Shared memories reveal shared structure in neural activity across individuals

J. Chen[1,2,*,†], **Y.C. Leong**[3,†], **C.J. Honey**[4,5], **C.H. Yong**[5], **K.A. Norman**[1,2], and **U. Hasson**[1,2]

[1]Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey, USA

[2]Department of Psychology, Princeton University, Princeton, New Jersey, USA

[3]Department of Psychology, Stanford University, Stanford, California, USA

[4]Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, Maryland, USA

[5]Department of Psychology, University of Toronto, Toronto, Ontario, Canada

Abstract

Our lives revolve around sharing experiences and memories with others. When different people recount the same events, how similar are their underlying neural representations? Participants viewed a fifty-minute movie, then verbally described the events during functional MRI, producing unguided detailed descriptions lasting up to forty minutes. As each person spoke, event-specific spatial patterns were reinstated in default-network, medial-temporal, and high-level visual areas. Individual event patterns were both highly discriminable from one another and similar between people, suggesting consistent spatial organization. In many high-order areas, patterns were more similar between people recalling the same event than between recall and perception, indicating systematic reshaping of percept into memory. These results reveal the existence of a common spatial organization for memories in high-level cortical areas, where encoded information is largely abstracted beyond sensory constraints; and that neural patterns during perception are altered systematically across people into shared memory representations for real-life events.

*Correspondence to: janice@princeton.edu.
†These authors contributed equally to this work.

We tend to think of memories as personal belongings, a specific set of episodes unique to each person's mind. Every person perceives the world in his or her own way and describes the past through the lens of individual history, selecting different details or themes as most important. On the other hand, memories do not seem to be entirely idiosyncratic; for example, after seeing the same list of pictures, there is considerable correlation between people in which items are remembered[1]. The capacity to share memories is essential for our ability to interact with others and form social groups. The macro and micro-processes by which shared experiences contribute to a community's collective memory have been extensively studied across varied disciplines[2–5], yet relatively little is known about how shared experiences shape memory representations in the brains of people who are engaged in spontaneous natural recollection. If two people freely describe the same event, how similar (across brains) are the neural codes elicited by that event?

Human brains have much in common with one another. Similarities exist not only at the anatomical level, but also in terms of functional organization. Given the same stimulus, an expanding ring for example, regions of the brain that process sensory (visual) stimuli will respond in a highly predictable and similar manner across different individuals. This predictability is not limited to sensory systems: shared activity across people has also been observed in higher-order brain regions (e.g., the default mode network[6] [DMN]) during the processing of semantically complex real-life stimuli such as movies and stories[7–13]. Interestingly, shared responses in these high-order areas seem to be associated with narrative content and not with the physical form used to convey it[11,14,15]. It is unknown – at any level of the cortical hierarchy – to what extent the similarity of human brains during shared perception is recapitulated during shared recollection. This prospect is made especially challenging when recall is spontaneous and spoken, and the selection of details left up to the rememberer (rather than the experimenter), as is often the case in real life.

In memory, details may be lost or changed, motives may be re-framed, and new elements may be inserted. Although a memory is an imperfect replica of the original experience, the imperfection might serve a purpose. As demonstrated by Jorge Luis Borges in his story "Funes the Memorious"[16], a memory system that perfectly recorded all aspects of experience, without the ability to compress, abstract, and generalize the to-be-remembered information, would be useless for cognition and behavior. In other words, perceptual representations undergo some manner of beneficial modification in the brain prior to recollection. Therefore, memory researchers can ask two complementary questions: 1) to what extent a memory resembles the original event; and 2) what alterations take place between perceptual experience and later recollection. The first question has been extensively explored in neuroscience; many studies have shown that neural activity during perception of an event is reactivated to some degree during recollection of that event[17–19]. However, the laws governing alterations between percept and recollection are not well understood.

