

A Naturalistic Paradigm to Investigate Postencoding Neural Activation Patterns in Relation to Subsequent Voluntary and Intrusive Recall of Distressing Events

Renée M. Visser, Richard N. Henson, and Emily A. Holmes

ABSTRACT

BACKGROUND: While neuroimaging has provided insights into the formation of episodic memories in relation to voluntary memory recall, less is known about neural mechanisms that cause memories to occur involuntarily, for example, as intrusive memories of trauma. Here, we investigated brain activity shortly after viewing distressing events as a function of whether memories for those events later intruded involuntarily. The postencoding period is particularly important because it is a period when clinical interventions could be applied.

METHODS: A total of 32 healthy volunteers underwent functional magnetic resonance imaging while viewing distressing film clips, interspersed with 5 minutes of awake (postencoding) rest. Voluntary memories of the films were assessed using free recall and verbal and visual recognition tests after a week, while intrusive (involuntary) memories were recorded in a diary throughout that week.

RESULTS: When analyzing functional magnetic resonance imaging responses related to watching the films, we replicated findings that those “hotspots” (salient moments within the films) that would later become intrusive memories elicited higher activation in parts of the brain’s salience network. Surprisingly, while the postencoding persistence of multivoxel correlation structures associated with entire film clips predicted subsequent voluntary recall, there was no evidence that they predicted subsequent intrusions.

CONCLUSIONS: Results replicate findings regarding the formation of intrusive memories during encoding and extend findings regarding the consolidation of information in postencoding rest in relation to voluntary memory. While we provided a first step using a naturalistic paradigm, further research is needed to elucidate the role of postencoding neural processes in the development of intrusive memories.

<https://doi.org/10.1016/j.bpsc.2021.08.006>

Memories of highly distressing events are often stronger and more vivid than memories of neutral events. Sometimes, such memories spring to mind involuntarily (1) and unwanted, rendering them intrusive (2,3). Intrusive memories are a cardinal symptom of posttraumatic stress disorder (4) and occur in various psychopathologies (5–7), causing emotional distress (8) and functional impairment (9). It is important to understand how intrusions arise to inform clinical interventions (10,11). Here, we ask the following question: What happens in the brain immediately after witnessing highly distressing events that leads to intrusive memories?

Most neuroimaging studies of emotional memory, and intrusive memory in particular, have focused on encoding or recall (12–21). However, much less is known about the consolidation of intrusive memories, i.e., the time-dependent process that stabilizes memory traces after initial encoding (22,23). This postencoding period is important because this is when clinical interventions could be applied.

A challenge for studying intrusive memories is that traumatic events comprise rich, multimodal experiences that take time to unfold, and only some moments within an event tend to intrude. Clinically, these moments are referred to as trauma “hotspots” (24). For example, after experiencing a car crash, one may have intrusions of the worst moments (e.g., windshield breaking) or something preceding the worst moments (e.g., car headlights approaching). Measuring neural responses to those moments is challenging because they can only be identified retrospectively and idiosyncratically. For example, diary entries show that hotspots within distressing films [an experimental model for psychological trauma (25)] can become intrusive for one individual but not another (12,13). Using such diaries, functional magnetic resonance imaging (fMRI) studies revealed that during encoding, hotspots that later intrude elicit more activity in the salience network (26,27), particularly the inferior frontal gyrus (IFG) and middle temporal gyrus (MTG) (12,13).

To study memory consolidation with fMRI, previous studies have examined intrinsic functional connectivity, such as between the amygdala and hippocampus, in relation to later memory (28,29). While such analyses may reveal general markers for consolidation, here we are interested in event-specific markers for consolidation. Using two types of multivariate approaches that characterize representational geometries (30), we aimed to assess the degree to which neural patterns associated with encoding persist or recur after encoding. In the first approach, those patterns are pairwise correlations between the time series of all voxels within a specific brain region, so-called multivoxel correlation structure (MVCS), which can also be viewed as a form of local functional connectivity. This approach is suitable for examining consolidation of a prolonged experience, such as a minute or so of a film clip. Previous work demonstrated that the persistence of MVCS in the hippocampus over a postencoding rest period predicted subsequent voluntary recall (31) and conditioned fear responses (32). The second approach uses multivoxel activation patterns associated with short events (e.g., a few seconds) and tests whether these patterns recur during a postencoding period. This approach is better suited to examine consolidation of specific hotspots. Research showed that such postencoding reactivation of item-specific patterns predicted better voluntary memory (33,34) and individual differences in fear/extinction memory (35,36).

We combined fMRI scanning on 2 consecutive days (sessions 1 and 2) with subsequent voluntary memory recall tests (session 3, 1 week later) and a daily diary of intrusive memories between sessions 1 and 3 (Figure 1). During scanning (session 1), participants watched six distressing film clips, alternated by rest periods of a few minutes, allowing assessment of offline (i.e., postencoding) processes related to the initial consolidation of each clip.

We tested two a priori hypotheses: first, based on previous findings (12,13,20), we predicted that univariate activation in the salience network while watching distressing clips would predict which hotspots later intrude during daily life. A priori

regions of interest (ROIs) included the hippocampus and amygdala based on their role in emotional memory (23,37–40), and left IFG and MTG for their previous implication in intrusive memories (12,13). Second, following the study by Tambini and Davachi (31), we predicted that persistence during the post-encoding rest period of hippocampal MVCS associated with encoding of an entire film clip would be related to voluntary recall of that clip (free recall, verbal/visual recognition 1 week later); an open question was whether this would also be the case for intrusive memories. Finally, given that intrusive memories typically only involve specific moments (hotspots) within a film, a third analysis explored whether the post-encoding reactivation of multivoxel patterns related to specific hotspots (rather than the whole clip) would predict whether these hotspots become intrusive.

METHODS AND MATERIALS

Participants

Participants were recruited via a volunteer panel and screened via telephone interview on criteria, including current mental health problems (see Supplemental Methods and Materials). Of 36 eligible participants, 1 did not finish the experiment and data from 3 participants could not be analyzed owing to excessive head motion. The final sample comprised 32 participants (age [in years]: mean = 23.2, SD = 4.5; 21 female, 11 male). All provided written informed consent in a manner approved by the Cambridge Psychological Research Ethics Committee (reference 2017.009).

