiffer, A. M., Stadler, W., and von Cramon, D. Y. (2012). The fraction of an action is more than a movement: Neural signatures of event segmentation in fMRI. *Neuroimage* 61, 1195–1205. doi: 10.1016/j.neuroimage.2012.04.008

Schubotz, R. I., Wurm, M. F., Wittmann, M. K., and von Cramon, D. Y. (2014). Objects tell us what action we can expect: Dissociating brain areas for retrieval and exploitation of action knowledge during action observation in fMRI. *Front. Psychol.* 5, 1–15. doi: 10.3389/fpsyg.2014.00636

Scully, I. D., Napper, L. E., and Hupbach, A. (2017). Does reactivation trigger episodic memory change? A meta-analysis. *Neurobiol. Learn. Mem.* 142, 99–107. doi: 10.1016/j.nlm.2016.12.012

Sederberg, P. B., Gershman, S. J., Polyn, S. M., and Norman, K. A. (2011). Human memory reconsolidation can be explained using the temporal context model. *Psychon. Bull. Rev*. 18, 455–468. doi: 10.3758/s13423-011-0086-9

Siestrup, S., Jainta, B., El-Sourani, N., Trempler I., Wolf, O. T., Cheng, S., & Schubotz, R. I. Data from: What happened when? Brain and behavioral responses to modified structure and content in episodic cueing. *OSF*. (2021) https://osf.io/34knw/?view_only=3194f4ad2fa54a0fa47f4189b2aa1f77

Siestrup, S., Jainta, B., El-Sourani, N., Trempler I., Wolf, O. T., Cheng, S., & Schubotz, R. I. (2021). What happened when? Brain and behavioral responses to modified structure and content in episodic cueing. *bioRxiv* [Preprint]. doi: 10.1101/2021.11.10.467871

Sinclair, A. H., and Barense, M. D. (2018). Surprise and destabilize: Prediction error influences episodic memory reconsolidation. *Learn. Mem.* 25, 369–381. doi: 10.1101/lm.046912.117

Sinclair, A. H., and Barense, M. D. (2019). Prediction Error and Memory Reactivation: How Incomplete Reminders Drive Reconsolidation. *Trends Neurosci.* 42, 727–739. doi: 10.1016/j.tins.2019.08.007

Smith, T. A., Hasinski, A. E., and Sederberg, P. B. (2013). The context repetition effect: Predicted events are remembered better, even when they don't happen. *J. Exp. Psychol. Gen.* 142, 1298–1308. doi: 10.1037/a0034067

Snodgrass, J. G., and Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *J. Exp. Psychol. Gen.* 117, 34–50. doi: 10.1037/0096-3445.117.1.34

St. Jacques, P. L., Olm, C., Schacter, D. L. (2013). Neural mechanisms of 987 reactivation-Induced updating that enhance and distort memory. *Proc. Natl. Acad. Sci. U. S. A.* 110, 19671–19678. doi: 10.1073/pnas.1319630110

Stachenfeld, K. L., Botvinick, M. M., and Gershman, S. J. (2017). The hippocampus as a predictive map. *Nat. Neurosci.* 20, 1643–1653. doi: 10.1038/nn.4650

Stadler, W., Schubotz, R. I., von Cramon, D. Y., Springer, A., Graf, M., and Prinz, W. (2011). Predicting and Memorizing Observed Action: Differential Premotor Cortex Involvement. *Hum. Brain Mapp.* 32, 677–687. doi: 10.1002/hbm.20949

Tamber-Rosenau, B. J., Esterman, M., Chiu, Y. C., and Yantis, S. (2011). Cortical Mechanisms of Cognitive Control for Shifting Attention in Vision and Working Memory. *J. Cogn. Neurosci.* 23, 2905–2919. doi: 10.1162/jocn.2011.21608

Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annu. Rev. Psychol.* 53, 1–25. doi: 0084-6570/02/0201-0001

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273-289. doi: 10.1006/nimg.2001.0978