In this paper, we introduce a novel inter-subject pattern correlation framework that reveals shared memory representations and shared memory alteration processes across the brain. Participants watched a movie and then were asked to verbally recount the full series of events, aloud, in their own words, without any external cues. Despite the unconstrained nature of this behavior, we found that spatial patterns of brain activity observed during

movie-viewing were reactivated during spoken recall (movie-recall similarity). The reactivated patterns were observed in an expanse of high-order brain areas that are implicated in memory and conceptual representation, broadly overlapping with the DMN. We also observed that these spatial activity patterns were similar across people during spoken recall (recall-recall similarity) and highly specific to individual events in the narrative (i.e., discriminable), suggesting the existence of a common spatial organization or code for memory representations. Strikingly, in many high-order areas, which partially overlap with the DMN, we found that recall-recall similarity was stronger than movie-recall similarity, indicating that neural representations were transformed between perception and recall in a systematic manner across individuals.

Overall, the results suggest the existence of a common spatial organization for memory representations in the brains of different individuals, concentrated in high-level cortical areas (including the DMN) and robust enough to be observed as people speak freely about the past. Furthermore, neural representations in these brains regions were modified between perceptual experience and memory in a systematic manner across different individuals, suggesting a shared process of memory alteration.

RESULTS

Spontaneous spoken recall

Seventeen participants were presented with a 50-minute segment of an audio-visual movie (BBC's "Sherlock") while undergoing functional MRI (Fig. 1A). Participants were screened to ensure that they had never previously seen any episode of the series. They were informed before viewing that they would later be asked to describe the movie. Following the movie, participants were instructed to describe aloud what they recalled of the movie in as much detail as they could, with no visual input or experimenter guidance (Fig. 1B), during brain imaging. Participants were allowed to speak for as long as they wished, on whatever aspects of the movie they chose, while their speech was recorded with an MR-compatible microphone.

Without any guidance from the experimenters, participants were able to recall the events of the movie with remarkable accuracy and detail, with the average spoken recall session lasting 21.7 minutes (min: 10.8, max: 43.9, s.d. 8.9) and consisting of 2657 words (min: 1136, max: 5962, s.d. 1323.6; Table S1). Participants' recollections primarily concerned the plot: characters' actions, speech, and motives, and the locations in which these took place. Additionally, many participants described visual features (e.g., colors and viewpoints) and emotional elements (e.g., characters' feelings). The movie was divided into 50 "scenes" (11 – 180 [s.d. 41.6] seconds long), following major shifts in the narrative (e.g., location, topic, and/or time, as defined by an independent rater; see Experimental Procedures). The same "scenes" were identified in the auditory recordings of the recall sessions based on each participant's speech (see Methods). On average, 34.4 (s.d. 6.0) scenes were successfully recalled. A sample participant's complete recall behavior is depicted in Fig. 1C; see Fig. 1D for a summary of all participants' recall behavior. Scenes were recalled largely in the correct temporal order, with an average of 5.9 (s.d. 4.2) scenes recalled out of order. The temporal compression during recall (i.e., the duration of recall relative to the movie; see slopes in Fig.

1D) varied widely, as did the specific words used by different participants (see Table S2 for examples).

Neural reinstatement within participants

Before examining neural patterns shared across people, we first wished to establish to what extent, and where in the brain, the task elicited similar activity between movie-viewing (encoding) and spoken recall within each participant, i.e., movie-recall neural pattern reinstatement. Studies of pattern reinstatement are typically performed within-participant, using relatively simple stimuli such as single words, static pictures, or short video clips, often with many training repetitions to ensure successful and vivid recollection of studied items[19–24]. Thus, it was not known whether pattern reinstatement could be measured after a single exposure to such an extended complex stimulus and unconstrained spoken recall behavior.

