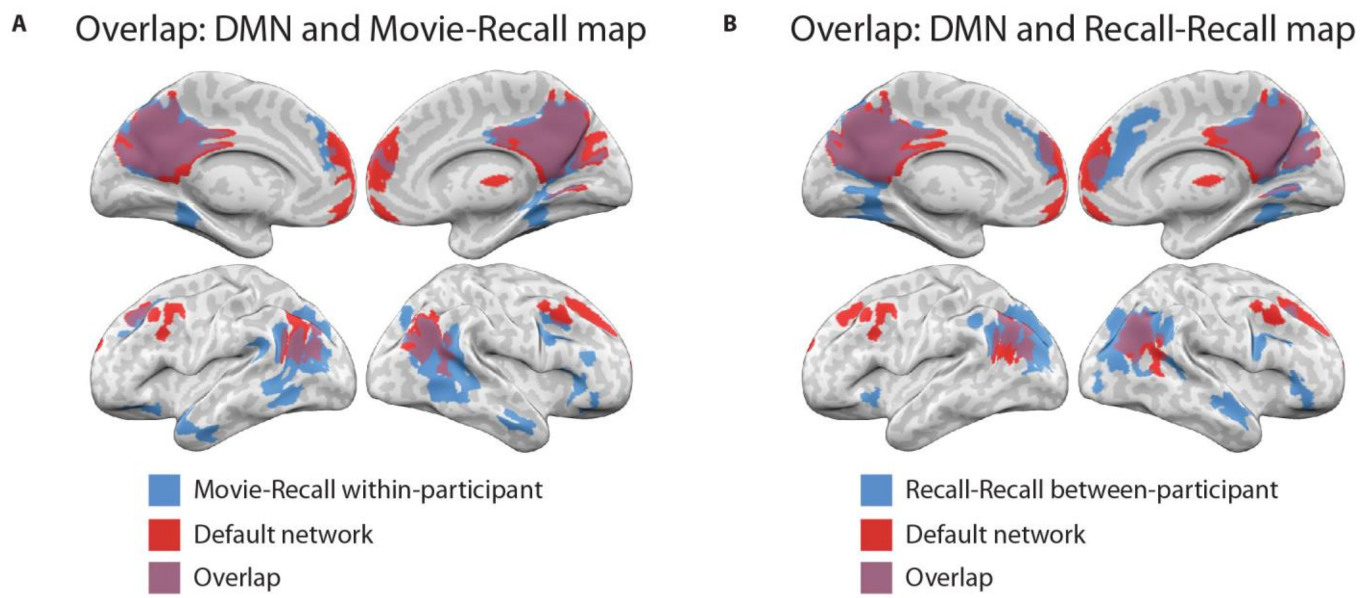
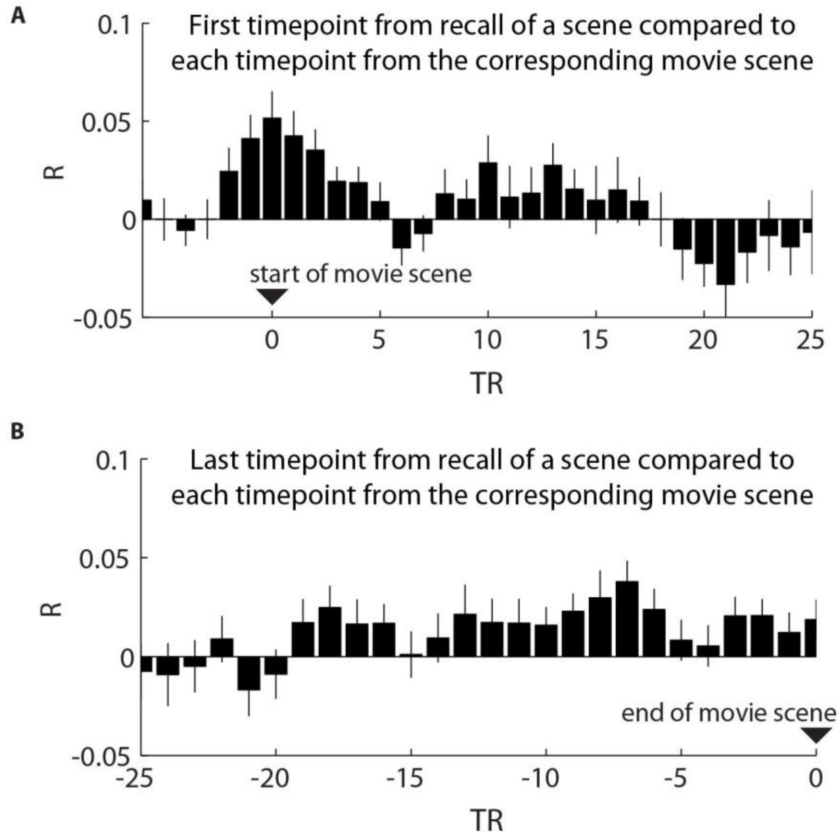


Supplementary Figure 1

Overlap between default mode network (DMN) and movie/recall maps.