Wahlheim, C. N., and Zacks, J. M. (2019). Memory guides the processing of event changes for older and younger adults. *J. Exp. Psychol. Gen.* 148, 30–50. doi: 10.1037/xge0000458

Wiggett, A. J., and Downing, P. E. (2011). Representation of Action in Occipito-temporal Cortex. *J. Cogn. Neurosci.* 23, 1765–1780. doi: 10.1162/jocn.2010.21552

Wobbrock, J. O., Findlater, L., Gergle, D., and Higgins, J. J. (2011). The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures. CHI '11: CHI Conference on Human Factors in Computing Systems, 143–146. doi: 10.1145/1978942.1978963

Worsley, K. J., and Friston, K. J. (1995). Analysis of fMRI Time-Series Revesited - Again. *Neuroimage* 2, 173–181. doi: 10.1006/nimg.1995.1023

Wurm, M. F., and Schubotz, R. I. (2018). The role of the temporoparietal junction (TPJ) in action observation: Agent detection rather than visuospatial transformation. *Neuroimage*, 165, 48–55. doi: 10.1016/j.neuroimage.2017.09.064

Wurm, M. F., Schubotz, R. I. (2012). Squeezing lemons in the bathroom: Contextual information modulates action recognition. *Neuroimage* 59, 1551–1559. doi: 10.1016/j.neuroimage.2011.08.038

TABLES

Table 1. Peak activations from second-level whole-brain analyses of modification effects.

Localization	H	Cluster extent	MNI Coordinates			*t*-value
			x	y	z	
str* > *ori (uncorrected at $p < 0.001$)						
Superior frontal sulcus	R	15	27	23	59	3.72
Dorsal premotor cortex	R	52	39	2	56	4.80
Inferior frontal gyrus (pars opercularis, BA 44)	R	76	54	14	20	4.67
con* > *ori (FDR-corrected at $p < 0.05$)						
Pre-supplementary motor area	L	49	-6	14	53	4.16
	R	64	9	17	47	4.10
Intraparietal sulcus	R	292	30	-61	50	6.23
	L	244	-27	-64	38	5.47
Inferior frontal junction	L	321	-39	2	32	6.43
Inferior frontal sulcus, extending into inferior frontal gyrus (pars triangularis, BA 45)	L		-42	20	26	5.54
	R	508	45	29	20	7.04
Anterior insula	R	132	33	20	-1	5.57
	L	138	-30	20	-4	5.68
Occipitotemporal cortex	L	239	-45	-55	-13	7.55
Fusiform gyrus	L		-33	-46	-16	4.39
Occipitotemporal cortex, including fusiform gyrus and extending into parahippocampal gyrus	R	266	33	-52	-13	5.12

Note: H = Hemisphere, MNI = Montreal Neurological Institute, L = Left, R = Right, *str* = structure modification, *con* = content modification, *ori* = original.