For each participant, brain data were transformed to a common space and then data from movie-viewing and spoken recall were each divided into the same 50 scenes as defined for the behavioral analysis. This allowed us to match time periods during the movie to time periods during recall. All timepoints within each scene were averaged, resulting in one pattern of brain activity for each scene. The pattern for each movie scene ("movie pattern") was compared to the pattern during spoken recall of that scene ("recollection pattern"), within-participant, using Pearson correlation (Fig. 2A). The analysis was performed in a searchlight (centered on every voxel in the brain) across the brain volume ($5 \times 5 \times 5$ cubes, 3-mm voxels). Statistical significance was evaluated using a permutation analysis[25] that compares neural pattern similarity between matching scenes against that of non-matching scenes, corrected for multiple comparisons using FDR (q = 0.05, Fig 2A). All statistical tests reported throughout the paper are two-tailed unless otherwise noted. This analysis reveals regions containing scene-specific reinstatement patterns, as statistical significance is only reached if matching scenes (same scene in movie and recall) can be differentiated from non-matching scenes.

The searchlight analysis revealed a large set of brain regions in which the scene-specific spatial patterns observed during movie-viewing were reinstated during the spoken recall session (Fig. 2B), including posterior medial cortex (PMC), medial prefrontal cortex (mPFC), parahippocampal cortex (PHC), and posterior parietal cortex (PPC). This set of regions corresponds well with the DMN[6,26], and encompasses areas that are known to respond during cued recollection in more traditional paradigms[27]. (See Fig. S1A and Table S3 for overlap.) Individual participant correlation values for the PMC are shown in Fig. 2C. PMC was selected for illustration purposes because the region is implicated as having a long (on the order of minutes) memory-dependent integration window in studies that use real-life stimuli such as movies and stories[10,28]. Movie-recall similarity cannot be explained by a time varying signal evolving independently of the stimulus (see *Methods*). These results show that during verbal recall of a 50-minute movie, neural patterns associated with individual scenes were reactivated in the absence of any external cues. For analysis of reinstatement at a finer temporal scale, see Fig. S2.

Pattern similarity between participants

The preceding results established that freely spoken recall of an audio-visual narrative could elicit reinstatement in an array of high-level cortical regions, including those that are typically observed during episodic memory retrieval[27]. Having mapped movie-recall correlations *within* individual participants, we next examined correlations *between* participants during both movie and recall.

Previous studies have shown that viewing the same movie, or listening to the same story, can induce strong between-participant similarity in the timecourses of brain activity in many different regions[7,29]. Now examining spatial (rather than temporal) similarities between participants, we found that scene-specific spatial patterns of activity were highly similar across participants during movie-viewing in areas spanning the cortical hierarchy, from low-level sensory areas to higher-level association areas (Fig. S3). These results also echo prior studies using cross-participant pattern analysis during shared perceptual stimulation in simpler paradigms[30–34].

Next, we compared scene-specific movie patterns and scene-specific recollection patterns between participants. The analysis was identical to the reinstatement analysis described above (Fig. 2B), but performed *between* brains rather than *within* each brain. For each participant, the recollection pattern for each scene was compared to the pattern from the corresponding movie scene, averaged across the remaining participants (Fig. 2D). The searchlight revealed extensive movie-recall correlations between participants, (Fig. 2E–F). These results indicate that in many areas that exhibited movie-recall reinstatement effects within an individual, neural patterns elicited during spoken recollection of a given movie scene were similar to neural patterns in other individuals watching the same scene. (See Table S3 for overlap with DMN.)

Shared spatial patterns between participants during recall

The preceding results showed that scene-specific neural patterns were shared across brains 1) during movie viewing, when all participants viewed the same stimulus (Fig. S3), and 2) when one participant's recollection pattern was compared to other participants' movie-induced brain patterns (Fig. 2D–F). These results suggested that between-participant similarities might also be present during recollection, despite the fact that during the recall session no stimulus was presented and each participant described each movie scene in their own words and for different durations. Thus, we next examined between-brain pattern similarity during the spoken recall session.

As before, brain data within each recall scene were averaged across time in each participant, resulting in one pattern of brain activity for each scene. The recollection pattern from each scene for a given participant was compared (using Pearson correlation) directly to the recollection pattern for the same scene averaged across the remaining participants (Fig. 3A), in every searchlight across the brain. The analysis revealed an array of brain regions that had similar scene-specific patterns of activity between participants during spoken recall of shared experiences (Fig. 3B), including high-order cortical regions throughout the DMN and category-selective high-level visual areas, but not low-level sensory areas (see Fig. S4 for

overlap with visual areas; see Fig. S1 and Table S3 for overlap with DMN). Individual participant correlation values for PMC are shown in Fig. 3C. Pattern similarity between brains could not be explained by acoustic similarities between participants' speech output, and there was no relationship to scene length (see *Methods*).