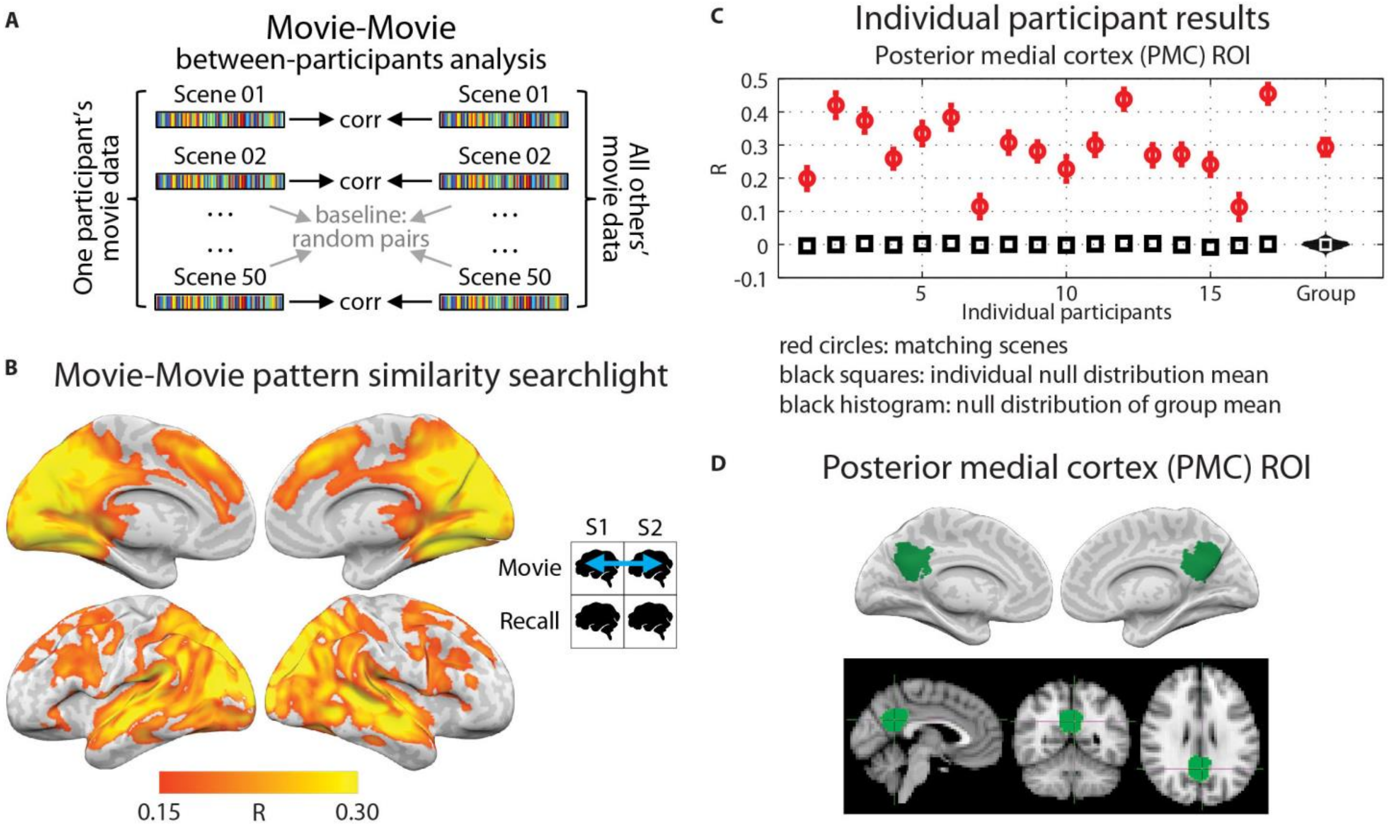
We defined the DMN for each individual using the posterior medial cortex ROI as a seed for functional connectivity during the first scan of the movie (23 minutes), thresholded at $R > 0.4$; a group-level DMN map was then created by averaging across participants. While the DMN is typically defined using resting state data, it has been previously demonstrated that this network can be mapped either during rest or during continuous narrative with largely the same results. See Table S3 for overlap calculations for all searchlight maps (from Figs. 2B, 2E, 3B, and 7B). Note that this DMN definition procedure is independent from the calculations of the searchlight maps, because functional connectivity is calculated across time (and during movie only), while the searchlight analyses were spatial pattern comparisons (between movie and recall). **A)** Overlap between the group DMN and the within-participant movie-recall searchlight map from Fig. 2B. 39.7% of this map falls within the DMN. **B)** Overlap between the group DMN and the between-participant recall-recall map from Fig. 3B. 50.7% of this map falls within the DMN.



Supplementary Figure 2

Within-participant pattern reinstatement at a finer temporal scale.

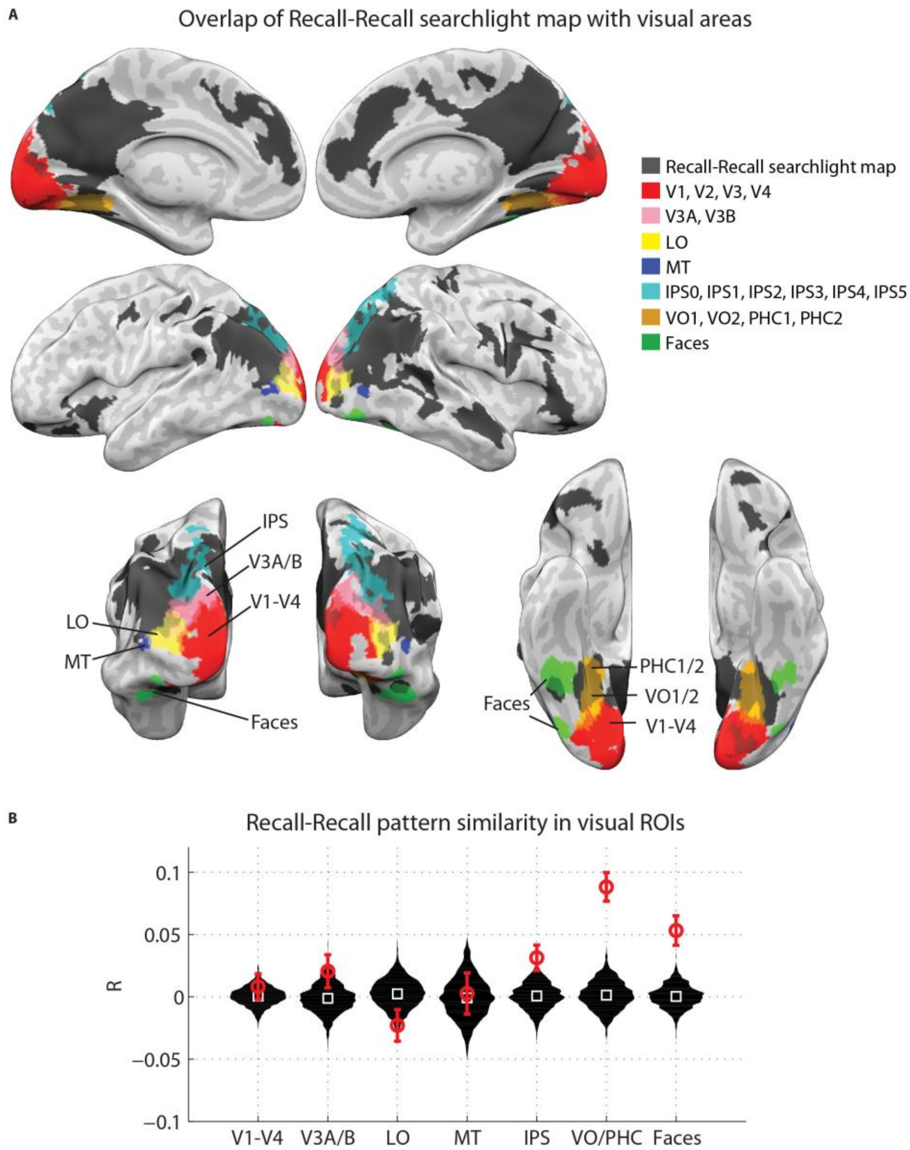
While averaging at the scene level was effective for observing neural reinstatement, the behavior of mnemonic recollection that we observed unfolded over time at a finer scale than the scene level. For example, participant 8 used 131 words over 67 seconds to describe scene 13. Here, we further examined reinstatement effects at individual timepoints. **A)** For each scene for a given participant, we compared the pattern of activity at each timepoint in the movie scene with the pattern from the *first* timepoint of recall of that scene in the posterior medial cortex ROI. These correlation values were averaged across all scenes and all participants. Correlations with the earliest timepoints of encoding scenes tended to be higher than correlations with later timepoints, suggesting sub-scene level specificity of reinstatement. Error bars represent standard error across subjects. **B)** For each scene for a given participant, we compared the pattern of activity at each timepoint in the movie scene with the pattern from the *last* timepoint of recall of that scene in the posterior medial cortex ROI. These correlation values were averaged across all scenes and all participants. Error bars represent standard error across subjects.