Experimental Design

The experiment comprised three sessions (Figure 1), combining the trauma film paradigm (25) with fMRI scanning and completion of a diary to record the daily frequency of intrusive memories over a week. During session 1, six film clips previously shown to elicit intrusive memories (13) were presented. Clips depicted scenes involving actual/threatened death or serious injury (Table S1). Each clip was approximately

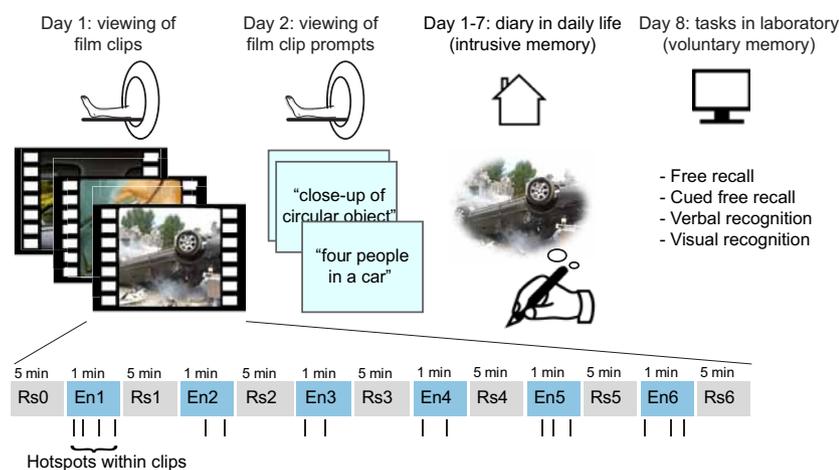


Figure 1. Within-subject design consisting of a naturalistic event-related functional magnetic resonance imaging paradigm and two follow-up sessions. Participants viewed six 1-minute film clips with distressing content while they were being scanned (indicated with “En” for encoding). Resting-state blood oxygen level-dependent activity during the postencoding period between the film clips was measured for 5 minutes (indicated with “Rs” for resting-state). Each film clip could contain several hotspots (indicated with “Hotspots within clips”). The time during and immediately after viewing each of the distressing clips is the analog peritraumatic period. For the next 7 days, participants recorded in a daily diary any intrusive memories, that is, mental images of the films that spontaneously popped into their mind in naturalistic settings, yielding a general intrusive memory frequency score. In addition, based on the diary entries, we were able to retrospectively, per individual, define intrusive and

nonintrusive hotspots within the film clips (naturalistic event-related functional magnetic resonance imaging paradigm). A day later (day 2), they returned for a second scan (not part of this paper). One week later, they returned to the lab to perform tests of voluntary memory recall, including free recall and verbal and visual recognition tests.

1 minute, followed by a 5-minute rest during which participants were encouraged to close their eyes. They could think about anything, with no restrictions. Clips were presented on a screen viewed via a mirror attached to the head coil. Sounds were presented via noise-shielding headphones. After the encoding phase, an additional 6-minute resting-state scan was acquired during which participants pressed a button if an image of the film popped into mind. A pen-and-paper diary was used to record film-related intrusive memories for 1 week (41), defining them as involuntarily occurring mental images. Participants noted their occurrence (morning/afternoon/evening) and content (Figure S1; Table S1).

Approximately 24 hours after session 1, participants returned (session 2) and were shown brief descriptions of some of the film clips (Figure 1). We do not report on this further; the retrieval procedure did not influence later memory (Supplemental Results).

Session 3 took place in a laboratory 7 to 8 days after session 1, starting with voluntary memory recall tests, i.e., (cued) free recall, verbal recognition, and visual recognition, followed by an intrusion provocation task. In addition, several self-report measures were administered across the 3 days. See Supplemental Methods and Materials for a full description of experimental procedure and assessments and a detailed description of how free recall and diary data were analyzed and how hotspots were defined.

MRI Acquisition

A Siemens 3T Prisma MRI scanner with 32-channel head coil was used to acquire functional data with a gradient-echo, echo-planar, multiband sequence (repetition time = 1190 ms; echo time = 30 ms; flip angle = 74°; 64 axial slices, acceleration factor 4, ascending acquisition; 2 × 2 × 2 mm voxel size; 192 × 192 × 128 field of view) covering the whole brain. Only the first run in session 1 (film encoding task), containing 2100 volumes, is analyzed here. High-resolution three-dimensional T1-weighted images (repetition time = 2250.0 ms, echo time = 3.02 ms, flip angle = 9°; 1 × 1 × 1 mm voxel size; 256 × 256 × 192 field of view) were collected for normalization.

Whole-Brain Parcellation and ROIs

ROIs were defined using the Brainnetome Atlas (BNA), containing 246 nonoverlapping cortical and subcortical regions (<http://atlas.brainnetome.org>). Multivariate analyses were performed both intra- and interregional. A priori ROIs included the hippocampus (BNA 215–218; 2261 voxels), amygdala (BNA 211–214; 607 voxels), left IFG (BNA 29, 31, 33, 35, 37, 39; 2161 voxels), and MTG (BNA 81–88). A salience network mask was created by creating 5-mm spheres around voxel coordinates reported by Hermans *et al.* [(27); Table S4, Expl, p. 21].

fMRI Analysis

fMRI preprocessing steps are described in Supplemental Methods and Materials. For the univariate analysis, 16 iconic hotspot moments were modeled as separate regressors using a double-gamma hemodynamic response function implemented in FEAT version 6.0, part of FSL (FMRIB Software Library, <http://www.fmrib.ox.ac.uk/fsl>). The duration varied per hotspot moment (Table S1; Figure S2). Six motion parameters

were included as regressors of no interest. Per individual, we contrasted intrusive with nonintrusive hotspots (Supplemental Methods and Materials). Seven participants who did not have any nonintrusive hotspots (only intrusive and ambiguous hotspots) were excluded from this analysis. Mixed-effects analyses were conducted to assess differences in intrusive versus nonintrusive hotspots at the group level. Activation was thresholded using clusters determined by $Z > 3.1$ ($p < .001$) and a (corrected) cluster significance threshold of $p = .05$ (Figure 2A). In addition, ROI analyses were performed after converting to percentage signal change (Figure 2B).

For the multivariate analysis within each ROI, we calculated neural correlation profiles uniquely related to film clips (6) and resting-state blocks (7); we extracted the time series for each ROI voxel and discarded the first 25 seconds of each rest epoch. While some initial consolidation could already occur during these first 25 seconds, discarding these data was necessary to ensure that resting-state data only contained activity related to the consolidation, not the actual encoding, of a clip, given the prolonged nature of the hemodynamic response function. To match temporal distance between the encoding epoch and the pre- and postencoding rest epochs, we also removed the last 25 seconds of each resting epoch (for results without removal of the first and last 25 seconds, see Supplemental Results). Next, correlations between the time series of every pair of voxels were calculated to obtain a correlation matrix, i.e., MVCS (31) (Figure 3A), and then Fisher z-transformed. The upper triangle of this matrix was then correlated across phases to determine MVCS similarity. To control for baseline levels of temporal autocorrelation [e.g., (42,43)], we subtracted the similarity between a film MVCS and the preceding rest MVCS from the similarity between a film MVCS and the following rest MVCS (Figure 3B). This yielded a persistence index, indicating the degree to which correlation patterns related to the encoding of a film clip persisted/reappeared during rest. In addition to this intraregional analysis, we repeated the same MVCS analyses across all brain regions after averaging the time series across voxels within each Brainnetome ROI.

