Consolidation of complex events via reinstatement in posterior cingulate cortex

Article (Published Version)


This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/57862/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
Behavioral/Cognitive

Consolidation of Complex Events via Reinstatement in Posterior Cingulate Cortex

Chris M. Bird,1,4* James L. Keidel,1,4* Leslie P. Ing,2,3 Aidan J. Horner,2,4 and Neil Burgess2,4

1School of Psychology, University of Sussex, Falmer BN1 9QH, United Kingdom, 2UCL Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, United Kingdom, 3King’s College Hospital, Denmark Hill, London SE5 9RS, United Kingdom, and 4UCL Institute of Neurology, London WC1N 3BG, United Kingdom

It is well-established that active rehearsal increases the efficacy of memory consolidation. It is also known that complex events are interpreted with reference to prior knowledge. However, comparatively little attention has been given to the neural underpinnings of these effects. In healthy adults humans, we investigated the impact of effortful, active rehearsal on memory for events by showing people several short video clips and then asking them to recall these clips, either aloud (Experiment 1) or silently while in an MRI scanner (Experiment 2). In both experiments, actively rehearsed clips were remembered in far greater detail than unrehearsed clips when tested a week later. In Experiment 1, highly similar descriptions of events were produced across retrieval trials, suggesting a degree of semanticization of the memories had taken place. In Experiment 2, spatial patterns of BOLD signal in medial temporal and posterior midline regions were correlated when encoding and rehearsing the same video. Moreover, the strength of this correlation in the posterior cingulate predicted the amount of information subsequently recalled. This is likely to reflect a strengthening of the representation of the video’s content. We argue that these representations combine both new episodic information and stored semantic knowledge (or “schemas”). We therefore suggest that posterior midline structures aid consolidation by reinstating and strengthening the associations between episodic details and more generic schematic information. This leads to the creation of coherent memory representations of lifelike, complex events that are resistant to forgetting, but somewhat inflexible and semantic-like in nature.

Key words: episodic memory; fMRI; hippocampus; MVPA

Inteoduction

Memory consolidation, whereby memories are stabilized via interactions between medial temporal lobe regions and the neocortex (Dudai, 2004; Wixted, 2004), is often considered to be a passive process, occurring in the absence of explicit, active recall of the memoranda. (Skaggs, 1925; Della Sala et al., 2003; Dewar et al., 2012).

A proposed mechanism for consolidation is via offline reinstatement of neuronal activity elicited during encoding (Marr, 1971; McClelland et al., 1995). Neuronal reinstatement relating to consolidation has been described in rodents (Davidson et al., 2009; Karlsson and Frank, 2009; Foster and Wilson, 2006) and...
Humans (via rhinal cortex “ripple” events; Axmacher et al., 2008). Multivariante fMRI studies have shown that, for simple stimuli, reinstatement of patterns of BOLD activity occurs spontaneously between encoding and test, and that the degree of reinstatement is associated with subsequent memory strength (Deiker et al., 2013; Staresina et al., 2013). Moreover, several other studies have used fMRI to show that BOLD activity following memory encoding is related to subsequent memory performance despite there being no overt instruction to rehearse (Tambini et al., 2010; Ben-Yakov et al., 2013; Elman et al., 2013; Staresina et al., 2013; Tambini and Davachi, 2013; Ben-Yakov et al., 2014).

In contrast to this apparently effortless mechanism of consolidation, Conway (2009) noted that detailed memories of events are typically forgotten within a week; they are presumably not fully consolidated or are lost in a transformation to a gist-like representation. Nevertheless, even detailed memories for some events are well retained for long periods. A feature of such memories is that they are often actively retrieved many times; and indeed, active rehearsal of information is a powerful method of boosting later recall (Roediger and Karpicke, 2006). In addition, memory for complex events differs from memory for simple stimuli in that, to comprehend the sequence of unfolding actions, it is necessary to interpret them with reference to our prior knowledge of similar situations, sometimes referred to as memory “schemas” or “scripts” (Bartlett, 1932; Bransford and Johnson, 1972; Bower et al., 1979; Brewer and Treyns, 1981). Therefore, memory for a complex lifelike event is never a straightforward representation of the incoming information, but is instead a combination of this and our stored semantic knowledge.