Supplementary Figure 3

Pattern similarity between participants during the movie.

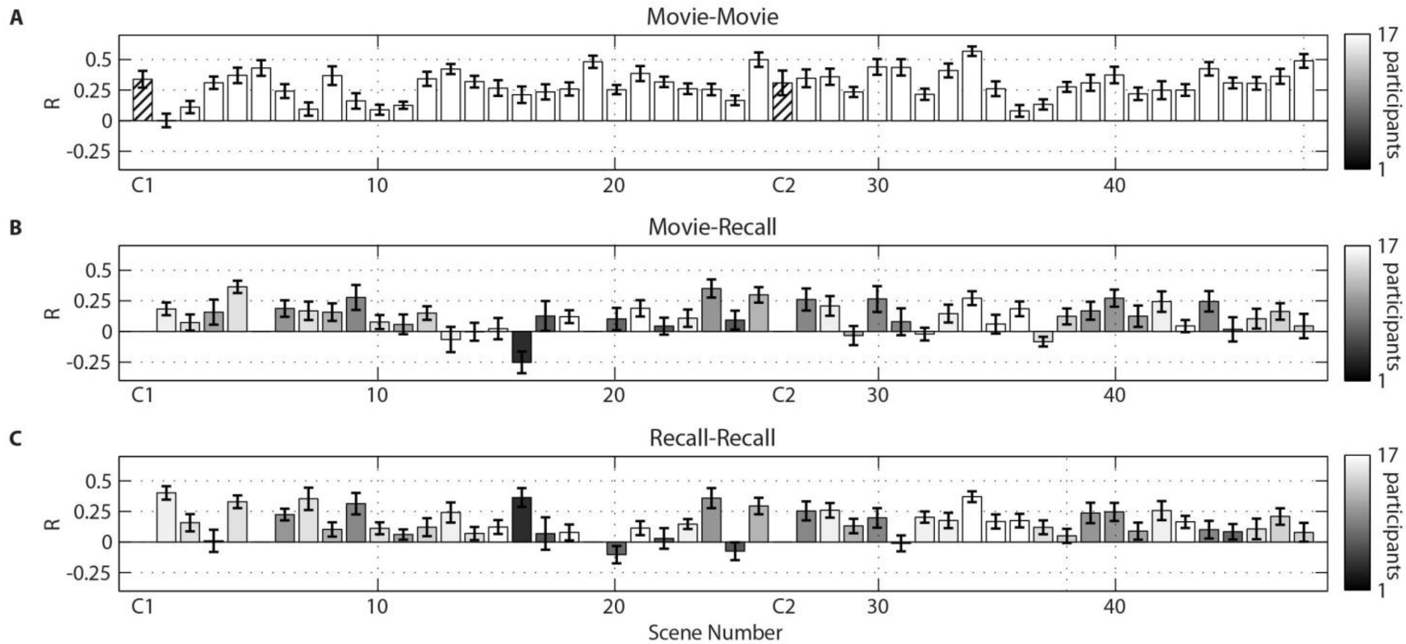
A) Schematic for between-participant movie-movie analysis. BOLD data from the movie were divided into scenes, then averaged across time within-scene, resulting in one vector of voxel values for each movie scene and each recalled scene. Correlations were computed between matching pairs of movie scenes *between* participants. **B)** Searchlight map showing correlation values for across-participant pattern similarity during the movie. Searchlight was a 5x5x5 voxel cube. **C)** Correlation values for all 17 participants in independently-defined PMC (posterior medial cortex). Red circles show average correlation of matching scenes and error bars show standard error across scenes; black squares show average of the null distribution for that participant. At far right, the red circle shows the true participant average and error bars show standard error across participants; black histogram shows the null distribution of the participant average; white square shows mean of the null distribution. **D)** Posterior medial cortex region of interest, cluster in the "dorsal default mode network" set (http://findlab.stanford.edu/ functional_ROIs.html).



Supplementary Figure 4

Overlap of recall–recall map with visual areas.