Next, we tested whether the persistence of film-related correlation patterns (intraregional or whole brain) into post-encoding rest was associated with memory performance. Persistence indices were Fisher z-transformed prior to statistical testing. The independent variable was average persistence index across the six clips; the dependent variable was proportion of correctly recognized items on the verbal/visual recognition tests or total number of diary intrusions (Figure 3C). In addition, we used a Steiger test implemented in R (<https://cran.r-project.org/web/packages/cocor/cocor.pdf>) to directly compare (correlated) correlation coefficients, testing whether the MVCS persistence index more strongly predicts one type of memory, e.g., voluntary versus involuntary.

While MVCS analysis yields a neural profile reflecting activity over a prolonged period of time (here, a 1-min film clip), intrusive memories typically concern mental images of specific moments of only a few seconds within that event (hotspots). Our final analysis explored whether the degree of relative reactivation of a pattern related to a hotspot in the rest period following the clip that contained that particular hotspot reflected whether the hotspot would become intrusive. Relative

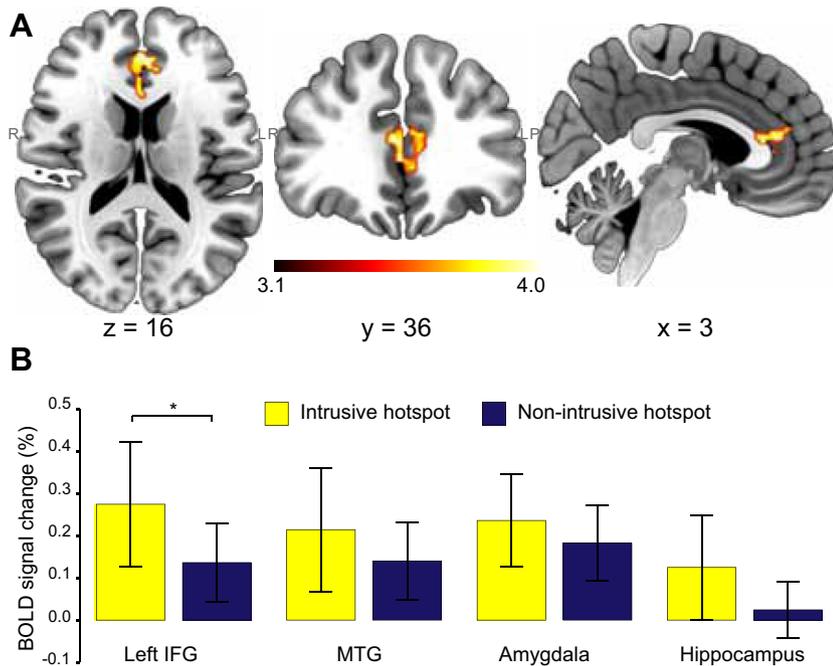


Figure 2. Neural correlates of intrusive memory encoding. Univariate activation at time of film viewing distinguishes hotspots that become intrusive vs. those that do not (nonintrusive). **(A)** Whole-brain analysis of the encoding of intrusive vs. nonintrusive hotspots showing a large cluster in the anterior cingulate cortex (314 voxels; 2512 mm³; center of gravity: $x = -0.5$, $y = 35.5$, $z = 15.5$; $Z_{\text{max}} = 4.30$). Activation is thresholded using clusters determined by $Z > 3.1$ ($p < .001$) and a (corrected) cluster significance threshold of $p = .05$. **(B)** Region of interest analysis for the left inferior frontal gyrus (IFG), middle temporal gyrus (MTG), hippocampus, and amygdala showing the blood oxygen level-dependent (BOLD) percentage signal change for intrusive and nonintrusive hotspots relative to implicit baseline. Values are means, with error bars representing the standard error of the mean. * $p < .05$. Image overlay created using MRICroGL (<https://www.nitrc.org/projects/mricrogl/>). A, anterior; L, left; P, posterior; R, right.

reactivation was assessed by correlating each pattern with each postencoding resting-state volume (Figure S3), correcting for baseline levels of temporal autocorrelations. We then used a linear mixed-effect model where relative reactivation was predicted by intrusiveness of the hotspot as fixed effect, including participants ($N = 32$) and hotspots (16) as random effects (for details, see Supplemental Methods and Materials).

Multiple comparisons correction (4 for a priori ROIs; 246 for exploratory whole-brain analyses) was done by limiting the false discovery rate (44). p values are reported two sided where hypotheses were nondirectional ($\alpha = 0.05$).

RESULTS

Behavioral Results: Intrusive and Voluntary Memory

Diaries of intrusive memories were scored by 2 independent raters (interrater reliability: $r = 0.94$); only descriptions of a clear mental image of a specific moment in a film clip were counted as intrusive memories. All clips were effective in eliciting intrusive memories. Further analysis of diary data yielded 16 iconic hotspot moments within the six clips (i.e., moments that appeared as separate intrusive memories for at least 3 participants) (examples in Table S1). For the univariate neuroimaging analyses and the hotspot pattern analysis, the crucial comparison was between hotspots that would become clear intrusive memories for a particular individual versus hotspots that would not, but would for other individuals [previously labeled potentials (12,13)]. The distribution of intruding and voluntary recalled hotspots is shown in Figure S4.

Summary statistics and correlations among memory measures are presented in Table 1 (for self-report measures, see Table S2). Across individuals, measures of voluntary recall were related to each other, as were measures of involuntary recall, but as expected, no significant correlations were

observed between voluntary and involuntary memory [in line with (41)]. While the relationship between overall voluntary and involuntary memory across participants was weak, intrusive hotspots were better remembered than nonintrusive hotspots when analyzed within-participant (Supplemental Results).

Activation During Film Viewing Predicts Hotspots That Become Intrusive

We had predicted higher activity in the salience network (12,13), alongside the left IFG and MTG, during encoding of intrusive compared with nonintrusive hotspots. Because 7 participants only had intrusive or ambiguous hotspots (i.e., no nonintrusive hotspots), results are reported for the remaining 25 participants (Figure S5).

In the whole-brain search, only the anterior cingulate cortex showed greater activity during encoding of intrusive versus nonintrusive hotspots (Figure 2A). The opposite contrast did not reveal any significant clusters. ROI analyses (Figure 2B) showed higher activation for intrusive hotspots in the left IFG ($t_{24} = 2.40$, $p = .012$, one sided; $d = 0.48$), but not MTG ($t_{24} = 1.03$, $p = .156$, one sided; $d = 0.21$) or amygdala ($t_{24} = 0.91$, $p = .187$, one sided; $d = 0.18$). The hippocampus showed no significant difference between intrusive and nonintrusive hotspots ($t_{24} = 1.46$, $p = .158$, two sided; $d = 0.29$), consistent with previous findings (12,13). To facilitate comparison with studies that used more liberal thresholds (12,13), Table S3 and Figure S6 display whole-brain results using a cluster defining threshold of $Z > 2.3$ ($p = .01$), indicating widespread activation in the salience network. Moreover, a salience network mask created from reported coordinates (27) also revealed higher activation for intrusive hotspots ($t_{24} = 1.79$, $p = .043$, one sided; $d = 0.36$).