These between-brain similarities during recall were observed despite there being no stimulus present while participants spoke, and individuals' behavior – the compression factor of recollection and the words chosen by each person to describe each event– varying dramatically (Fig. 1D and Table S2). The direct spatial correspondence of event-specific patterns between individuals suggests the existence of a spatial organization to the neural representations underlying recollection that is common across brains.

Classification accuracy

How discriminable were the neural patterns for individual scenes during the movie? To address this question we performed a multi-voxel classification analysis. Participants were randomly assigned to one of two groups (N=8 and N=9), and an average pattern for each scene was calculated within each group for PMC. Classification accuracy across groups was calculated as the proportion of scenes correctly identified out of 50. Classification rank was calculated for each scene. The entire procedure was repeated using 200 random combinations of groups (Fig. 4A, green markers) and averaged. Overall classification accuracy was 36.7%, $p < 0.001$ (Fig. 4A, black bar; chance level [2.0%] plotted in red); classification rank for individual scenes was significantly above chance for 49 of 50 scenes (Fig. 4B; FDR corrected $q = 0.001$; mean rank across scenes 4.8 out of 50).

How discriminable were the neural patterns for individual scenes during spoken recall? Using the data from recall, we conducted classification analyses identical to the above, with the exception that data were not available for all 50 scenes for every participant (34.4 recalled scenes on average). Thus, group average patterns for each scene were calculated by averaging over extant data; for 41 scenes there were data available for at least one participant in each group, considering the 200 random combinations of participants into groups of N=8 and N=9. Overall classification accuracy was 15.8%, $p < 0.001$ (Fig. 4C; chance level 2.4%). Classification rank for individual scenes was significantly above chance for 40 of 41 possible scenes (Fig. 4D; FDR corrected $q = 0.001$; mean rank across scenes 9.5 out of 41). See Fig. S5 for scene-by-scene pattern similarity values.

To explore factors that may have contributed to discriminability of neural patterns between scenes, we examined 1) how many dimensions of information are encoded in the shared neural patterns that support scene classification, and 2) what kinds of information may be encoded in the shared neural patterns. Our analyses suggest that, at a minimum, 15 dimensions of information are encoded in the patterns shared across individuals during the movie, and approximately 12 generalized between movie and recall in PMC (Fig. 5). We found that the presence of characters' speech, of written text, the number of locations visited and of persons onscreen, arousal, and valence each contributed to PMC activity patterns during movie viewing (see *Methods: Encoding model* and Fig. S6).

Visualization of BOLD activity in individual scenes

In order to visualize the neural signal shared across subjects, we randomly split the movie-viewing data into two equally sized independent groups (N=8 each) and averaged BOLD values across participants within each group, and across timepoints within-scene; the same was done for the recall data, creating one brain image per group per scene. For movie data see Fig. 6A (leftmost 2 panels) and 6B–C; for recall see Fig. 6A (rightmost 2 panels) and 6D–E. This averaging procedure reveals the component of the BOLD signal that is shared across brains, i.e., if a similar activity pattern can be observed between the two independent groups for an individual scene, it indicates a common neural response across groups. Visual inspection of these images suggests replication across groups for individual scenes, and differentiation between scenes, as quantified in the classification analysis above (see Fig. 4; also Fig. S5 for scene-by-scene correlation values). While the current results reveal a relatively coarse spatial structure that is shared across people, they do not preclude the existence of finer spatial structure in the neural signal that may be captured when comparisons are made within-participant or using more sensitive methods such as hyperalignment. Further work is needed to understand the factors that influence the balance of idiosyncratic and shared signals between brains. See 'Spatial resolution of neural signals' in Online Methods.