A network of brain regions is involved in long-term memory, including the hippocampus and medial temporal lobes as well as parts of the thalamus, frontal lobes, and medial parietal regions, such as the precuneus, posterior cingulate, and retrosplenial cortex (Scolville and Milner, 1957; Rudge and Warrington, 1991; Squire, 1992; Aggleton and Brown, 1999; Eichenbaum, 2001; Wagner et al., 2005). It has been suggested that these latter regions are important for spatially coherent visual imagery of environments (Byrne et al., 2007), and they have also been associated with semantic memory processes (Binder et al., 2009).

To investigate memory for extended lifelike events, we used videos as memoranda and probed memory over 1 or 2 weeks. After watching the videos, participants rehearsed their content, either by describing them aloud (Experiment 1) or silently to themselves while in an MRI scanner (Experiment 2). Experiment 1 investigated the effect of active rehearsal on the durability of memories. Experiment 2 aimed at identifying whether reinstatement of BOLD activity can be detected when the memoranda are unfamiliar tri-unique videos and whether the strength of reinstatement during periods of active rehearsal is associated with subsequent memory recall.

Materials and Methods

Participants

All participants gave written consent and were paid for participating, as approved by the local Research Ethics Committee. All were right-handed with normal or corrected-to-normal vision and reported to be in good health with no history of neurological disease. Experiment 1 involved 13 participants (6 female, mean age 22.3 years, range 18–30 years). Experiment 2 involved 16 participants (10 female, mean age 26.6 years, range 20–34 years).

Stimuli

A total of 26 videos were used. The videos were clips taken from short films or videos posted on www.YouTube.com. The clips lasted on average 38 s (range = 29–48 s) and were presented without sound. All the videos depicted live action, taking place outside (15 videos) or inside (11 videos). Of the 26 videos used, 10 involved multiple characters interacting, 11 involved interactions between two main characters, and 4 involved just one character or no characters (1 video). Thirteen videos involved humorous content. None of the videos involved distressing or highly emotionally negative content.

Procedure for Experiment 1

Twenty-one of the 26 videos were used, divided into three sets of seven (Sets 1, 2, and 3). Pilot studies ensured that the average number of details recalled for the videos in each set was approximately the same. The experiment used a within-subjects design with three recall conditions (Fig. 1A). The assignment of Sets 1, 2, and 3 to the three conditions was counterbalanced across participants. All participants watched all 21 videos in a single encoding session. The seven videos in Condition 1 were then recalled on days 1, 8, and 18. The seven videos in Condition 2 were recalled on days 1 and 18. The seven videos in Condition 3 were recalled on days 8 and 18.

Participants were told that their task was to watch videos and try to remember their content for a memory test. They were then shown a practice video similar to the ones used in the main task. Immediately after watching the practice video, participants were asked to describe the video in as much detail as possible, while the experimenter scored their recall for the content of the video. Following this, the participants were shown a checklist of details contained in the video, with the experimenter’s “ticks” indicating those details that had been successfully recalled. The purpose of this was to emphasize that descriptions of the videos should be as detailed as possible and that credit would be given for the recall of any detail that was specific to the video.

In the encoding phase, participants watched all 21 videos, presented on a computer monitor via PowerPoint. After the action finished, the screen would freeze, and the participants would press a key to start the next video. Each video had a title that was presented above the video throughout its duration. Participants were instructed to attend to the title as well as the video. In the rehearsal/recall phase, participants were prompted with the video’s title and asked to describe it in as much detail as possible. The order of the videos in the test phase was pseudo-randomized according to the day and the experimental conditions (see design). Responses were audio recorded and later scored for the number of individual details recalled.