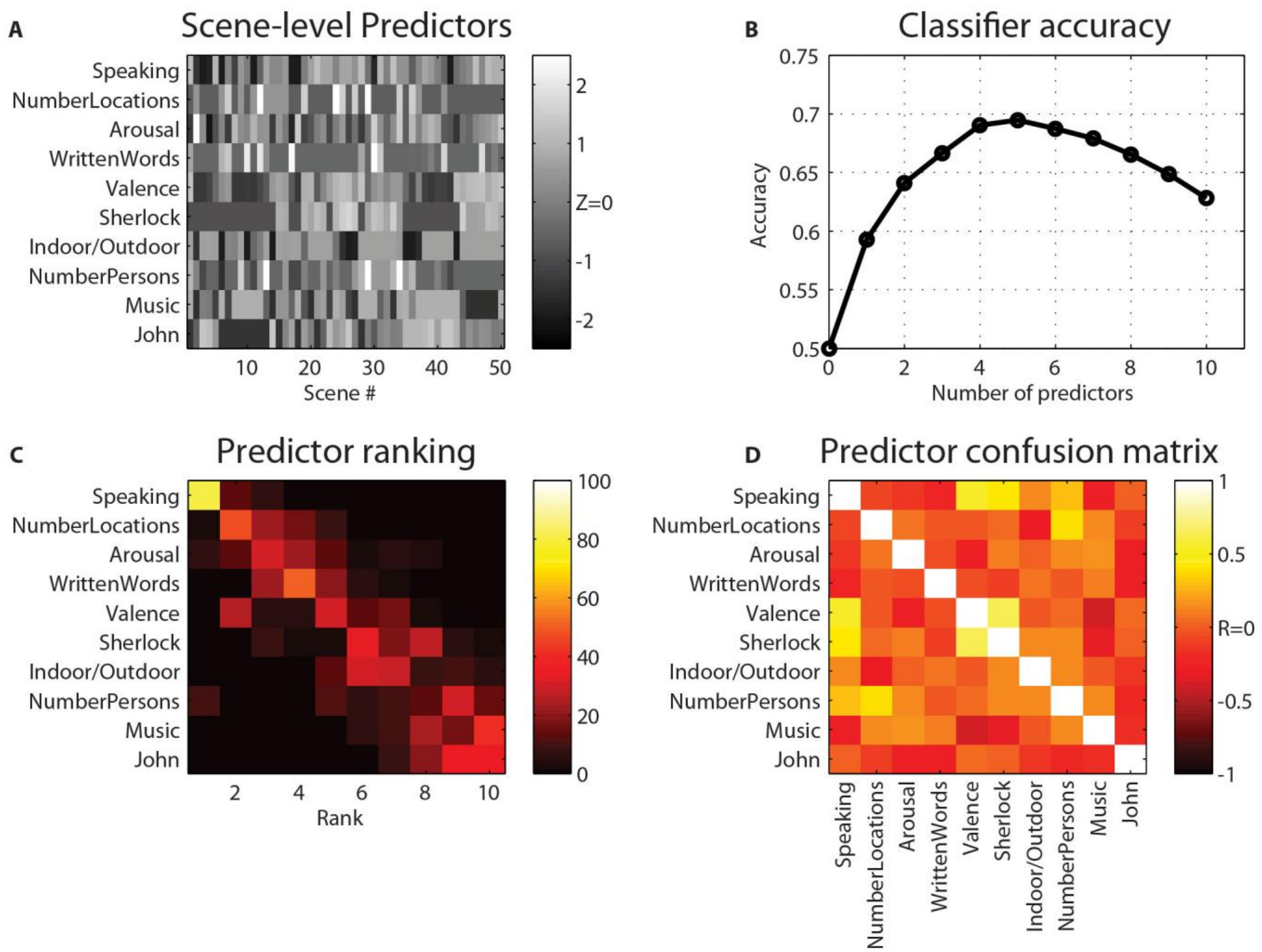
To what extent did spoken recollection in this study engage visual imagery? Movie-recall reinstatement effects were not found in low level visual areas, but instead were located in high level visual areas, and extensively in higher order brain regions outside of the visual system. Our observation of reinstatement in high level visual areas is compatible with studies showing reinstatement in these regions during cued visual imagery. The lack of reinstatement effects in low-level areas may be due to the natural tendency of most participants to focus on the episodic narrative (the plot) when recounting the movie, rather than on fine visual details. See also *Methods: Visual imagery.* **A)** In gray, brain areas where recollection patterns were significantly similar across participants (map from Fig. 3B). In other colors, commonly studied visual areas. Retinotopic visual areas were taken from a published probabilistic atlas (Wang et al., 2014, *Cereb. Cortex*). Face-selective areas were generated using Neurosynth (Yarkoni et al., 2011, *Nat. Methods*). **B)** For each of the visual area ROIs shown in [A], similarity of scene-level recollection patterns was calculated between participants in the same manner as Fig. 3. Statistical significance was determined by shuffling scene labels to generate a null distribution of the participant average. For each region, red circle shows the true participant average, error bars show standard error across participants; black histogram shows null distribution of the participant average; white square shows mean of the null distribution. In low-order visual regions, recall-recall pattern similarity was not different from chance; however, significant recall-recall pattern similarity was observed in higher-order visual regions (VO/PHC and face-selective areas).



Supplementary Figure 5

Between-participants pattern similarity in PMC, scene by scene.