Alteration of neural patterns from perception to recollection

A key question of the experiment was how neural representations change between perception (movie viewing) and memory (recollection). If a given scene's recollection pattern is simply a noisy version of the movie scene pattern, recall-recall correlation between brains cannot be higher than the movie-recall correlation. This is illustrated in Fig. 7A (left) by adding uncorrelated noise patterns to the simulated movie patterns for each brain; in this scenario, recollection patterns for a given scene necessarily become *more dissimilar* to the recollection patterns of other people than they are to the movie pattern. In contrast, if the movie patterns are altered in systematic manner, it is possible for the recall-recall correlations to be higher than movie-recall correlations; in Fig. 7A, right, adding correlated alteration patterns to the movie patterns within each brain can result in recollection patterns becoming *more similar* to the recollection patterns of other people than to the original movie scene pattern (Fig. 7A, right). Note that the movie-movie correlation values are irrelevant for this analysis. See also Fig. S7 and *Methods: Simulation of movie-to-recall pattern alteration*.

To search for systematic alterations of neural representations between movie and recall, we looked for brain regions in which, for individual scenes, recollection activity patterns were more similar to recollection patterns in other individuals than they were to movie patterns. In order to ensure a balanced contrast, we compared the between-participants recall-recall values to the *between*-participants movie-recall values (rather than to *within*-participant movie-recall). Statistical significance of the difference was calculated using a resampling test (see Methods). The analysis revealed that recall-recall pattern similarity was stronger than movie-recall pattern similarity in PHC and other high-level visual areas, right superior temporal pole, PMC, right mPFC, and PPC (Fig. 7B). (See Table S3 for overlap with DMN.) If recall patterns were merely a noisy version of the movie events, we would expect the exact opposite result: lower recall-recall similarity than movie-recall similarity. No regions in the

brain showed this opposite pattern. Thus, the analysis revealed alterations of neural representations between movie-viewing and recall that were systematic (shared) across subjects.

How did global differences between movie and recall impact alteration?

There were a number of global factors that differed between movie and recall: e.g., more visual motion was present during movie than recall; recall involved speech and motor output while movie-viewing did not. A possible concern was that the greater similarity for recall-recall relative to movie-recall might arise simply from such global differences. To test this concern, we examined the *discriminability* of individual scenes within the regions that exhibited robust pattern alteration, i.e., the searchlights shown in the Fig. 7B map. We asked whether each participant's individual scene recollection patterns could be classified better using a) movie data from other participants, or b) the recall data from other participants. Classification rank was found to be higher when using the recall data as opposed to the movie data in 99% of the searchlights; see Fig. 7C for the distribution of t-values resulting from a test between recall-recall and movie-recall. Thus, global differences between movie and recall could not explain our observation of systematic pattern alterations. (See Fig. S8 for the scene-by-scene difference in pattern similarity between recall-recall and movie-recall).

Interestingly, in PMC, the degree of alteration predicted the memorability of individual scenes (Fig. S9). See Fig. S10 for analysis of subsequent memory effects in hippocampus; Fig. S11 for analysis of hippocampal sensitivity to the gap between part 1 and part 2 of the movie.

Reinstatement within vs. between participants

While our between-participant analyses explored the shared component of memory representations, neural patterns may also contain information reflecting more fine-grained individual differences in memory representations[35]. If so, one would expect movie-recall similarity to be stronger within-participant than between-participant. A simple comparison of within-participant movie-recall pattern similarity (Fig. 2B–C) to between-participant movie-recall pattern similarity (Fig. 2E–F) does not suffice, due to anatomical registration being better within-participant than between-participants. Thus, we performed a second-order similarity analysis: correlation of representational dissimilarity matrices (RDMs) within and between participants[25,35]. Each RDM was composed of the pairwise pattern correlations for individual scenes in the movie ("movie-RDM") and during recall ("recall-RDM") calculated within-brain. These RDMs could then be compared within and between participants.