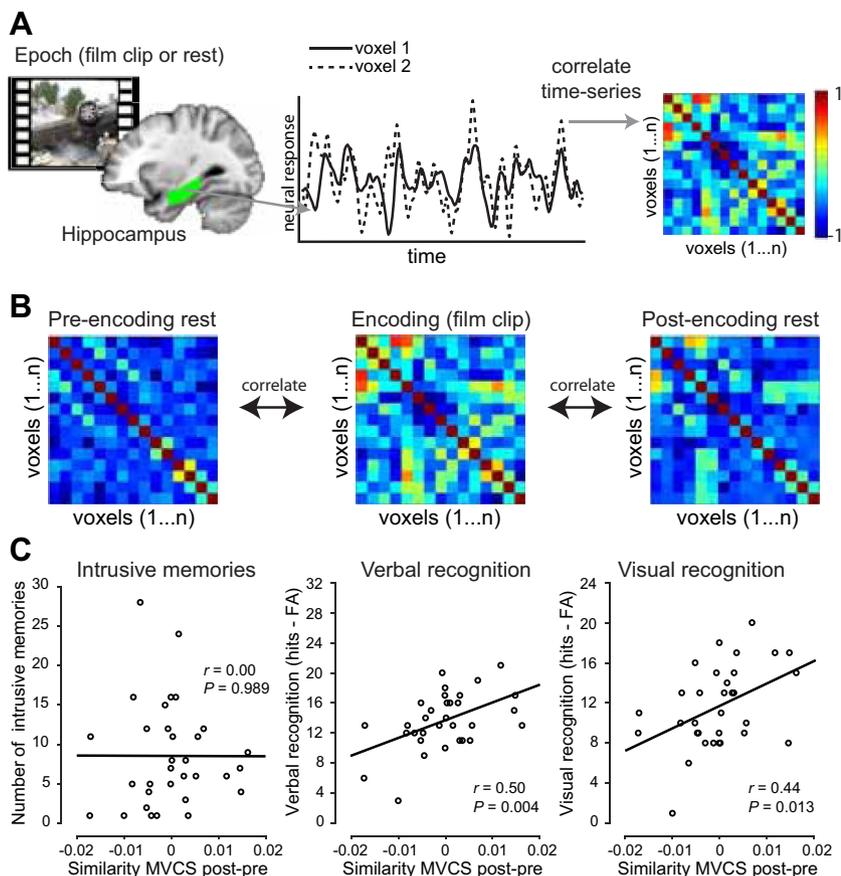


Figure 3. The persistence of film-related multivoxel correlation patterns into postencoding rest in the hippocampus predicts subsequent voluntary recall but not subsequent intrusive memory frequency. **(A)** To obtain neural response profiles related to encoding film clips (6) and postencoding resting-state blocks (7), we extracted the time series related to the epoch for each voxel in the hippocampus. Then all voxel time series were correlated with one another to obtain a multivoxel correlation structure (MVCS) per epoch (film clip or resting-state block), discarding the first and last 25 s of the resting state to avoid blurring across contiguous blocks. **(B)** Next, the similarity between MVCSs was assessed by calculating pairwise Pearson correlations between different Fisher z-transformed MVCSs related to the different film clips and resting-state blocks. The crucial comparison was the similarity between a film MVCS and the postencoding rest MVCS compared with film MVCS and pre-encoding rest MVCS. This yielded a relative persistence index indicating the degree to which correlation patterns related to the encoding of a film clip persist during postencoding rest. **(C)** In the hippocampus, this index did not predict the frequency of intrusive memories over the week. It did, however, predict verbal and visual recognition 1 week later across individuals. FA, false alarm.

Persistence of Film-Related Multivoxel Correlation Patterns Into Postencoding Rest Predicts Subsequent Voluntary Recall but Not Intrusive Memory

We examined the degree to which neural activity related to encoding persists during the 5-minute postencoding rest by calculating MVCS (31) (Figure 3A). Next, we compared the

similarity (Pearson correlation) between MVCS for a film clip and MVCS for the subsequent rest with that between MVCS for a film clip and MVCS for the preceding rest (Figure 3B), yielding a relative persistence index of neural patterns related to the encoding of a film clip.

While there was no main effect of persistence in the hippocampal ROI ($t_{31} = 0.15, p = .885$; but see Supplemental

Table 1. Summary Statistics and Pairwise Pearson Correlations Among Measures of Voluntary and Involuntary Recall (N = 32)

Measure	Mean (SD)	Range	1	2	3	4	5	6	7
1 Intrusive Memories Reported in Diary Over 1 Week	8.6 (6.6)	1–28	1.00	–	–	–	–	–	–
2 Intrusions During Resting State Day 1	6.5 (5.2)	0–21	0.31 ^a	1.00	–	–	–	–	–
3 Intrusions During Laboratory Task Day 8	6.7 (5.7)	1–25	0.65 ^b	0.39 ^c	1.00	–	–	–	–
4 FR, Total Number of Scenes Recalled	39.9 (7.4)	22–51	0.10	0.03	–0.13	1.00	–	–	–
5 FR, Total Number of Event Details	106.6 (29.2)	57–178	0.19	0.21	0.07	0.81 ^b	1.00	–	–
6 Verbal Recognition, Hits–FA (Out of 32)	13.7 (3.8)	3–21	0.18	0.28	0.11	0.50 ^d	0.50 ^d	1.00	–
7 Visual Recognition, Hits–FA (Out of 24)	11.7 (4.1)	1–20	0.14	0.15	–0.02	0.59 ^b	0.49 ^d	0.51 ^d	1.00

FA, false alarm; FR, free recall.

^a $p = .05$ to $.08$

^b $p < .001$.

^c $p < .05$.

^d $p < .01$.

Results), the hippocampal persistence index predicted, across participants, both verbal ($r = 0.50$; $p = .004$) and visual ($r = 0.44$; $p = .013$) recognition memory 1 week later (Figure 3C), replicating previous research (31). Hippocampal MVCS persistence was not related to free recall ($r = 0.29$; $p = .110$), but analysis across all 246 ROIs revealed such an effect in two regions in the IFG (BNA 35, 40) (Figure 4A). Numerous other brain areas, mostly in the medial temporal lobe and anterior midline (Figure 4B, C), showed a relationship between the persistence index and later verbal and visual recognition memory (Tables S4–S7), although only visual areas showed a relationship with visual recognition memory (Figure 4C).