Recall of the videos was scored in a similar way to commonly used “prose recall” tests of memory, where a point is awarded for every “idea” correctly recalled (e.g., Wilson et al., 1991). Pilot work identified a number of details that were consistently recalled for each of the videos. These constituted a “checklist” of details that served as a framework for scoring recall of each video. Points were awarded for correctly recalling actions (e.g., “someone swiped their card to open the door”) and specific descriptions (e.g., “a balding man” but not “a man”). In some cases, points awarded for specific descriptions were capped (e.g., if the video involved 4 adults in their twenties, one point was awarded for “adults in their twenties” rather than 4 points for describing each individual in turn as “in their twenties”). Correct details were always awarded a point regardless of whether they were included on the checklist. Redundant information, for example, simply restating information that was in the video’s title, was not credited.

If a participant was unable to recall anything about the video when given the title, then “hints” were provided to cue recall. These hints were based on the checklist of details for each video. The first hint would describe a salient feature of the opening scene; for example, it would describe the characters present or the location of the opening scene. The purpose of giving hints was to cue recall of any details of the video that the participant could remember. Therefore, hints would be provided until (1) the majority of the video had been described or (2) the participant indicated that they had no memory of the particular video. Points were awarded for any details recalled that were additional to the information contained in the hints. On subsequent recall sessions, credit would not be
given for recalling details that had been provided in a hint during a previous session. Errors in recall were not formally analyzed. The responses were all scored independently by two raters (the authors C.M.B. and L.P.I.), and the mean score was used as the measure of recall. One rater (C.M.B.) was blind to the experimental conditions from which each recording came.

Procedure for Experiment 2
Before scanning, participants performed a practice trial with the examiner according to the same procedure as for Experiment 1. They were then instructed that the task they would perform in the scanner would be similar to this. However, rather than describe the videos out loud, they were requested to rehearse the content of the videos silently to themselves. Participants were informed that their recall of the videos would be tested a week after the scan. Participants were instructed that the task they would perform in the scanner would be similar to this. However, rather than describe the videos out loud, they were requested to rehearse the content of the videos silently to themselves. Participants were informed that their recall of the videos would be tested a week after the scan. The experiment was performed in two runs, with each run comprising an encoding phase and a rehearsal phase (Fig. 1 B, C). Each encoding trial involved showing a video while the title was displayed at the top of the screen throughout. At the end of the video, the final frame remained on screen until the participant pressed a button. This was to ensure that the participants concentrated on the task. A blank screen was shown during the intertrial interval (ITI) between the button press and the start of the next video. The ITI was jittered between 6 and 10 s (mean = 8 s). The rehearsal phase used a cued recall paradigm, with the video title serving as the rehearsal cue. On each rehearsal trial, the title of the video appeared on screen for 2.5 s and then faded but remained visible. Participants were instructed to spend approximately the same amount of time rehearsing the video as the video had originally lasted. After rehearsal, the participant pressed a button to indicate they had finished. They were then asked to rate how vividly they could remember the video on a visual analog scale from 1 to 5. A blank screen was shown during the ITI between rating the vividness and the subsequent cue. The ITI was jittered between 4 and 9 s (mean = 6.5 s). There were a total of 26 encoding trials and 20 rehearsal trials. The 6 unrehearsed trials served as a baseline for memory performance after a week given no instruction to rehearse.

In the testing phase, performed a week after scanning, participants were prompted with the video’s title and asked to describe it in as much detail as possible. Responses were audio recorded and later scored for the number of details correctly recalled. Scoring followed the same procedure as for Experiment 1. Responses were scored by one rater (C.M.B. or L.P.I.) who was not blind to the experimental condition because the same videos were allocated to the two experimental conditions for all participants. These scores were then checked, but not scored independently, by a third rater (C.O., see Acknowledgments).