We calculated correlations between movie-RDM and recall-RDM, within-participant, in a searchlight analysis across the brain volume (Fig. 8A). The same analysis was performed between all pairs of participants (Fig. 8B). Of critical interest was the difference between the within-participant comparison and the between-participant comparison. Statistical significance of the difference was evaluated using a permutation analysis that randomly swapped condition labels for within- and between-participant RDM correlation values, and

FDR corrected across all voxels in the brain (q = 0.05). This analysis revealed a single cluster located in the right temporoparietal junction (two searchlight cubes centered on MNI coordinates [48, −48, 9] and [48, −48, 6]; Fig. 8C) for which within-participant movie-RDM to recall-RDM correlation was significantly greater than between-participants movie-RDM to recall-RDM correlation, i.e., in which individual-unique aspects of neural patterns contributed to reinstatement strength above and beyond the shared representation. (See also *Methods: Comparison between semantic similarity and neural pattern similarity*.)

DISCUSSION

In this study, we found that neural patterns recorded during movie-viewing were reactivated, in a scene specific manner, during free and unguided verbal recollection. Furthermore, the spatial organization of the recall neural patterns was preserved across people. This shared brain activity was observed during free spoken recall as participants reported the contents of their memories (a movie they had watched) in their own words, in the absence of any sensory cues or experimental intervention. Reactivated and shared patterns were found in a large set of high-order multimodal cortical areas, including default mode network areas, high-level visual areas in ventral temporal cortex, and intraparietal sulcus. In a subset of regions, brain activity patterns were modified between perception and recall in a consistent manner across individuals. Interestingly, the magnitude of this modification predicted the memorability of individual movie scenes, suggesting that alteration of brain patterns between percept and recollection may have been beneficial for behavior (Fig. S9). Overall, these findings show that the neural activity underlying memories for real-world events has a common spatial organization across different brains, and that neural activity is altered from initial perception to recollection in a systematic manner, even as people speak freely in their own words about past events.

Our findings suggest that memory representations for real-world events, similar to sensory representations, are spatially organized in a functional architecture that is shared across brains. The well-studied spatial organization of sensory responses takes the form of "topographic maps" in the brain, e.g., retinotopic or tonotopic maps of visual or auditory features, which are preserved across individuals[36–38]. In contrast, little is known about the consistency of memory-related cortical patterns across individuals. Memory-relevant areas such as hippocampus and entorhinal cortex represent an animal's spatial location and trajectory[39], but these representations are remapped for new spatial layouts and do not seem to be consistent across brains[40]. Previous studies have compared activity between brains using the powerful representational similarity analysis (RSA) approach[25], in which the overall structure of inter-relationships between stimulus-evoked neural responses is compared between participants[35,41]. While this type of RSA is a second-order comparison (a correlation of correlations), our approach uses direct comparison of spatial activity patterns between brains; this establishes that the spatial structures of neural patterns underlying recollected events are common across people (as opposed to the stimulus inter-relationship structure). The shared responses are local in that similarity can be detected in a small patch of brain (a 15×15×15 mm searchlight; see *Methods: Spatial resolution of neural signals* and Fig. 6), but also widespread, encompassing default mode network areas, high-level visual areas in ventral temporal cortex (but not low-level visual areas; Fig. S4 and *Methods: Visual*

imagery), and intraparietal sulcus. Future work will explore the mapping between specific mnemonic content and the structure of neural responses in these areas (initial analyses, Fig. S6).

The brain areas in which we observed shared representations during recall overlap extensively with the "default mode network" (DMN)[6]. The DMN has been implicated in a broad range of complex cognitive functions, including scene and situation model construction, episodic memory, and internally focused thought[26,27,42,43]. Multiple studies have shown that, during processing of real-life stimuli such as movies and stories, DMN activity timecourses are synchronized across individuals and locked to high-level semantic information in the stimulus, but not to low-level sensory features or mid-level linguistic structure. For example, these regions evince the same narrative-specific dynamics whether a given story is presented in spoken or written form[14,15], and whether it is presented in English or Russian[11]. Dynamics are modulated according to the perspective of the perceiver[12], but when comprehension of the narrative is disrupted (while keeping low-level sensory features unchanged), neural activity becomes incoherent across participants[10,44]. Together, these results suggest that DMN activity tracks high-level information structure (e.g., narrative or situational elements[43]) in the input. The current study extends prior findings by demonstrating that synchronized neural responses among individuals during encoding later give rise to shared neural responses during recollection, reinstated at will from memory without the need for any guiding stimulus, even when each person describes the past in his or her own words. Note that, while there was considerable overlap between the DMN and the recall-recall map in posterior medial areas, there is less overlap in frontal and lateral parietal areas (Fig. S1 and Table S3). Our results show that some DMN regions evince shared event-specific activity patterns during recollection; more work is needed to probe functional differentiation within these areas.