In contrast, persistence in the hippocampal ROI did not predict intrusive memory frequency (Figure 3C; Table S4), nor did any of the 246 BNA ROIs (45) (see Supplemental Results for effects at $p < .01$ uncorrected and additional analyses). Indeed, direct comparison of the correlation coefficients revealed that the relationship between MVCS persistence in the hippocampus and verbal recognition was significantly

stronger than the relationship between MVCS persistence and the frequency of intrusive memories ($z = 2.26$, $p = .024$); the same contrast for visual recognition memory showed a marginally significant difference ($z = 1.90$, $p = .058$).

Finally, apart from voxel-by-voxel correlation patterns per region, we assessed region-by-region (246×246) correlation patterns, using voxel-averaged time series per ROI to calculate a correlation matrix. There was no evidence that persistence of interregional correlations (i.e., whole-brain correlation structures) predicted intrusive memory frequency ($r = 0.08$; $p = .653$), verbal recognition ($r = 0.04$; $p = .814$), or free recall ($r = 0.03$; $p = .854$), but it did predict visual recognition memory ($r = 0.39$; $p = .027$).

Reactivation of Hotspot-Related Neural Response Patterns and Hotspot Intrusiveness

Our final analysis focused on the offline reactivation of multi-voxel patterns related to specific hotspots rather than the

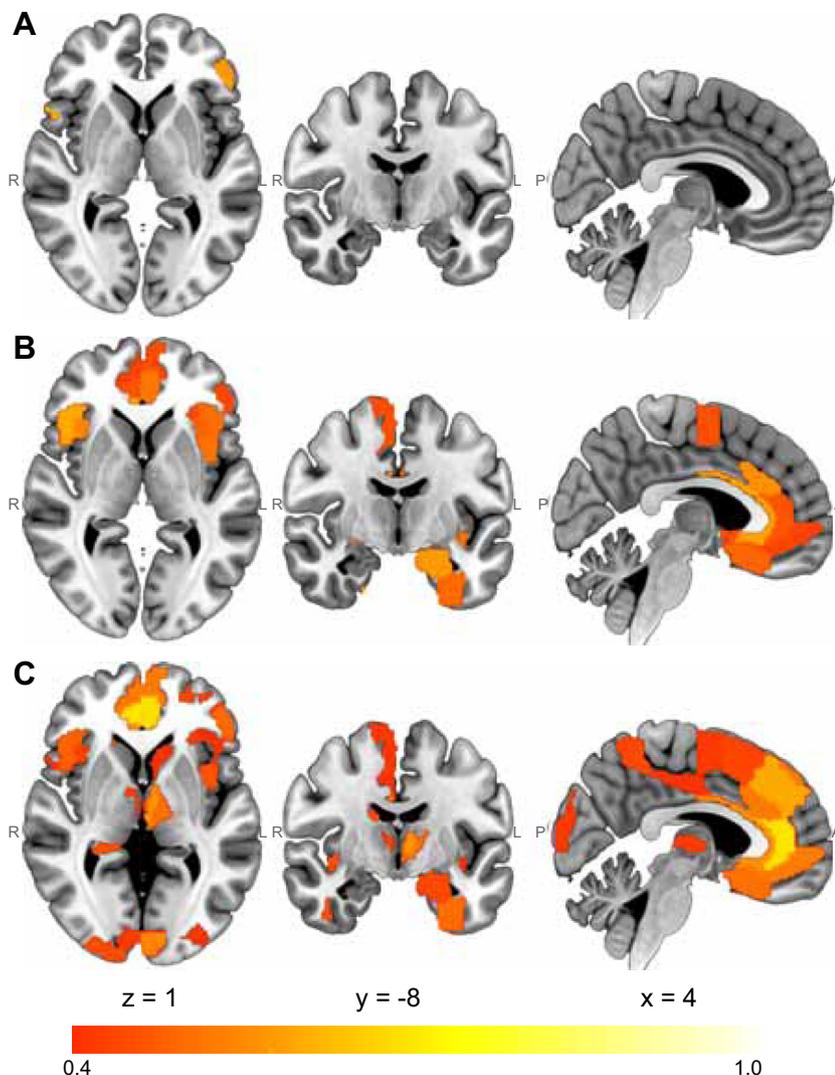


Figure 4. Persistence of film-related multivoxel correlation patterns into postencoding rest in relation to subsequent voluntary recall. Areas where the persistence of multivoxel correlation patterns related to the encoding of an entire film clip in postencoding rests predicts the number of scenes recalled during free recall (A), verbal recognition (B), or visual recognition (C) of the material from the film clips 1 week later ($N = 32$). Colors depict Pearson correlation values (r). Only areas that survive false discovery rate correction for number of regions (246) are displayed (free recall: $r > 0.60$, $p < .001$; verbal recognition: $r > 0.47$, $p < .007$; visual recognition: $r > 0.44$, $p < .013$). A, anterior; L, left; P, posterior; R, right.

whole clip. None of the effects survived correction for multiple comparisons. Uncorrected effects at $p < .01$ are reported in [Supplemental Results](#).

DISCUSSION

We used fMRI to examine which neural processes related to the initial consolidation of a memory of a distressing event render it intrusive. While previous studies focused on neural processes during distressing events or when involuntarily recalling events, our study focuses on processes that occur immediately after distressing events (i.e., while the experience is being consolidated into long-term memory). While our data showed that voluntary recall can be predicted from neural patterns in postencoding rest periods, none of the postencoding neural measures predicted subsequent intrusive memories. Instead, only activation at the time of encoding indicated whether a moment would become intrusive.

Findings build on previous work. First, we found activation in the anterior cingulate cortex during the encoding of negative events that subsequently become intrusive memories. The anterior cingulate cortex is part of the salience network that has previously shown similar effects (12,13); indeed, we found more widespread activation in other regions of this network, such as the midcingulate cortex, insular cortex, and IFG, at the more liberal threshold used in those previous studies.

Second, while the persistence of encoding-related correlation patterns (MVCS) into postencoding rest did not predict intrusive memory frequency over the subsequent week, it did predict voluntary memory recall, in line with previous findings (31). We extend these findings in four ways: 1) instead of testing on the same day, we tested voluntary memory 1 week after encoding, suggesting that previously described postencoding processes (31) reflect processes important for longer-term memory consolidation; 2) we used more naturalistic stimuli (films), with more distressing content, than the words/pictures used in previous memory consolidation fMRI studies (which may help clinical translation), alongside multiple measures of voluntary and involuntary memory; 3) we show that the prediction of subsequent voluntary memory retrieval is not restricted to the hippocampus but involves a large-scale network including areas in the medial temporal lobe and along the anterior midline; and 4) as we correct for preceding rest correlations, the MVCS persistence index only contains variance unique to consolidation of the film clip of interest, and we show that it is this unique variance that predicts better voluntary memory across individuals.