MRI acquisition
BOLD-sensitive T2*-weighted fMRI measurements were acquired on a 3T Siemens Allegra scanner using a gradient-echo EPI pulse sequence with the following parameters: repetition time, 2880 ms; echo time, 30 ms; flip angle, 90°; slice thickness, 2 mm; interslice gap, 1 mm; in-plane resolution, 3 × 3 mm; field of view, 192 mm2; 48 slices per volume. The sequence was optimized to minimize signal dropout in the medial temporal lobes (Weiskopf et al., 2006). In addition, a field map using a double-echo fast, low-angle shot sequence was recorded for distortion correction of the acquired echo planar images (Weiskopf et al., 2006). After the functional scans, a T1-weighted structural image (1 mm3 resolution) was acquired for coregistration preprocessing steps.

Image preprocessing
The first five EPI volumes collected in each run were discarded to allow for T1 equilibration. For GLM analysis, the remaining functional images were then spatially realigned to the first image in the times series and were corrected for distortions based on the field map (Hutton et al., 2002) and the interaction of motion and distortion using the Realign and Unwarp routine in SPM 8 (Andersson et al., 2001; Hutton et al., 2002). Data were then corrected for the offset time of slice activation with reference to the middle slice of the first volume. Using AFNI (Cox, 1996), each subject’s structural scan was then registered to the first functional volume acquired and warped into Talairach space. The transform parameters esti-
mated in this normalization step were then applied to the functional data. The functional data were smoothed with a 6 mm FWHM Gaussian kernel, and scaled to percentage signal change. The data for the representational similarity analyses were preprocessed in the same way, except that the functional data were not normalized or smoothed.

**Data analysis**

GLM. BOLD responses were estimated using a GLM implemented in AFNI’s 3DDeconvolve program. Task regressors included video presentation and retrieval periods, rating cues, ratings, and button presses. All experimental periods were modeled as boxcars whose duration matched the individual length of the modeled period, except for button presses at the end of each video, which were modeled as impulse functions. In addition, the six motion parameters estimated from the realignment step were included in the model, as were baseline Legendre polynomials up to the eighth order to account for scanner drift. The group-level effects for were included in the model, as were baseline Legendre polynomials up to the eighth order. To account for the fact that some videos were recalled better overall than others, the mean number of details recalled by all participants for a video was subtracted from the individual’s score for that video (although note that highly similar results were obtained when simply using the raw number of details recalled).

**Representational similarity.** For the representational similarity analyses (RSA), each of the two runs was analyzed in separate GLMs as described above, with the exception that each video and retrieval period was modeled with its own regressor. Searchlight maps for each participant were then generated as follows: at each voxel, a sphere was created consisting of all the voxels within 10 mm of the voxel (on average, 160 voxels per sphere). The vectors of t statistics within this sphere for all the encoding periods of rehearsed videos (i.e., 20 of the 26 presented; see above), and all the rehearsal periods were then correlated, and the resulting Fisher-transformed r value was assigned to the center voxel of the sphere for each specific encoding-rehearsal pairing. Thus, each voxel had 400 values associated with it, 20 of which represented the voxelwise correlation between matching encoding and rehearsal periods (e.g., watching “Encounters at the office” and rehearsing “Encounters at the office”); whereas the remaining 380 represented nonmatching encoding-rehearsal pairs (e.g., watching “Encounters at the office” and rehearsing a different video).

To identify brain regions whose representations were more similar for matching than nonmatching pairs, we calculated for each participant the mean correlation for matching pairs minus the mean of the nonmatching pairs and assigned this value to each voxel. These maps were tested against a null hypothesis of zero using a one-sample t test across subjects. To assess how participants’ memory for individual videos affected the degree of representational similarity between the encoding and retrieval periods for that video, we directly contrasted the RSA above with an analysis in which the contribution of each matching encoding-retrieval pair was weighted by the number of details a given participant remembered for a given video compared with the participants as a group. For each subject, we calculated the difference between the weighted mean of the matching pairs and the unweighted mean of the matching pairs. If the degree of memory for a given video is unrelated to the similarity of the encoding-rehearsal pair, then the expected value of this contrast is zero. However, if the correlation between encoding-rehearsal pairs is increased when more details from that video are remembered, then the expected value of the contrast is greater than zero. This was assessed by a one sample t test across subjects.