FIGURE LEGENDS

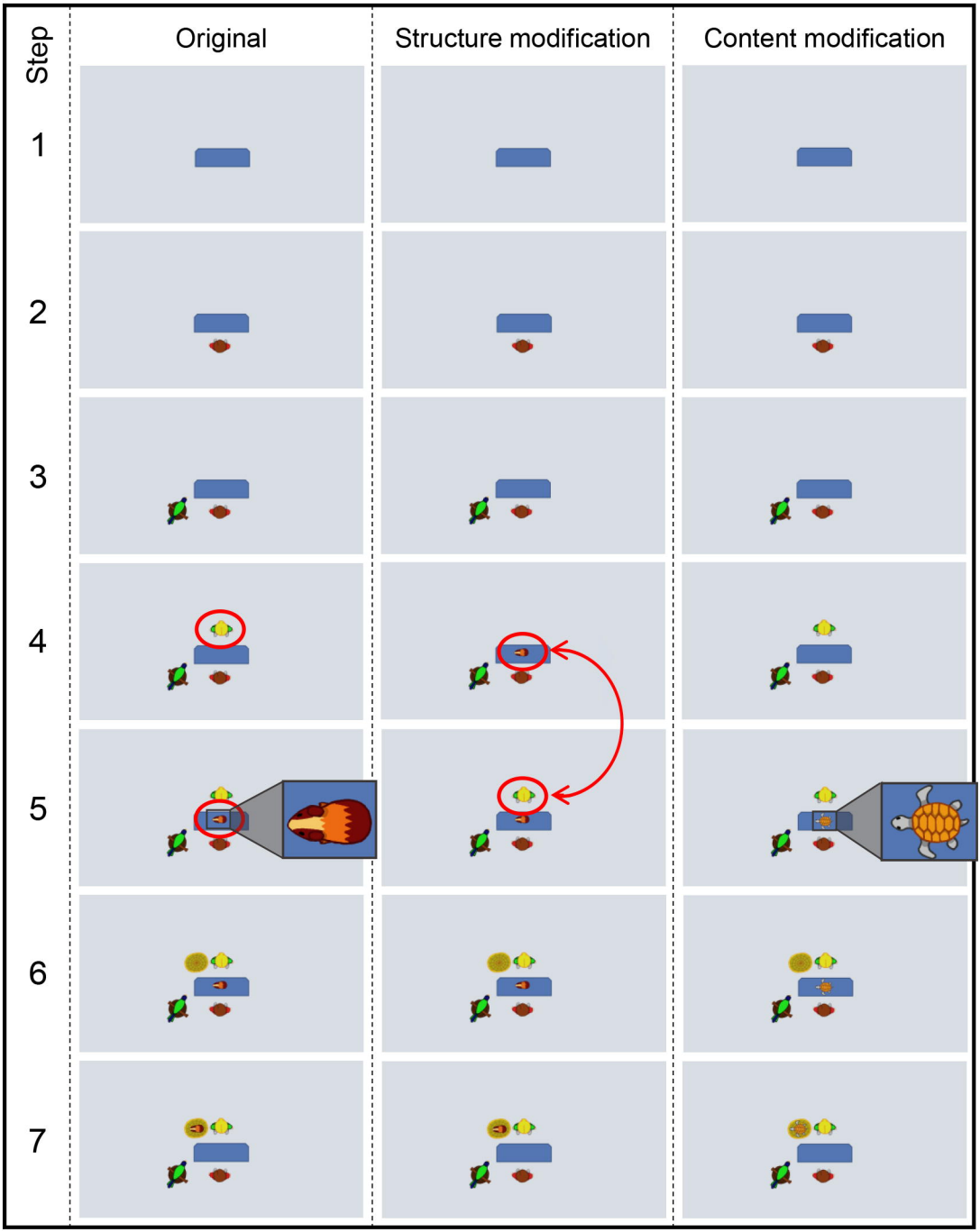
Figure 1. Example of an original episode and its modified versions, shown by the sequence of the main event steps. Twenty-four stories existed in three different versions each: an original, a structure modification, and a content modification. For the structure modifications, two adjacent action steps were switched compared to the original. In this example, the original shows the blonde woman join the scene before the guinea pig is positioned on the sales counter; in the structure modification, the guinea pig appears before the blonde woman (red circles). For the content modifications, an object was exchanged compared to the original (here: tortoise instead of guinea pig on the sales counter in step 5). Note that in the fMRI experiment, each participant was only presented with one of the three versions of a story. We do not reproduce photos of our stimulus material because it is copyrighted material (PLAYMOBIL® figures); instead, we provide schematic images.