Our third analysis exploratively examined whether narrowing the focus on offline reactivation of multivoxel response patterns related to specific hotspots (rather than the whole clip) would indicate whether these hotspots become intrusive or not. After correction for multiple comparisons (246 brain areas), no relationship was found between postencoding reactivation of hotspots and their subsequent intrusiveness. Limitations inherent to the current naturalistic design restrict conclusions from this null finding. First, our interest in film-specific consolidation processes required resting-state periods in between film clips. Therefore, a small number of films was used, and only those were selected that had proved to be most effective in eliciting intrusions. This resulted in a low number of

nonintrusive control scenes compared with previous studies, corroborated by the fact that we observed weaker univariate activation when comparing intrusive versus nonintrusive hotspots during encoding (12,13), suggesting less power. Second, the temporal distance between events influences neural pattern similarity because of autocorrelations in fMRI data. This is problematic when examining naturally occurring hotspots in film clips because the timing of a hotspot within a clip with respect to the prefilm or postfilm rest period may affect the relative strength of the similarity values and thus the degree of reactivation as indexed by similarity values. Despite allowing us a means by which to study the idiosyncratic nature of intrusive memory (individuals have intrusions of different events), longer film clips mean that there is more variation in timing and duration of possible intrusive moments; this variation makes it harder to match moments across individuals or identify clean yet equally distressing controls. Future research could include extending measurements to the hours after film viewing to better understand how consolidation processes unfold over time. In addition, intrusions could be weighed by the distress they elicit (46) to facilitate clinical translation, although intrusions of distressing film clips [unless acquired in the line of work (4)] will never be more than a proxy of intrusions of real-life trauma.

A final limitation of this naturalistic event-related design, where hotspots can only be identified retrospectively and no interventions are applied, is that it restricts direct comparisons of the neural mechanisms underlying voluntary versus intrusive memory recall. The question whether voluntary and intrusive memory rely on distinct memory systems is part of an ongoing debate (41,47–52). Behaviorally, voluntary recall after 1 week was better for those moments that had intruded during that week. This is in line with recent findings (53) and suggests that intrusive memories may sometimes serve as a type of rehearsal, either directly or because it prompts further reflection on the content of the intrusion. Thus, some dependencies can be found between voluntary and intrusive memory. Yet at the same time, people who experienced more intrusions did not appear to have better overall memory. That is, different measures of voluntary recall (free recall, verbal and visual recognition) were strongly related to each other, as were the measures of involuntary recall (intrusion frequencies in laboratory tasks and the diary), but no significant correlations were observed between measures of voluntary and involuntary recall across individuals. Furthermore, the persistence of multivoxel correlation patterns was related to voluntary memory only and was significantly more strongly related than to the total frequency of intrusions. The latter observations are more in line with proposals that the two types of memory are independent to a certain degree. For example, manipulations such as engagement in a visuospatial task (the computer game Tetris) immediately after watching trauma films have been found to reduce the frequency of intrusive memories over a week, while leaving voluntary memory intact, providing strong support for separate memory systems or processes underlying voluntary versus intrusive memories (41).

While so far only revealing mechanisms predicting voluntary recall, the present applications of multivariate fMRI hold promise. Several clinical theories make predictions with regard

to how stressful events are processed and represented in the brain (5,49,54–56), and so far, it has been difficult to test these ideas in humans. Compared with nonspecific assessments of consolidation (e.g., connectivity between two brain regions), applications assessing similarities in representational geometries (30,33–36,57–59) allow us to track neural encoding profiles, uniquely related to one particular hotspot or film clip, during postencoding periods and establish their relationship to subsequent memory. Such approaches may also have translational relevance by shedding light on neural processes occurring in the immediate aftermath of traumatic events. These processes could potentially be targeted by mechanistically informed interventions to prevent later symptoms, i.e., intrusive memories of trauma building up (11). This idea has been exploited in experimental research, where the administration of pharmacological and behavioral manipulations immediately following analog trauma resulted in fewer intrusive memories (60,61). While immediate interventions are not always possible, they are applicable when exposure to trauma can be anticipated, such as in the emergency room or when police officers or journalists watch distressing film footage in the line of duty (4). First indications from studies in hospital settings suggest promising clinical translation of very early interventions (62–64).

In summary, while the neural underpinnings of intrusive memory formation remain enigmatic, this naturalistic paradigm provides a first step that may inspire future research into factors that weaken and strengthen memories for highly distressing events in controlled laboratory settings. Such research may eventually elucidate the mechanisms underlying the efficacy of early interventions and help increase their effectiveness.

ACKNOWLEDGMENTS AND DISCLOSURES

This study is supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie agreement (Grant No. 705641; SUAI/023 RG92025 [to RMV]), the Netherlands Organisation for Scientific Research (Grant No. Veni 016.195.246 [to RMV]), the UK Medical Research Council Programme (Grant No. SUAG/046 G101400 [to RNH]), the Swedish Research Council (VR; 2017-00957 [to EAH]), the Lupina Foundation (to EAH), and the Oak Foundation (Grant No. OCAY-18-442 [to EAH]).

All authors were involved in the design of the study. RMV collected the data. RMV and EAH developed guidelines for scoring of diary data and scored these accordingly. RMV and RNH developed tools for the imaging data analysis. RMV analyzed the data. RMV, RNH, and EAH wrote the paper. All authors approved the final version.

We are grateful to the Holmes lab for their input during the design and setting up of the study and in particular Ella James for the independent assessment of the diary data. Furthermore, we would like to thank Josephine Newman and Nicole Creasy for help with developing and piloting the stimulus material; Boris Szarzynski, Dora Bosnjak, and Robin van Dijkhuizen for transcribing audio files and independent assessment of free recall data; Laura Singh for proofreading; H. Steven Scholte for providing analytic tools; and Nikolaus Kriegeskorte and Erno Hermans for useful discussions about design and data analysis.

Raw neuroimaging data are publicly available at <https://openneuro.org/datasets/ds003721/versions/1.0.0>. Other raw data, as well as summary data files, materials, and code, are publicly available at <https://osf.io/ucen5/>.

EAH reports serving on the board of trustees of the charity MQ: Transforming Mental Health but receives no remuneration for this role. All other authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Clinical Psychology (RMV), University of Amsterdam, Amsterdam, The Netherlands; Medical Research Council Cognition and Brain Sciences Unit and Department of Psychiatry (RNH), University of Cambridge, Cambridge, United Kingdom; Department of Psychology (EAH), Uppsala University, Uppsala; and Department of Clinical Neuroscience (EAH), Karolinska Institutet, Stockholm, Sweden.

Address correspondence to Renée M. Visser, Ph.D., at r.m.visser@uva.nl.

Received Apr 24, 2020; revised Jul 20, 2021; accepted Aug 11, 2021.