**Results**

**Experiment 1**

The inter-rater reliability for the scores from Experiment 1, as calculated by the Pearson correlation between the two raters’ scores for each participant’s recall of each video, was 0.93 (p < 0.001). The results from Experiment 1 are shown in Figure 2.

Because the study did not use a fully factorial design, performance within conditions was compared across days and performance within days was compared across conditions using planned paired-sample t tests. There were 9 separate comparisons, giving a Bonferroni-corrected p value of 0.0056 to be significant at a level of α = 0.05. For Set 1, performance was significantly lower on day 8 compared with day 1 (t(12) = 4.65, p < 0.001), but there was no difference between performance on day 8 and day 18. For Set 2, there was a significant reduction in performance on day 18 compared with day 1 (t(12) = 10.8, p < 0.001). For Set 3, there was no difference in performance on day 8 and day 18. On day 1, recall in Sets 1 and 2 was not different. On day 8, there was a highly significant difference in the number of details recalled between Set 1 and Set 3 (t(12) = 10.3, p < 0.0001). On day 18, the difference between Sets 1 and 2 was significant (t(12) = 3.4, p = 0.0050), and the differences between Set 1 and Set 3 and between Set 2 and Set 3 were significant (t(12) > 5.5, p < 0.001 for both comparisons).

It is possible that our findings reflect large numbers of videos being forgotten (no details recalled) in Condition 3, whereas the remaining videos were vividly recalled. To investigate this possibility, we excluded from the analysis any video where a hint had been provided to aid recall. Hints were only provided for 6.0% of the video descriptions, although the majority of hints were provided for videos in Condition 3 (29 of 38 hints). Removing these video descriptions had negligible effects on the data; the biggest differences were for Condition 3, where the day 8 recall scores were 5.4 versus 5.7 and the day 18 recall scores were 5.7 versus 6.1 (latter scores reflect the mean number of details recalled after removing videos where hints were provided). Therefore, rehearsal appears predominantly to boost the number of details recalled in the videos, rather than result in fewer forgotten videos.

When scoring the descriptions of the videos, it was striking that they were highly consistent within individuals across the testing sessions. This was not only the case for correctly recalled information but also for incorrect details. For example, one participant falsely recalled a kiss between two characters when tested on day 1, and then repeated this on day 8. To quantify the simi-
larity in recall across sessions, we calculated the correlation between details recalled by the same participants across sessions with different participants who performed at the same level on the first session (calculated separately for each video). The mean within-participant correlation between details recalled on day 1 and day 8 was $r = 0.79$ (SD = 0.11), whereas the mean between-participant correlation was $r = 0.27$ (SD = 0.19), with the difference between these values being highly significant ($t_{(20)} = 11.2, p < 0.001$).

Follow-up analyses revealed no significant differences in subsequent memory that were related to the content of the videos (e.g., “outside” vs “inside”; details available upon request).

Experiment 2

Behavioral data

This experiment investigated the brain regions involved in remembering the content of video clips. In total, there were 20 videos that were rehearsed while in the scanner on day 1 and 6 that were not. The mean number of details recalled on day 7 from the rehearsed videos was 8.50 (SEM = 0.35), whereas the mean number of details from the nonrehearsed videos was 2.65 (SEM = 0.31). This difference is highly significant ($t_{(15)} = 14.1, p < 0.001$), replicating Experiment 1. A follow-up analysis investigated whether there was a relationship between recall vividness ratings from day 1 and subsequent memory scores on day 7. At an individual level, there was a significant ($p < 0.05$) correlation between vividness ratings given in the scanner on day 1 and subsequent memory on day 7 in 7 of 16 participants. To analyze the significance of the relationship between vividness and subsequent memory across the whole group, we Fisher-transformed the Pearson correlation coefficients for each individual and tested this against 0, using a one-sample $t$ test. This was significant ($t_{(15)} = 5.03, p < 0.001$), indicating that there was a robust relationship between vividness and subsequent memory at the group level.