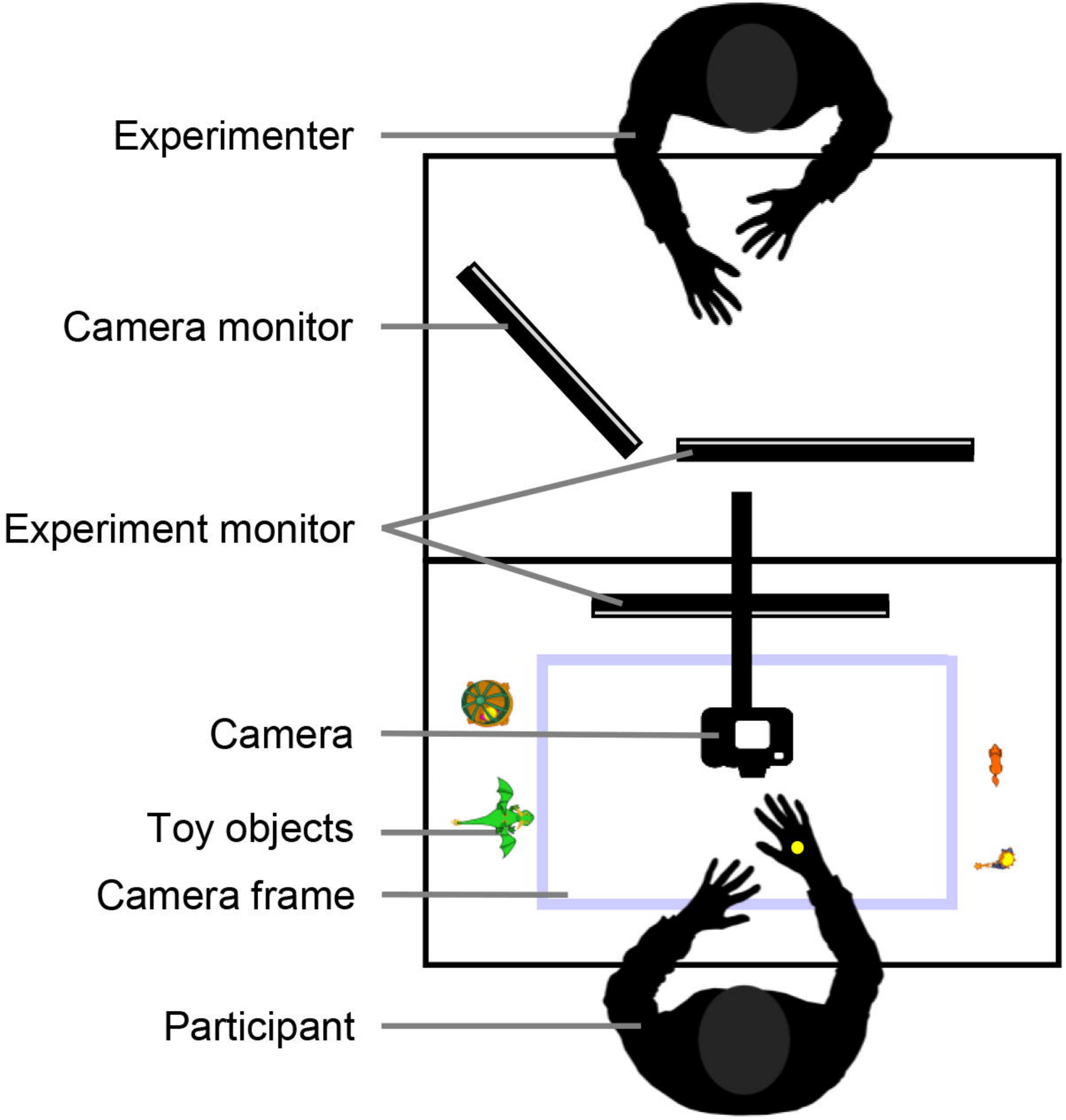
Figure 2. Training setup. During training participants observed and imitated toy stories, while sitting at the filming setup. Their performance was monitored by the experimenter.