A memory is not a perfect replica of the original experience; perceptual representations undergo modification in the brain prior to recollection that may increase the usefulness of the memory, e.g., by emphasizing certain aspects of the percept and discarding others. What laws govern how neural representations change between perception and memory, and how might these modifications be beneficial for future behavior? We examined whether the alteration of neural patterns from percept to memory was idiosyncratic or systematic across people, reasoning that if percept-based activity changed into memory in a structured way, then patterns at recall should become more similar to each other (across individuals) than they are to the original movie patterns. Such systematic transformations were observed in several brain regions, including PHC and other high-level visual areas, superior temporal pole, PMC, mPFC, and PPC (Fig. 6B). Not only the similarity, but also the discriminability of events was increased during recall, indicating that the effect was not due to a common factor (e.g., speech) across recalled events (Fig. 6C). Interestingly, scenes that exhibited more neural alteration in PMC were also more likely to be recalled (Fig. S9A; this should be interpreted cautiously, as the number of data points, 45, is relatively small). A possible interpretation of these findings is that participants shared familiar notions of how certain events are structured (e.g., what elements are typically present in a car chase scene), and these existing schemas guided recall. Such forms of shared knowledge might improve

memory by allowing participants to think of schema-consistent scene elements, essentially providing self-generated memory cues[45].

In the current study, we simplified the continuous movie and recall data by dividing them into scenes, identified by major shifts in the narrative (e.g., location, topic, time). This "event boundary" segmentation was necessary for matching time periods in the movie to periods during recall. Our approach follows from a known property of perception: people tend to segment continuous experience into discrete events in similar ways[46]. While there are many reasonable ways to split the movie, matching movie scenes to recall audio becomes difficult when the number of boundaries increases, as some descriptions are more synoptic (Table S2). To overcome the temporal misalignment between movie and recall, and between recalls, each event was averaged across time. These averaged event patterns nonetheless retained complex information: movie scene patterns contained at least 15 dimensions that contributed to classification accuracy (Fig. 5A), and approximately 12 dimensions generalized from movie to recall (Fig. 5B). Furthermore, in an exploration of what semantic content might underlie these dimensions (Fig. S6), we found that the presence of speech, of written text, number of locations visited and persons onscreen, arousal, and valence each contributed to PMC movie activity patterns. Thus, while some information was necessarily lost when we averaged within-scene, substantial multidimensional structure was preserved. (For consideration of the role of visual imagery, see *Methods: Visual imagery*; Fig. S4.) Recollection is obviously more complicated than a simple compression of original perception; a future direction is to formalize the hypothesis that narrative recall is "chunked" into scenes and use this heuristic to enable data-driven discovery of optimal scene boundaries during both movie and recall, without relying on human manual definition[47].

Together, these results show that a common spatial organization for memory representations exists in high-level cortical areas (e.g., the DMN), where information is largely abstracted beyond sensory constraints; and that perceptual experience is altered before recall in a systematic manner across people, a process that may benefit memory. These observations were made as individuals engaged in natural and unguided spoken recollection, testifying to the robustness and ecological validity of the phenomena. The ability to use language to reactivate, at will, the sequence of neural responses associated with the movie events, can be thought of as a form of conscious replay. Future studies may investigate if and how such volitional cortical replay is related to the compressed and rapid forward and reverse sequential replay observed in the hippocampus during sleep and spatial navigation[48]. Future work may also explore whether these shared representations facilitate the spoken communication of memories to others[49,50], and how they might contribute to a community's collective memory[2–5].