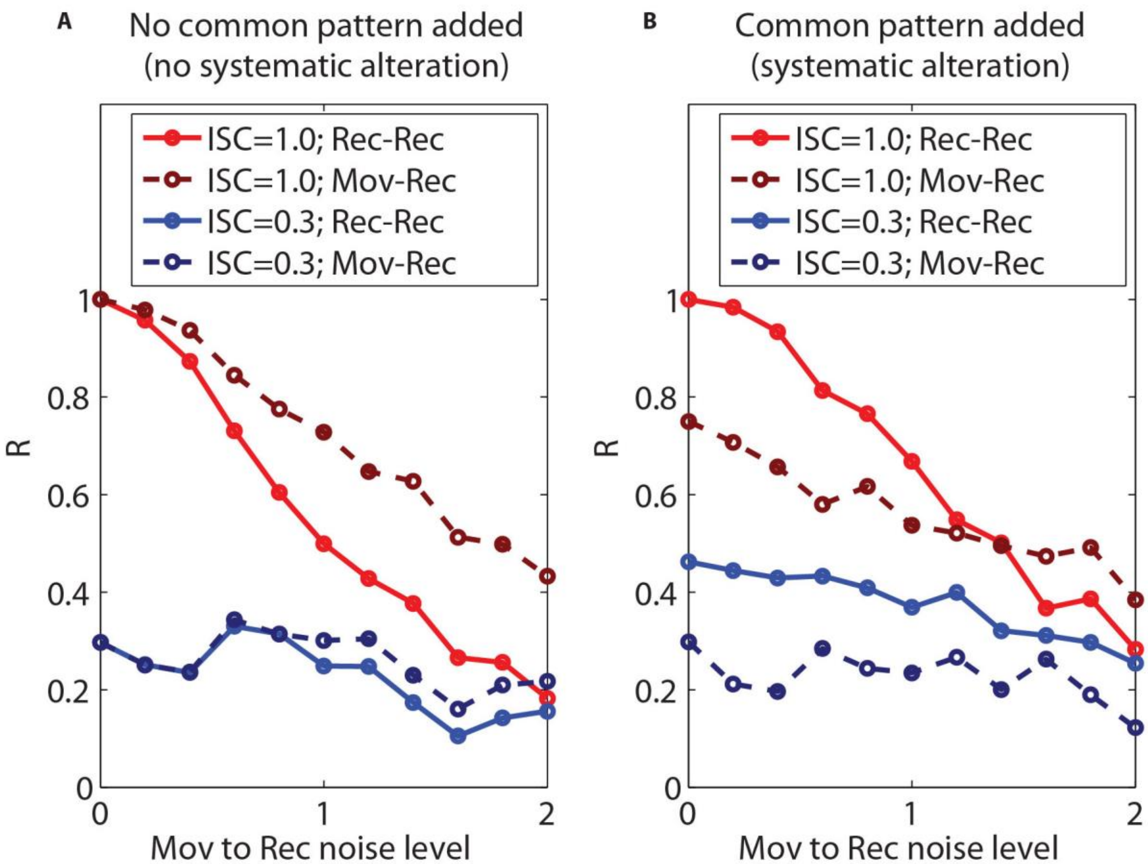
A) Between-participants movie-movie correlation values for 50 individual scenes in the posterior medial cortex (PMC) ROI (same ROI as Fig. 2C, 2F, 3C). For each scene, each participant's movie pattern from that scene was compared to the pattern from the corresponding movie scene averaged across the remaining participants. The bars show the average across participants for each scene. Error bars represent the standard error across participants. Striped bars indicate introductory video clips at the beginning of each functional scan (see Methods). **B)** Between-participants movie-recall correlation values for individual scenes in the PMC ROI (46 scenes were recalled by two or more participants). **C)** Between-participants recall-recall correlation values for individual scenes in the PMC ROI.



Supplementary Figure 6

Encoding model.

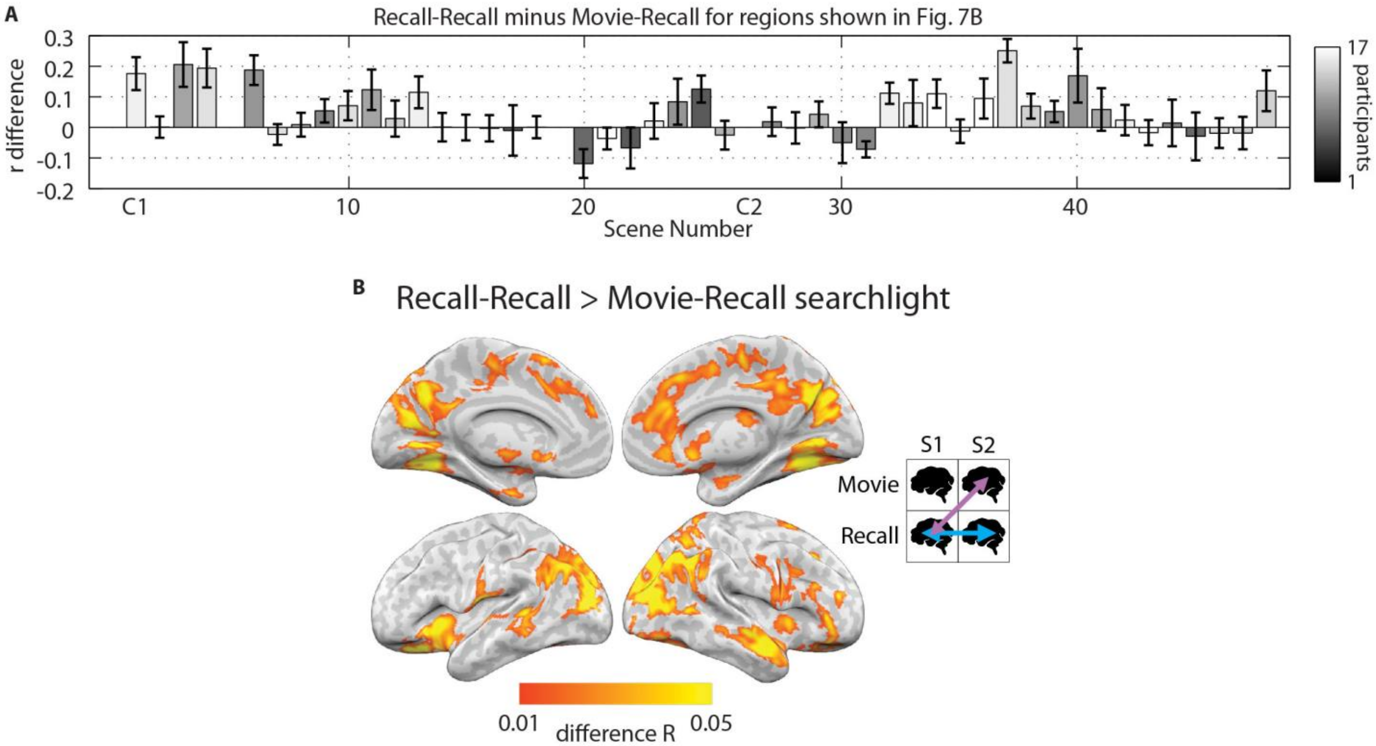
To explore what semantic information is represented in the shared neural patterns that supports our ability to discriminate patterns of activity between scenes, we constructed an encoding model to predict neural activity patterns from semantic content. See Supp. Note 4 for additional details. **A)** Detailed semantic labels were generated by an independent coder: 1000 time segments spanning the entire movie stimulus, and 10 labels for each segment. A score was derived for each of the 50 scenes for each label, creating 50-element predictor vectors. It should be noted that the list of 10 labels is by no means comprehensive, and is intended merely to serve as a starting point for future analyses. **B)** Predicted patterns in PMC were generated by regressing voxel activity on label values, and scene-level classification accuracy was assessed using a hold-2-out procedure validated across 100 combinations of independent groups (N=8 and N=9). Classification accuracy increased as predictors (labels) were added to the model, peaking at 69.5% with five predictors (chance level 50%). **C)** Predictors were ranked according to how much they improved accuracy for each of the 100 combinations; the most successful predictor was the proportion time during a scene that speech was present (*Speaking,* ranked first for 80% of combinations), followed by the number of different locations visited during a scene (*NumberLocations*, ranked 2nd for 48%), arousal (Arousal, ranked 3rd for 31%), proportion time that written words were present (*WrittenWords*, ranked 4th for 51%), and valence (*Valence*, ranked 5th for 31%). The number of persons in a scene (*NumberPersons*) was ranked first for 10% of combinations. **D)** Confusion matrix for the 10 predictors. Note that when two predictors are correlated, one may dominate in the predictor rankings.