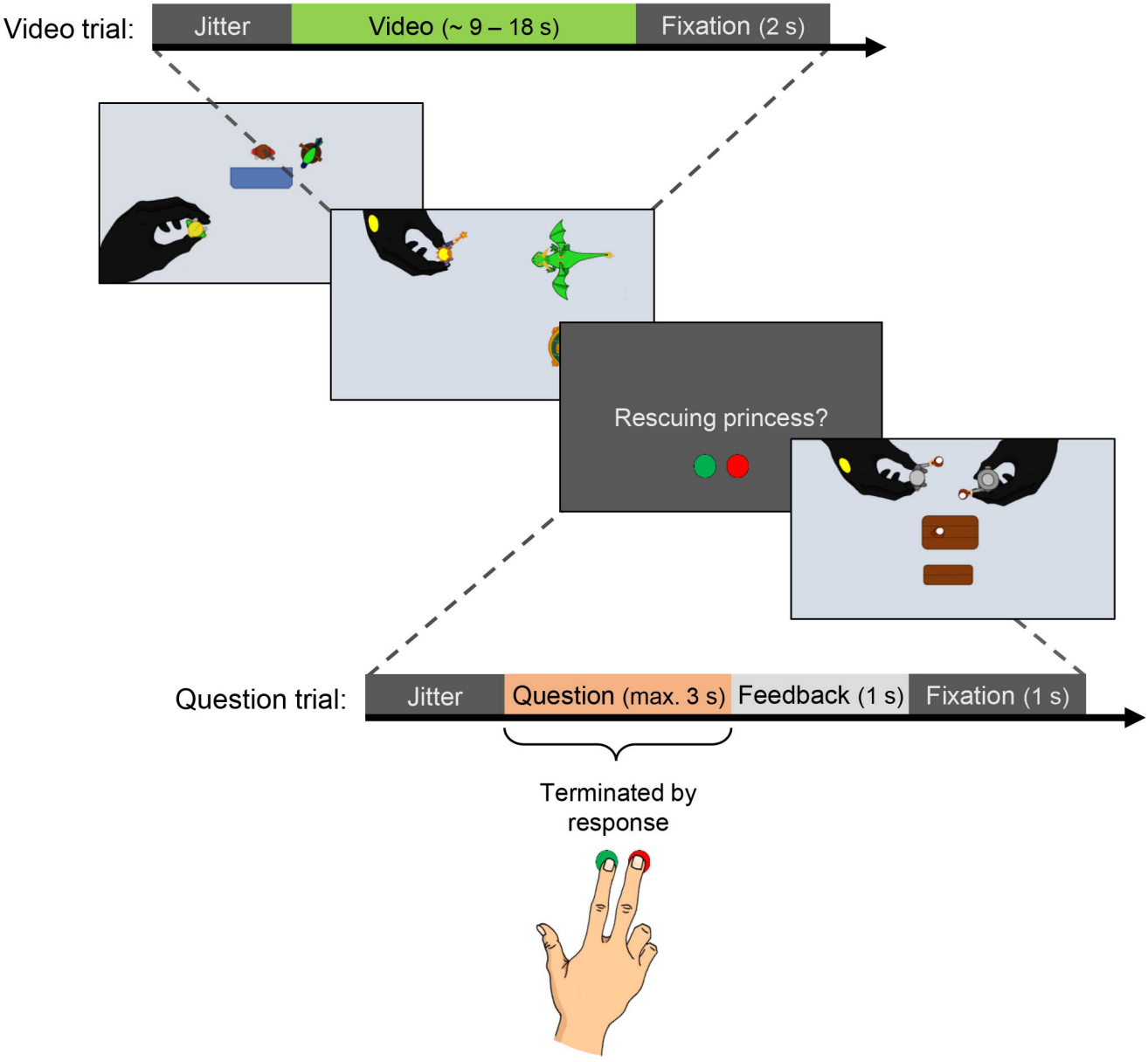
Figure 3. Schematic depiction of task during fMRI session. Video trials consisted of a variable jitter (0, 0.5, 1, or 1.5 s of fixation), a video showing a toy story (ca. 9 – 18 s) and a 2 s interstimulus interval (fixation). Question trials included a variable jitter, a question regarding the story shown in the preceding video (maximally 3 s long or terminated by response), and a 2 s interstimulus interval. The interstimulus interval after question trials was divided into 1 s feedback (“correct”, “incorrect”, “too late”) and 1 s fixation. Aside from the question, it was depicted which button should be pressed to accept (left, green) or reject (right, red) the description.