Neuroimaging data: univariate analyses of brain regions involved in encoding and rehearsal

These first analyses aimed to identify regions independently involved in both encoding and rehearsal, separately comparing the encoding and rehearsal periods with baseline (the unmodeled ITI and rest periods; Fig. 3). This analysis revealed very extensive regions of activity, largely in “visual” regions, including visual cortex, the ventral visual stream (Mishkin et al., 1983; Goodale and Milner, 1992), and large regions of the thalamus during encoding. In addition to these areas, the superior parietal cortex bilaterally and the right temporal pole showed significant activity, as did the left middle frontal gyrus. There were also large regions showing significant “deactivations” during task performance compared with baseline, which are discussed below.

The second analysis identified regions more active during rehearsal compared with baseline (Fig. 4). This analysis revealed extensive regions of posterior portions of the frontal lobes, both medially and laterally, although activations were greater in the left hemisphere. The superior parietal lobe bilaterally and left posterior lateral temporal lobe were also identified in this analysis.

A rather similar, though not identical, network of regions was significantly deactivated during both encoding and rehearsal relative to rest. These regions include medial prefrontal cortex and posterior midline areas, such as the precuneus and retrosplenial cortex. This finding is consistent with numerous previous reports of so-called “default network” activity during rest periods (Buckner et al., 2008).

Planned follow-up analyses of the univariate contrasts investigated whether BOLD activity during encoding and rehearsal periods of each video correlated with the number of details subsequently recalled for those videos. However, no brain regions showed a significant effect.

Multivariate analyses of memory reinstatement during rehearsal

We used representational similarity analyses (Kriegeskorte et al., 2008) to identify areas where the spatial pattern of activity across local groups of voxels was greater during encoding and rehearsal of the same video clip versus encoding and rehearsal of different video clips (see Fig. 6; Table 1). That is, we looked for regions in which the pattern of BOLD activity when encoding a particular video is reinstated to some extent when rehearsing that particular video. This analysis identified the medial parieto-occipital cortex (posterior cingulate cortex including the retrosplenial cortex, and precuneus), angular gyrus, and the posterior portion of the middle temporal gyrus extending into the parahippocampal gyrus and hippocampus on the left (Fig. 5).

It is interesting to note that the region of medial parietal cortex identified in the RSA overlaps considerably with areas showing deactivation during encoding and retrieval compared with rest.
A week after watching the videos, participants in Experiment 1 recalled approximately twice as many details about videos that had been actively rehearsed than those that had not been rehearsed. Was this simply because there was no (passive) rehearsal period for these other memories to consolidate in? We believe that this is unlikely to be the sole explanation. In a study of the effect of 10 min of wakeful rest on memory retention over a week, the improvement in recall was ~11 story units versus just <9 story units (Dewar et al., 2012). This contrasts with our study where the improvement due to rehearsal was ~10 details versus 5.

It is possible that rehearsal prompted the participants to decide upon a version of what happened in the video, and it was this version that was recalled later. This proposal is supported by the observation that the recalled descriptions were highly similar over time, sometimes repeating exactly the same phrases. Others have noted that individuals commonly retrieve their own descriptions of events rather than remember the events themselves (Williams et al., 2008).

It has been argued (e.g., Winocur and Moscovitch, 2011) that memories for events undergo transformation over time, changing from being episodic and context-specific to semantic or schematic. If our participants were creating a “story” of what happened, then this representation of the memory would necessarily be rather fixed and inflexible, which are characteristics of semantic memories rather than episodic memories (Tulving, 1972; Cermak, 1984). Our results suggest that rehearsal of events might accelerate this transformation process. We note that our participants’ descriptions remained highly detailed, which is contrary to the notion that semanticized memories are generic in nature (Winocur and Moscovitch, 2011), but is consistent with the observation that even very densely amnesic patients are often able to recall some stories from their pasts in considerable detail, although such anecdotes are typically repeated verbatim on each occasion (e.g., Cermak, 1984; Steinworth et al., 2005).