Supplementary Figure 7

Simulation of movie-to-recall pattern alteration.

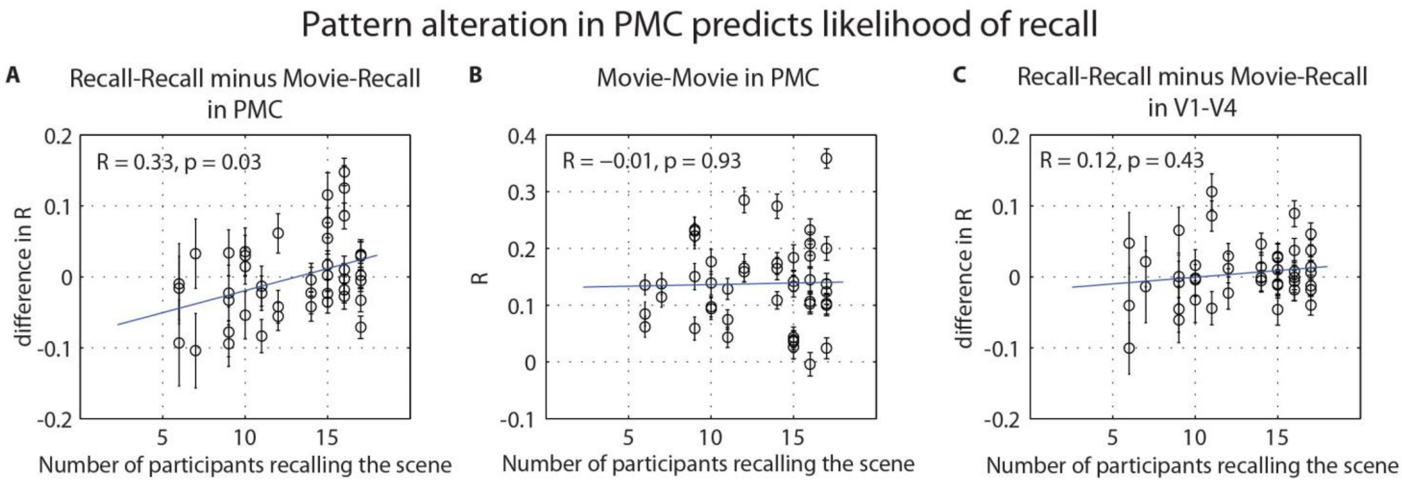
In this simulation, five 125-voxel random patterns are created (five simulated subjects) and random noise is added to each one, such that the average inter-subject correlation is R=1.0 (red lines) or R=0.3 (blue lines). These are the “movie patterns”. Next, we simulate the change from movie pattern to recall pattern by 1) adding random noise (at different levels of intensity, y-axis) to every voxel in every subject to create the “recall patterns”, which are noisy versions of the movie pattern; and 2) adding a common pattern to each movie pattern to mimic the “systematic alteration” from movie pattern to recall pattern, plus random noise (at different levels of intensity, x-axis). We plot the average correlation among the five simulated subjects’ recall patterns (Rec-Rec), as well as the average correlation between movie and recall patterns (Mov-Rec). **A)** Results when no common pattern is added, i.e., the recall pattern is merely the movie pattern plus noise (no systematic alteration takes place): Even as noise varies at the movie pattern stage and at the movie-to-recall change stage, similarity among recall patterns (*Rec-Rec*, solid lines) never exceeds the similarity of recall to movie (*Mov-Rec*, dotted lines). **B)** Results when a common pattern is added to each subject’s movie pattern, in addition to the same levels of random noise, to generate the recall pattern. Now, it becomes possible (even likely, under these simulated conditions) for the similarity among recall patterns (*Rec-Rec*, solid lines) to exceed the similarity of recall to movie (*Mov-Rec*, dotted lines). In short, when the change from movie to recall involves a systematic alteration across subjects, recall patterns may become more similar to each other then they are to the original movie pattern. Note that the similarity of the movie pattern to each other (movie-movie correlation) does not impact the results. See *Methods: Simulation of movie-to-recall pattern alteration.*



Supplementary Figure 8

Scene-by-scene difference of recall–recall minus movie–recall in regions shown in Figure 7b.