Supplementary material cited in this article is available online at <https://doi.org/10.1016/j.bpsc.2021.08.006>.

3D neuroimaging content is available for this article online at <https://doi.org/10.1016/j.bpsc.2021.08.006>.

REFERENCES

- Richardson-Klavehn A, Bjork RA (1988): Measures of memory. *Annu Rev Psychol* 39:475–543.
- Berntsen D (1996): Involuntary autobiographical memories. *Appl Cogn Psychol* 10:435–454.
- Kalivas PW, Paulus MP (2020): *Intrusive Thinking Across Neuropsychiatric Disorders: From Molecules to Free Will*. Cambridge, MA: MIT Press.
- American Psychiatric Association (2013): *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. Arlington, VA: American Psychiatric Publishing.
- Brewin CR, Gregory JD, Lipton M, Burgess N (2010): Intrusive images in psychological disorders: Characteristics, neural mechanisms, and treatment implications. *Psychol Rev* 117:210–232.
- Visser RM, Anderson MC, Aron AR, Banich MT, Brady KT, Huys QJM, et al. (2020): Neuropsychological mechanisms of intrusive thinking. In: Kalivas PW, Paulus MP, editors. (2020), Lupp JR (series editor). *Strüngmann Forum Reports. Intrusive Thinking Across Neuropsychiatric Disorders: From Molecules to Free Will*, vol. 30: Cambridge, MA: MIT Press, 124–184.
- Holmes EA, Mathews A (2010): Mental imagery in emotion and emotional disorders. *Clin Psychol Rev* 30:349–362.
- Foa EB, Riggs DS (1993): Post-traumatic stress disorder in rape victims. In: Oldham J, Riba MB, Tasman A, editors. (1993), *American Psychiatric Press Review of Psychiatry*, vol. 12: Washington, DC: American Psychiatric Press, 273–303.
- Iyadurai L, Visser RM, Lau-Zhu A, Porcheret K, Horsch A, Holmes EA, James EL (2019): Intrusive memories of trauma: A target for research bridging cognitive science and its clinical application. *Clin Psychol Rev* 69:67–82.
- Holmes EA, Craske MG, Graybiel AM (2014): Psychological Treatments: A call for mental-health science. *Nature* 511:287–289.
- Holmes EA, Ghaderi A, Harmer CJ, Ramchandani PG, Cuijpers P, Morrison AP, et al. (2018): The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. *Lancet Psychiatry* 5:237–286.
- Bourne C, Mackay CE, Holmes EA (2013): The neural basis of flashback formation: The impact of viewing trauma. *Psychol Med* 43:1521–1532.
- Clark IA, Holmes EA, Woolrich MW, Mackay CE (2016): Intrusive memories to traumatic footage: The neural basis of their encoding and involuntary recall. *Psychol Med* 46:505–518.
- Hall SA, Rubin DC, Miles A, Davis SW, Wing EA, Cabeza R, Berntsen D (2014): The neural basis of involuntary episodic memories. *J Cogn Neurosci* 26:2385–2399.
- Hall SA, Brodar KE, LaBar KS, Berntsen D, Rubin DC (2018): Neural responses to emotional involuntary memories in posttraumatic stress disorder: Differences in timing and activity. *Neuroimage Clin* 19:793–804.
- Whalley MG, Kroes MCW, Huntley Z, Rugg MD, Davis SW, Brewin CR (2013): An fMRI investigation of posttraumatic flashbacks. *Brain Cogn* 81:151–159.
- Waldhauser GT, Dahl MJ, Ruf-Leuschner M, Müller-Bamouh V, Schauer M, Axmacher N, et al. (2018): The neural dynamics of deficient memory control in heavily traumatized refugees. *Sci Rep* 8:13132.