There are alternative explanations for why active rehearsal boosts recall. First, rehearsal of a subset of the videos may inhibit consolidation of the nonrehearsed set and this inhibition may result in unrehearsed videos being largely forgotten (“retrieval-induced forgetting”) (Anderson et al., 2000; Wimber et al., 2015). Although this explanation does not explain why active rehearsal is such a good method for retention of detail over long periods (Roediger and Karpicke, 2006), it may explain why recall for nonrehearsed videos was so poor. A second mechanism that might serve to stabilize the memories is the strengthening of a hippocampal-dependent “episodic” representation of the events. This is discussed further below.

Using RSA, Experiment 2 identified a network of regions where patterns of BOLD activity elicited during the encoding of a video were reinstated during active rehearsal of that specific video (Fig. 6; Table 1). This network included the hippocampus and posterior midline regions (posterior cingulate, retrosplenial cortex, and precuneus). These regions are all strongly implicated in episodic mem-

![Figure 4. Brain regions involved in memory rehearsal compared with rest periods. Orange regions were more active when rehearsing the videos. Blue regions were deactivated during rehearsal periods compared with rest. Increased opacity of the color corresponds to higher t values. Regions significant at p < 0.001 (uncorrected for multiple comparisons) are outlined in black.](image-url)
ory processes (e.g., Aggleton and Brown, 1999; Eichenbaum, 2001). Therefore, the results support the proposal that active rehearsal not only enables a putative transformation process to take place, but also strengthens the episodic representation of the memory. Moreover, within posterior midline regions, the strength of representational similarity in the posterior cingulate correlated with the number of details recalled from each video a week after scanning. Given that the rehearsal period appears to be critical for robust memory consolidation (Experiment 1) (see also Rodiger and Butler, 2011), this finding suggests that the posterior cingulate plays a crucial role in active consolidation of complex memories.

Posterior midline structures have long been associated with memory and visual imagery in humans (Rudge and Warrington, 1991; Fletcher et al., 1995; Wagner et al., 2005) and spatial memory in rodents (Sutherland et al., 1988; Vann and Aggleton, 2004). They are also a central component of the “default network” of brain regions that are commonly more active during rest periods compared with task periods (Shulman et al., 1997; Spreng et al., 2009) and have been associated with processing underpinning “self-projection” (Buckner and Carroll, 2007) and “scene construction” (Hassabis and Maguire, 2007).

A related specific computational role has been proposed for the retrosplenial cortex and precuneus (Burgess et al., 2001; Byrne et al., 2007): that the retrosplenial cortex translates between egocentric and allocentric representations of an environment and, together with the precuneus, acts as a buffer for this information, allowing it to form a visuospatial mental imagine. Consistent with this, the retrosplenial cortex codes for both imagined location and imagined heading direction when humans visualize spatial scenes (Marchette et al., 2014) and activity in this region relates to mental rotation of viewpoint (Lambrey et al., 2008). Importantly, this model predicts that common representations will be formed at encoding and retrieval. It is likely, therefore, that RSA in posterior midline regions between encoding and rehearsal is partly due to reinstatement of visual representations created during encoding.

Previous studies have shown persistence of, or reinstatement of, patterns of activity in the hippocampus and posterior midline regions during tests of object-scene or object-face associations (Staresina et al., 2013; Tambini and Davachi, 2013). In these