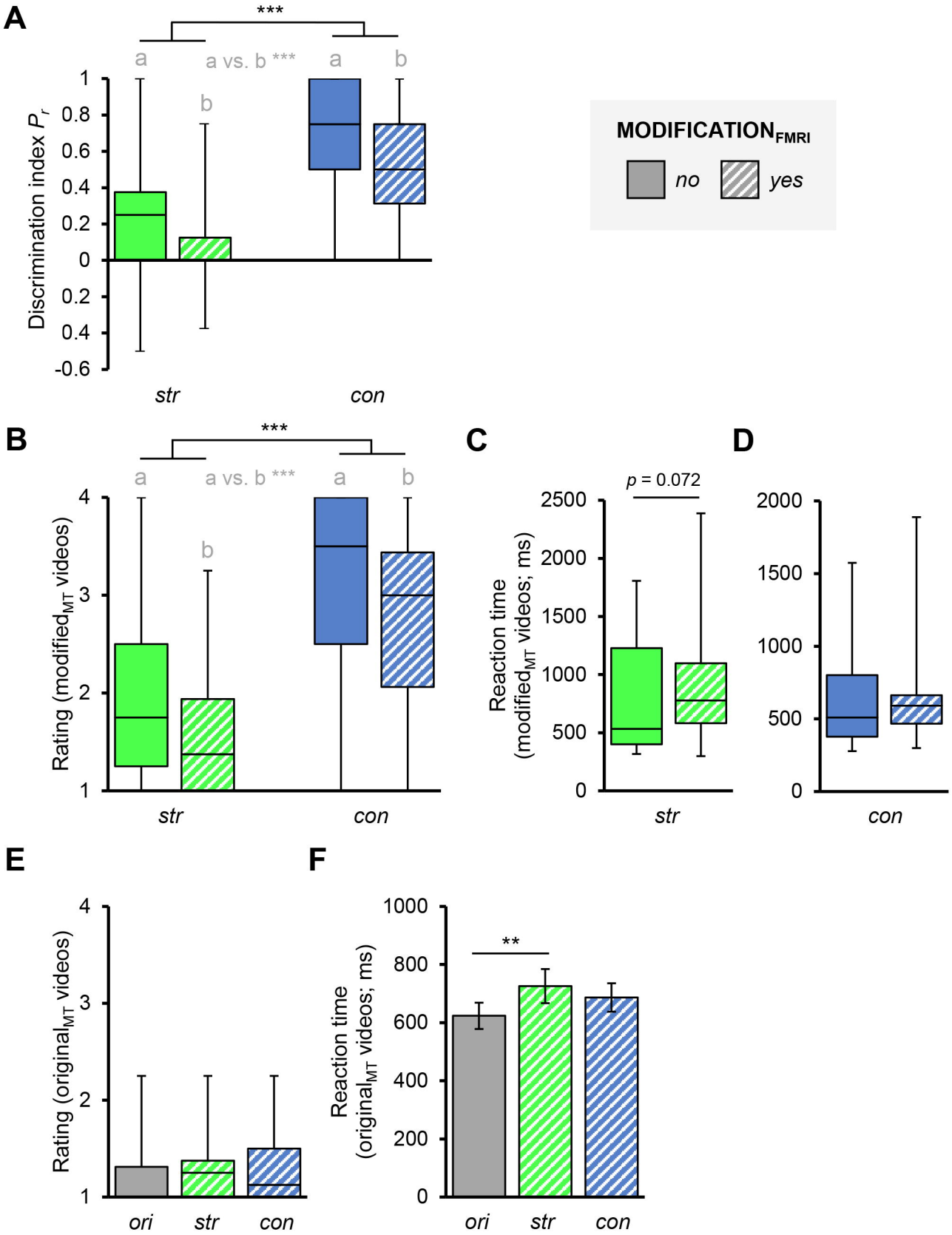
Figure 4: Behavioral results from post-fMRI memory test. For $\text{modified}_{\text{MT}}$ and $\text{original}_{\text{MT}}$ videos, participants rated whether it showed an originally encoded episode on a Likert scale including 1 (yes), 2 (rather yes), 3 (rather no) and 4 (no). **(A)** Discrimination index P_r from post-fMRI memory test. $\text{Original}_{\text{MT}}$ videos were the targets while $\text{modified}_{\text{MT}}$ videos were distractors. Statistics: non-parametric rmANOVA with the factors $\text{MODIFICATION}_{\text{FMRI}}$ (*yes, no*) and $\text{VERSION}_{\text{MT}}$ (*str, con*), $n = 35$. **(B)** Ratings for $\text{modified}_{\text{MT}}$ videos. Statistics: non-parametric rmANOVA based on aligned rank data with the factors $\text{MODIFICATION}_{\text{FMRI}}$ (*yes, no*) and $\text{VERSION}_{\text{MT}}$ (*str, con*), $n = 35$. **(C)** Reaction times for structure modified videos, $n = 16$. Statistics: Wilcoxon signed rank tests. **(D)** Reaction times for content modified videos, $n = 28$. **(E)** Ratings for $\text{original}_{\text{MT}}$ videos. Statistics: non-parametric rmANOVA with the factor $\text{VERSION}_{\text{FMRT}}$ (*ori, str, con*), $n = 35$. **(F)** Reaction times for $\text{original}_{\text{MT}}$ videos. Statistics: rmANOVA with the factor $\text{VERSION}_{\text{FMRT}}$ (*ori, str, con*), $n = 35$, post-hoc pairwise comparisons: paired *t*-tests, *p*-values were corrected for multiple comparisons. Box plots show medians, quartiles, minima and maxima. Bar plots show means and standard errors. ** = $p < 0.01$, *** = $p < 0.001$. *ori* = original, *str* = structure modification, *con* = content modification. a vs. b indicates the main effect of $\text{MODIFICATION}_{\text{FMRI}}$.