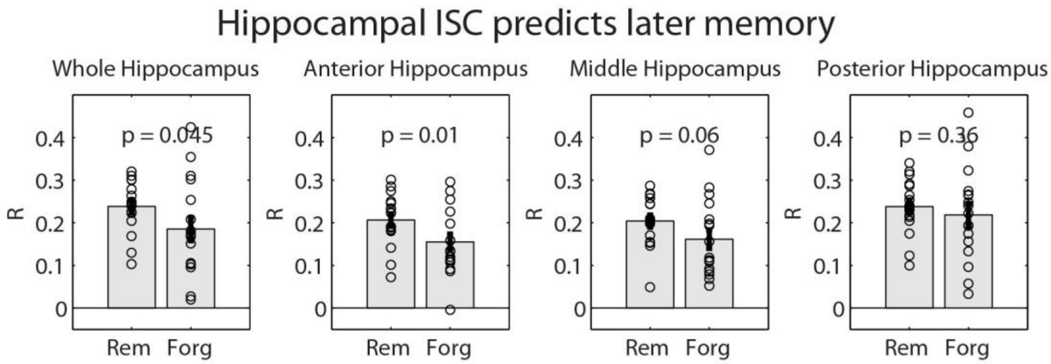
A) In Fig. 7B we plotted brain regions in which participants' recollection activity patterns were more similar to the recollection patterns in other individuals than they were to movie patterns ("neural alteration" effect). Here we show the results broken down scene-by-scene in the same regions. Error bars represent the standard error across participants. **B)** Recall-recall minus movie-recall difference values thresholded at 0.01.



Supplementary Figure 9

Subsequent memory analyses.

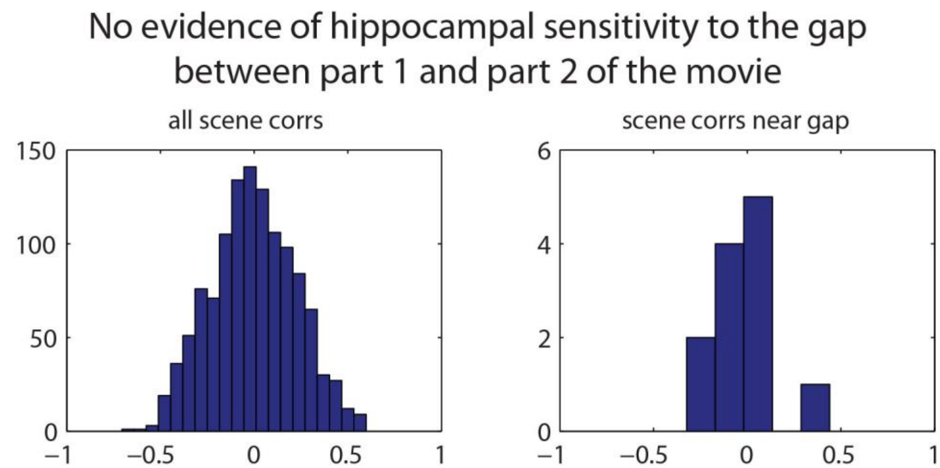
To examine how the systematic alteration of neural activity from movie to recall might be related to memorability, we divided scenes into remembered and forgotten for each participant. For each scene, the number of participants who had successfully recalled that scene was counted. We then extracted data from the PMC ROI and calculated the pairwise between-participants correlation during recall (same analysis as in Fig. 3A-C, except pairwise), the pairwise between-participants correlation between movie and recall (same analysis as in Fig. 2D-F, except pairwise), and used the difference as the degree of neural alteration (recall-recall similarity minus movie-recall similarity), at the scene level. Pairwise comparisons were used because the mean value of pairwise correlations is not affected by the number of participants (and the number of participants was different across data points in this analysis). **A)** We calculated Spearman's rank correlation for the number of participants who successfully recalled each scene vs. the average degree of neural alteration for each scene. The magnitude of neural alteration was significantly related to how many participants remembered that scene ($R = 0.33$, $p = 0.03$). In other words, the more that a given movie scene pattern was altered in systematic manner across subjects between perception and recall, the more likely that scene was to be remembered. **B)** A control analysis in PMC showing that between-participants movie-movie pattern similarity was not predictive of the likelihood of recall ($R = -0.01$, $p > 0.9$). **C)** A control analysis showing that the degree of neural alteration (i.e., recall-recall minus movie-recall) in early visual areas V1-V4 was not predictive of the likelihood of recall ($R = 0.12$, $p = 0.43$, same ROI as Fig. S4).



Supplementary Figure 10

Hippocampal inter-subject correlation (ISC).

We examined hippocampal contributions to recall success. During movie viewing, we calculated the correlation between a given participant's hippocampal timecourse (using an anatomically-defined whole hippocampus ROI) and the average hippocampal timecourse of all other participants, for individual scenes (i.e., the inter-subject correlation (ISC) for each scene). For each participant, scenes were binned by whether they were later remembered or forgotten. ISC was significantly greater for remembered scenes than forgotten scenes (left panel; 2-tailed paired t-test across participants, $t = 2.17$, $p = 0.045$), complementing previous results linking ISC in parahippocampal cortex to later recognition memory (Hasson et al., 2008, *Neuron*). The same analysis is shown for the hippocampus ROI split into anterior, middle, and posterior sections (second, third, and fourth panels from the left). A repeated-measures ANOVA with region (anterior, middle, posterior) and memory (remembered, forgotten) as factors revealed significant main effects of region $F(2,32) = 12.02$, $p < 0.0005$ and of memory $F(1,16) = 4.98$, $p = 0.04$, but not a significant region x memory interaction $F(2,32) = 1.69$, $p = 0.2$.



Supplementary Figure 11

No evidence of hippocampal sensitivity to the gap between part 1 and part 2 of the movie.

Evidence from time cells in the rodent hippocampus might predict that the hippocampus would be sensitive to the gap between the first segment and the second. In order to explore this question, we examined recall patterns in the hippocampus for the 3 scenes just before and 3 scenes just after the gap, specifically asking where the correlations of these patterns with their corresponding movie scenes fell in the distribution of all such movie-recall scene correlations. The left panel shows the distribution of movie-vs-recall pattern correlations for all 50 scenes (averaged across subjects), and the right panel shows the distribution of movie-vs-recall pattern correlations for the 3 scenes just before and 3 scenes just after the gap. There does not appear to be anything unusual about the scenes near the gap, in terms of their pattern similarity to the corresponding movie scenes (the near-gap values fall near the middle of the distribution). Thus, in this analysis, we did not find any evidence to support the hypothesis that the hippocampus is sensitive to the gap between part 1 and part 2 of the movie during recall.