<table>
<thead>
<tr>
<th>Region</th>
<th>Size (voxels)</th>
<th>Peak x</th>
<th>Peak y</th>
<th>Peak z</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSA between encoding and rehearsal of the same video</td>
<td>4579</td>
<td>-3</td>
<td>-55</td>
<td>20</td>
<td>9.2</td>
</tr>
<tr>
<td>Bilateral precuneus/posterior cingulate cortex</td>
<td>3747</td>
<td>-45</td>
<td>-59</td>
<td>8</td>
<td>8.86</td>
</tr>
<tr>
<td>Left middle temporal gyrus/parahippocampal gyrus/inferior parietal lobule/hippocampus/superior temporal sulcus</td>
<td>3606</td>
<td>45</td>
<td>-53</td>
<td>20</td>
<td>12.33</td>
</tr>
<tr>
<td>Right superior temporal gyrus/angular gyrus/inferior parietal lobule</td>
<td>757</td>
<td>35</td>
<td>-39</td>
<td>-6</td>
<td>7.34</td>
</tr>
<tr>
<td>Right parahippocampal gyrus</td>
<td>304</td>
<td>39</td>
<td>9</td>
<td>18</td>
<td>5.96</td>
</tr>
<tr>
<td>Right insula/middle frontal gyrus/white matter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modulation of regions showing RSA between encoding and rehearsal by the number of details subsequently recalled</td>
<td>Bilateral posterior cingulate cortex/precuneus</td>
<td>224</td>
<td>-3</td>
<td>-41</td>
<td>32</td>
</tr>
</tbody>
</table>

*All clusters are significant when whole-brain family-wise error was corrected for cluster size.*
studies, the stimuli were pairings of two pictures, and these stimuli could be retrieved as a single mental image. Nevertheless, our finding of representational similarity in posterior midline regions is likely to be driven by more than simply reinstating a single visual percept. Videos depict an unfolding sequence of actions that must be interpreted, with reference to prior knowledge or “schemas,” to create a coherent representation of the whole event (for relevant evidence from classic studies using complex memory, see Bartlett, 1932; Bransford and Johnson, 1972; Bower et al., 1979; Brewer and Treynor, 1981). Although prior knowledge was not quantified or manipulated in our study, the descriptions of the videos frequently referred to external information (for example, one video was described as being “like the film ‘Twilight,’” and in another, a character acted “like James Bond”). Critically therefore, our effect is likely to reflect the reinstatement of a coherent representation of the content of the video. It is interesting to note that, despite the clear evidence for reinstatement in medial parietal regions, overall BOLD activity in these regions did not significantly increase during encoding or retrieval; indeed, activity decreased in several areas (compare Figs. 4, 5 with Fig. 6). This is an important example of how multivariate analysis can identify stimulus-specific processing even in the absence of a positive univariate effect.

The posterior cingulate cortex has been identified as a candidate region for linking episodic and semantic information (e.g., Binder et al., 2009). For example, Maguire et al. (1999) scanned participants performing a reading comprehension and memory task where prior knowledge about the stories was manipulated. The authors concluded that the posterior cingulate cortex played a role in linking the narrative information with prior knowledge. Our results are compatible with this conclusion. It should be noted that our MRI findings relate to the degree of reinstatement of activity during encoding and the very early stages of episodic memory consolidation. Although Experiment 1 demonstrated that memory recall can be very similar across periods of days and weeks, it remains an open question whether recall reinstates similar patterns of brain activity after these delays.

In this paper, we have shown the significant effect of active rehearsal on retention of episodic detail over the period of a week, and that participants’ descriptions of the videos were highly similar across repeated recall sessions. We also showed that the pattern of brain activity during encoding of the videos was reinstated during retrieval throughout medial temporal and posterior midline regions, and that the degree of reinstatement in the posterior cingulate cortex correlated with recall of the videos following a delay of a week. Thus, in addition to their known role in recollection and visual imagery, these findings suggest that the posterior cingulate plays a crucial role in integrating incoming episodic experience with existing knowledge to create a coherent representation of the event (related to the ideas of schemas). Reinstatement of this representation aids consolidation by strengthening the associations between episodic details as well as more general schematic information, resulting in a memory that is resistant to forgetting, but rather inflexible and semanticized.

**References**


Bartlett FC (1932) Remembering. London: Cambridge UP.


Skaggs EB (1925) Further studies in retroactive inhibition. Psychol Monogr 34:1–60.