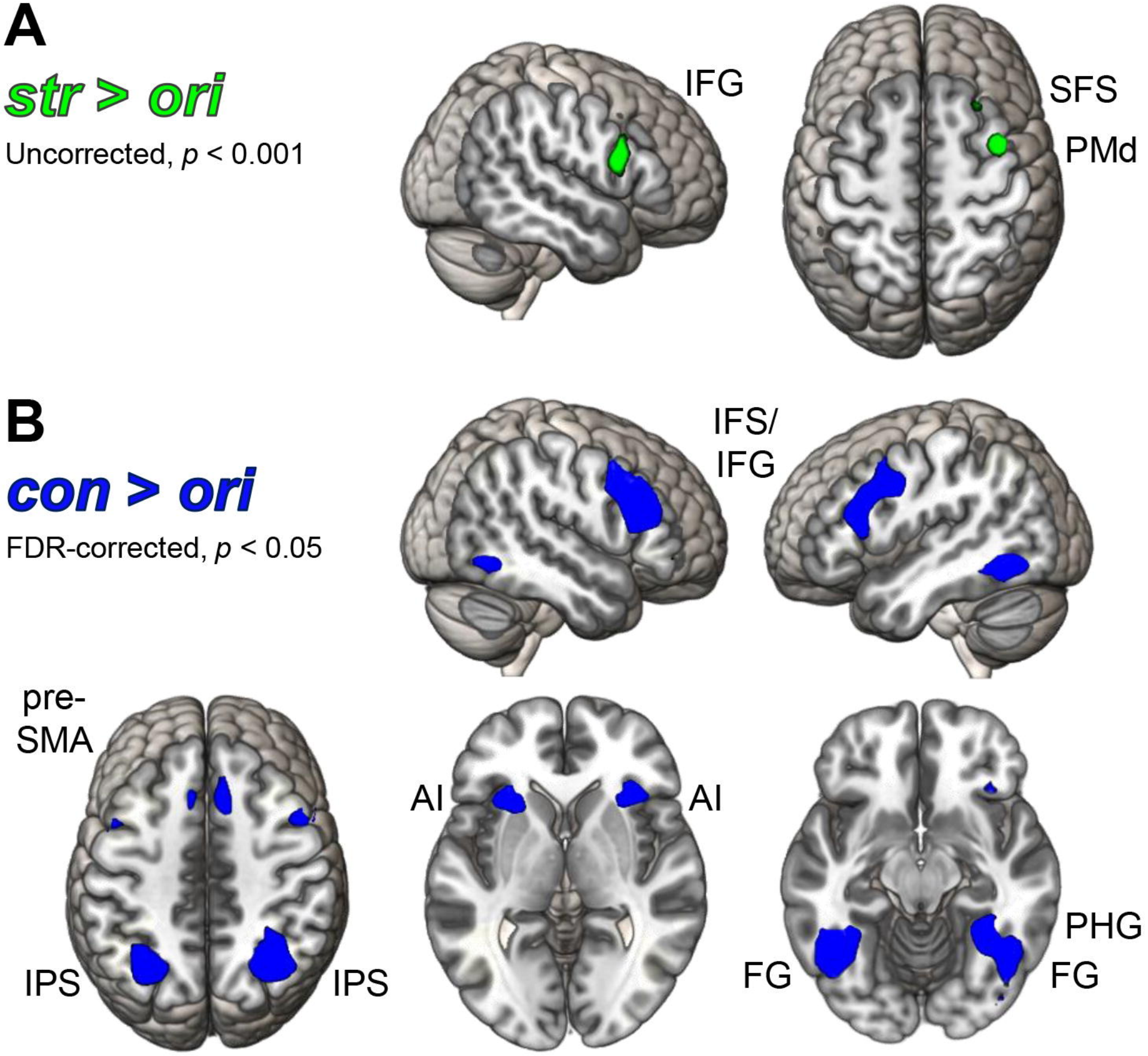
Figure 5. Whole brain activation for modification effects. (A) Uncorrected *t*-map ($p < 0.001$) for the *str* > *ori* contrast. (B) FDR-corrected *t*-map ($p < 0.05$) for the *con* > *ori* contrast. IFG: inferior frontal gyrus; SFS: superior frontal sulcus; PMd: dorsal premotor cortex; IFS: inferior frontal sulcus; pre-SMA: pre-supplementary motor area; IPS: intraparietal sulcus; AI: anterior insula; FG: fusiform gyrus; PHG: parahippocampal gyrus.