18. Catarino A, Küpper CS, Werner-Seidler A, Dalgleish T, Anderson MC (2015): Failing to forget: Inhibitory-control deficits compromise memory suppression in posttraumatic stress disorder. *Psychol Sci* 26:604–616.
19. Gvozdanovic GA, Stämpfli P, Seifritz E, Rasch B (2017): Neural correlates of experimental trauma memory retrieval. *Hum Brain Mapp* 38:3592–3602.
20. Battaglini E, Liddell B, Das P, Malhi G, Felmingham K, Bryant RA (2016): Intrusive memories of distressing information: An fMRI study. *PLoS One* 11:e0140871.
21. Miedl SF, Rattel JA, Franke LK, Blechert J, Kronbichler M, Spooemaker VI, Wilhelm FH (2020): Neural processing during fear extinction predicts intrusive memories. *Biol Psychiatry Cogn Neurosci Neuroimaging* 5:403–411.
22. McGaugh JL (2000): Memory—A century of consolidation. *Science* 287:248–251.
23. McGaugh JL (2004): The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annu Rev Neurosci* 27:1–28.
24. Holmes EA, Grey N, Young KAD (2005): Intrusive images and “hot-spots” of trauma memories in posttraumatic stress disorder: An exploratory investigation of emotions and cognitive themes. *J Behav Ther Exp Psychiatry* 36:3–17.
25. James EL, Lau-Zhu A, Clark IA, Visser RM, Hagensars MA, Holmes EA (2016): The trauma film paradigm as an experimental psychopathology model of psychological trauma: Intrusive memories and beyond. *Clin Psychol Rev* 47:106–142.
26. Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, *et al.* (2007): Dissociable intrinsic connectivity networks for salience processing and executive control. *J Neurosci* 27:2349–2356.
27. Hermans EJ, Van Marle HJF, Ossewaarde L, Henckens MJAG, Qin S, Van Kesteren MTR, *et al.* (2011): Stress-related noradrenergic activity prompts large-scale neural network reconfiguration. *Science* 334:1151–1153.
28. de Voogd LD, Klumpers F, Fernández G, Hermans EJ (2017): Intrinsic functional connectivity between amygdala and hippocampus during rest predicts enhanced memory under stress. *Psychoneuroendocrinology* 75:192–202.
29. Gvozdanovic GA, Seifritz E, Stämpfli P, Canna A, Rasch B, Esposito F (2021): Experimental trauma rapidly modifies functional connectivity. *Brain Imaging Behav* 15:2017–2030.
30. Kriegeskorte N, Mur M, Bandettini P (2008): Representational similarity analysis—Connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4.
31. Tambini A, Davachi L (2013): Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. *Proc Natl Acad Sci U S A* 110:19591–19596.
32. Hermans EJ, Kanen JW, Tambini A, Fernández G, Davachi L, Phelps EA (2017): Persistence of amygdala-hippocampal connectivity and multi-voxel correlation structures during awake rest after fear learning predicts long-term expression of fear. *Cereb Cortex* 27:3028–3041.
33. Staresina BP, Alink A, Kriegeskorte N, Henson RN (2013): Awake reactivation predicts memory in humans. *Proc Natl Acad Sci U S A* 110:21159–21164.
34. Deuker L, Olligs J, Fell J, Kranz TA, Mormann F, Montag C, *et al.* (2013): Memory consolidation by replay of stimulus-specific neural activity. *J Neurosci* 33:19373–19383.
35. de Voogd LD, Fernández G, Hermans EJ (2016): Awake reactivation of emotional memory traces through hippocampal-neocortical interactions. *Neuroimage* 134:563–572.
36. Gerlicher AMV, Tüscher O, Kalisch R (2018): Dopamine-dependent prefrontal reactivations explain long-term benefit of fear extinction. *Nat Commun* 9:4294.
37. Clewett D, Murty VP (2019): Echoes of emotions past: How neuro-modulators determine what we recollect. *eNeuro* 6:1263–1271.
38. van Rooij SJH, Stevens JS, Ely TD, Hinrichs R, Michopoulos V, Winters SJ, *et al.* (2018): The role of the hippocampus in predicting future posttraumatic stress disorder symptoms in recently traumatized civilians. *Biol Psychiatry* 84:106–115.
39. Harnett NG, Stevens JS, Fani N, van Rooij SJH, Ely TD, Michopoulos V, *et al.* (2022): Acute posttraumatic symptoms are associated with multimodal neuroimaging structural covariance patterns: A possible role for the neural substrates of visual processing in posttraumatic stress disorder. *Biol Psychiatry Cogn Neurosci Neuroimaging* 7:129–138.
40. Harnett NG, van Rooij SJH, Ely TD, Lebois LAM, Murty VP, Jovanovic T, *et al.* (2021): Prognostic neuroimaging biomarkers of trauma-related psychopathology: Resting-state fMRI shortly after trauma predicts future PTSD and depression symptoms in the Aurora study. *Neuropsychopharmacology* 46:1263–1271.
41. Lau-Zhu A, Henson RN, Holmes EA (2019): Intrusive memories and voluntary memory of a trauma film: Differential effects of a cognitive interference task after encoding. *J Exp Psychol Gen* 148:2154–2180.
42. Visser RM, de Haan MIC, Beemsterboer T, Haver P, Kindt M, Scholte HS (2016): Quantifying learning-dependent changes in the brain: Single-trial multivoxel pattern analysis requires slow event-related fMRI. *Psychophysiology* 53:1117–1127.
43. Mumford JA, Davis T, Poldrack RA (2014): The impact of study design on pattern estimation for single-trial multivariate pattern analysis. *Neuroimage* 103:130–138.
44. Benjamini Y, Hochberg Y (1995): Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300.
45. Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, *et al.* (2016): The human Brainnetome atlas: A new brain atlas based on connective architecture. *Cereb Cortex* 26:3508–3526.
46. Franke LK, Rattel JA, Miedl SF, Danböck SK, Bürkner PC, Wilhelm FH (2021): Intrusive memories as conditioned responses to trauma cues: An empirically supported concept? *Behav Res Ther* 143:103848.
47. Visser RM, Lau-Zhu A, Henson RN, Holmes EA (2018): Multiple memory systems, multiple time points: How science can inform treatment to control the expression of unwanted emotional memories. *Philos Trans R Soc B Biol Sci* 373:1742.
48. Berntsen D (2009): *Involuntary Autobiographical Memories. An Introduction to the Unbidden Past.* Cambridge: Cambridge University Press.
49. Bisby JA, Burgess N (2017): Differential effects of negative emotion on memory for items and associations, and their relationship to intrusive imagery. *Curr Opin Behav Sci* 17:124–132.
50. Brewin CR (2014): Episodic memory, perceptual memory and their interaction: Foundations for a theory of posttraumatic stress disorder. *Psychol Bull* 140:69–97.
51. Rubin DC, Berntsen D, Ogle CM, Deffler SA, Beckham JC (2016): Scientific evidence versus outdated beliefs: A response to Brewin (2016). *J Abnorm Psychol* 125:1018–1021.
52. Berntsen D (2010): The unbidden past: Involuntary autobiographical memories as a basic mode of remembering. *Curr Dir Psychol Sci* 19:138–142.
53. Herz N, Bar-Haim Y, Holmes EA, Censor N (2020): Intrusive memories: A mechanistic signature for emotional memory persistence. *Behav Res Ther* 135:103752.
54. Ehlers A, Clark DM (2000): A cognitive model of posttraumatic stress disorder. *Behav Res Ther* 38:319–345.
55. Horowitz MJ (1975): Intrusive and repetitive thoughts after experimental stress. A summary. *Arch Gen Psychiatry* 32:1457–1463.
56. Bisby JA, Burgess N, Brewin CR (2020): Reduced memory coherence for negative events and its relationship to posttraumatic stress disorder. *Curr Dir Psychol Sci* 29:267–272.
57. Visser RM, Kunze AE, Westhoff B, Scholte HS, Kindt M (2015): Representational similarity analysis offers a preview of the noradrenergic modulation of long-term fear memory at the time of encoding. *Psychoneuroendocrinology* 55:8–20.
58. Visser RM, Scholte HS, Beemsterboer T, Kindt M (2013): Neural pattern similarity predicts long-term fear memory. *Nat Neurosci* 16:388–390.

Postencoding Neural Processes and (In)Voluntary Memory

59. Undeger I, Visser RM, Olsson A (2020): Neural pattern similarity unveils the integration of social information and aversive learning. *Cereb Cortex* 30:5410–5419.
60. Das RK, Tamman A, Nikolova V, Freeman TP, Bisby JA, Lazzarino AI, Kamboj SK (2016): Nitrous oxide speeds the reduction of distressing intrusive memories in an experimental model of psychological trauma. *Psychol Med* 46:1749–1759.
61. Holmes EA, James EL, Kilford EJ, Deerprouse C (2010): Key steps in developing a cognitive vaccine against traumatic flashbacks: Visuo-spatial Tetris versus verbal Pub Quiz. *PLoS One* 5:e13706.
62. Horsch A, Vial Y, Favrod C, Harari MM, Blackwell SE, Watson P, *et al.* (2017): Reducing intrusive traumatic memories after emergency caesarean section: A proof-of-principle randomized controlled study. *Behav Res Ther* 94:36–47.
63. Iyadurai L, Blackwell SE, Meiser-Stedman R, Watson PC, Bonsall MB, Geddes JR, *et al.* (2018): Preventing intrusive memories after trauma via a brief intervention involving Tetris computer game play in the emergency department: A proof-of-concept randomized controlled trial. *Mol Psychiatry* 23:674–682.
64. Kanstrup M, Singh L, Göransson KE, Widoff J, Taylor RS, Gamble B, *et al.* (2021): Reducing intrusive memories after trauma via a brief cognitive task intervention in the hospital emergency department: An exploratory pilot randomised controlled trial. *Transl Psychiatry* 11:30.