Supplementary Table S1. Behavior during spoken recall.

Subject	# Scenes Recalled (out of 50)	Total Time (minutes)	Total # Words	# Scenes Out of Order
01	27	13.0	1705	1
02	24	22.0	1903	4
03	32	17.7	2478	2
04	33	10.8	1681	8
05	32	11.6	1501	0
06	39	21.7	1843	4
07	30	16.4	1136	14
08	39	24.6	3221	5
09	28	16.9	1680	5
10	40	20.8	2592	3
11	34	12.7	2019	6
12	38	29.5	4138	3
13	47	43.9	5962	13
14	38	30.6	3576	10
15	27	20.2	2029	3
16	37	21.2	2770	11
17	39	35.0	4939	9
Mean	**34.4 (6.0)**	**21.7 (8.9)**	**2657.2 (1323.6)**	**5.9 (4.2)**

Supplementary Table S2. Examples of four different participants' descriptions of two movie scenes.

Scene 13 Duration of original scene: 2 min 11 s	**S06:** So then later, at some point in the show, there is a press conference with the Police Chief. The press is asking about the deaths. The chief says they are suicides but feels that they are some how connected. Everyone receives a text that says wrong. The police chief goes on to talk about the deaths.
	S08: So we switch to this press conference where the chief of police, or the head of the police station, is answering these questions regarding the fact that these suicides have happened and there are these mass texts that are going on. During this press conference he is talking about the fact that all these suicides are connected in some way because of the similarity of the fact that they have all taken this medicine and they all are in places that they don't normally belong. And these text messages start occurring during the press conference as he's explaining the case that simply say "wrong" to all of the reporters. And he gets a text message that says he knows where to find him and its signed SH, so from Sherlock Holmes.
	S15: Yeah so then there's a scene with a detective and he's giving a press conference to some reporters, and its the detective and a woman who he's with. And during the press conference the reporters are asking him questions and as he's answering them, periodically there are texts that pop up and the say "Wrong!" and the reporters are all kind of taken aback. And people giving the press conference instruct the reporters to ignore the texts, but as the press conference goes along, it happens three different times and the third time the head inspector gets a text and its an invitation from SH to have him come seek his help with apparent suicides. During the press conference the reporters are asking if the suicides are actually, or like how they could be being investigated by homicide detectives? The detective doesn't really know much but he's just saying that they think they're all linked because all the people kill themselves the same way and they're all in odd locations when they do it, and there's no suicide notes.
	S17: So then we get the press conference, and the guy's saying that he thinks that these suicides are linked, and then a reporter says but how could suicides be linked? that doesn't make any sense. And then the head detect-, the head of the, sergeant or whatever says something like, I don't know but we're investigating it. Then everyone in the room gets a text saying Wrong, and everyone's kind of confused about it, and he said, the woman beside him said If you just got a text please ignore it. And then everyone's asking more questions, and then the sergeant says something that apparently Sherlock Holmes thinks is wrong, because they all get a text saying it's wrong again.
Scene 36 Duration of original scene: 1 min 55 s	**S06:** Watson he walks home and he hears a pay phone ring when a third one rings he answers. The voice says look at security cameras across the street and then asks him to get in the car.
	S08: So he's walking down this main road and the phone booth next to him rings. And then he continues walking, kind of ignoring that, and another phone rings in a business as he's walking by. And then it stops as someone else comes to answer it. And then he keeps walking, there's a third one that's on his left, and its ringing. So he goes into the phone booth, he answers, and this voice on the phone directs him to see the fact that there are these multiple security cameras that have been kind of tracking where he is. And it tells him to get into this car, to come where he's going.
	S15: So Dr. Watson is on his way back and he's passing telephone booths and first one rings, he ignores it, second one rings, he ignores it, third one rings, and he finally answers, there's a voice on the other end that tells him to look at security cameras in the area and each one gets averted. And then a car pull up and he says get in, I don't need to threaten you, you already know.
	S17: So then we see, right, kinda like across the street from the building where that lady was found dead, there was this red telephone booth. And it was ringing. Which is weird, and then Watson looks at it and hears it ringing but chooses not to answer it. So then he just ends up walking, we see him walking like on a sidewalk, busy sidewalk, and he's looking in a shop window. And there's a telephone ringing again. And then a store clerk is about to answer it but then doesn't. So then he has like a confused face on. And then he keeps walking and then on the third ringing at yet another phone booth, he is again confused and so okay well I'm just gonna answer it. So he goes in and answers it, and this man, whose voice we haven't heard yet, says, Do you see the camera to your left? And so then he kinda, he says Who is this? and the guys says Do you see this camera. So he looks up, and, the camera, so then we see from the camera's view, we're looking at Sherlock in the telephone booth. And then the camera pans over to the street. And then the person on the phone says, Do you see the camera across the street? And then we see the camera moving across the street. We see the camera on Watson's right. And the guy asks Do you see the camera. And so Watson notices all these cameras. And we get a shot of, in each corner of the screen, a view of the street where Watson is. So then the guy says, ok there's gonna be a car, it's gonna pick you up, do you understand the situation you're in? So get in the car, it's gonna pick you up. So then Watson understands the situation he's in and he gets in the car.

Supplementary Table S3. *Overlap between default mode network (DMN) and movie/recall maps*. We defined the DMN for each individual using the posterior medial cortex ROI as a seed for functional connectivity during the first scan of the movie (23 minutes), thresholded at R > 0.4; a group-level DMN map was then created by averaging across participants. See also Figure S1.

	Overlap with DMN (Jaccard index)	Percent of map that falls within the DMN
Within-subject movie-recall map (Fig. 2B)	0.3038	39.7%
Between-subject movie-recall map (Fig. 2E)	0.2794	34.4%
Between-subject recall-recall map (Fig. 3B)	0.2689	50.7%
Recall-recall > movie-recall map (Fig. 6B)	0.0926	26.3%