Step
Original
Structure modification
Content modification
1
2
3
4
5
6
7


Experimenter
Camera monitor
Experiment monitor
Camera
Toy objects
Camera frame
Participant


Video trial:
Jitter
Video (~ 9 – 18 s)
Fixation (2 s)
Rescuing princess?
Question trial:
Jitter
Question (max. 3 s)
Feedback (1 s)
Fixation (1 s)
Terminated by response


A
***
a
a vs. b ***
a
b
b
Discrimination index P_r
1
0.8
0.6
0.4
0.2
0
0.2
0.4
-0.6
str
con
MODIFICATION$_{FMRI}$
no
yes
B
***
a
a vs. b ***
a
b
b
Rating (modified$_{MT}$ videos)
4
3
2
1
str
con
C
p = 0.072
Reaction time (modified$_{MT}$ videos; ms)
2500
2000
1500
1000
500
0
str
D
2000
1500
1000
500
0
con
E
Rating (original$_{MT}$ videos)
4
3
2
1
ori
str
con
F
**
Reaction time (original$_{MT}$ videos; ms)
1000
800
600
400
200
0
ori
str
con


A
str > ori
Uncorrected, p < 0.001
IFG
SFS
PMd
B
con > ori
FDR-corrected, p < 0.05
IFS/
IFG
pre-
SMA
AI
AI
IPS
IPS
FG
FG
PHG