Online Methods

Participants

Twenty-two participants were recruited from the Princeton community (12 male, 10 female, ages 18–26, mean age = 20.8). All participants were right-handed native English speakers, reported normal or corrected-to-normal vision, and had not watched any episodes of *Sherlock*[1] prior to the experiment. All participants provided informed written consent prior

to the start of the study in accordance with experimental procedures approved by the Princeton University Institutional Review Board. The study was approximately two hours long and participants received $20 per hour as compensation for their time. Data from five out of the twenty-two participants were discarded due to excessive head motion (greater than one voxel; 2 participants), or because recall was shorter than 10 minutes (2 participants), or for falling asleep during the movie (1 participant). For one participant (#5 in Figs. 2C, 2F, 3C) the movie scan ended 75 seconds early (i.e., this participant was missing data for part of scene 49 and all of scene 50). No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications[9,11,14].

Stimuli

The audio-visual movie stimulus was a 48-minute segment of the BBC television series *Sherlock*[51], taken from the beginning of the first episode of the series (a full episode is 90 minutes). The stimulus was further divided into two segments (23 and 25 minutes long); this was done to reduce the length of each individual runs, as longer runs might be more prone to technical problems (e.g., scanner overheating).

At the beginning of each of the two movie segments, a 30-second audiovisual cartoon was prepended (*Let's All Go to the Lobby*[52]) that was unrelated to the *Sherlock* movie. In studies using inter-subject temporal correlation, it is common to include a short auditory or audiovisual "introductory" clip before the main experimental stimulus, because the onset of stimulus may elicit a global arousal response. Such a response could add noise to a temporal correlation across subjects, and thus experimenters often truncate the neural signal elicited during the introductory clip. In the current experiment, we used spatial instead of temporal correlations. Because spatial correlation in one scene does not necessarily affect the correlation in another scene, we decided not to remove the introductory clips. As can be seen in Fig. 4B, the introductory clips at the beginning of each movie run (bars C1 and C2, striped) are highly discriminable from the 48 scenes of the movie. Furthermore, one participant described the introductory clip during their spoken recall, and this event was included for the within-brain movie-recall analysis for that participant. In the absence of any obvious reason to exclude these data, we decided to retain the cartoon segments.

Experimental Procedures

Participants were told that they would be watching the British television crime drama series *Sherlock*[51] in the fMRI scanner. They were given minimal instructions: to attend to the audiovisual movie, e.g., "watch it as you would normally watch a television show that you are interested in," and that afterward they would be asked to verbally describe what they had watched. Participants then viewed the 50-minute movie in the scanner. The scanning (and stimulus) was divided into two consecutive runs of approximately equal duration.

Data collection and analysis were not performed blind to the conditions of the experiments. All participants watched the same movie before verbally recalling the plot of the movie. To preserve the naturalistic element of the experimental design, we presented the movie in its original order. The unit of analysis was the scenes that participants freely recalled, which was not within experimental control.

The movie was projected using an LCD projector onto a rear-projection screen located in the magnet bore and viewed with an angled mirror. The Psychophysics Toolbox [http://psychtoolbox.org] for MATLAB was used to display the movie and to synchronize stimulus onset with MRI data acquisition. Audio was delivered via in-ear headphones. Eyetracking was conducted using the iView X MRI-LR system (Sensomotoric Instruments [SMI]). No behavioral responses were required from the participants during scanning, but the experimenter monitored participants' alertness via the eyetracking camera. Any participants who appeared to fall asleep, as assessed by video monitoring, were excluded from further analyses.

At the start of the spoken recall session, which took place immediately after the end of the movie, participants were instructed to describe what they recalled of the movie in as much detail as they could, to try to recount events in the original order they were viewed in, and to speak for at least 10 minutes if possible but that longer was better. They were told that completeness and detail were more important than temporal order, and that if at any point they realized they had missed something, to return to it. Participants were then allowed to speak for as long as they wished, and verbally indicated when they were finished (e.g., "I'm done"). During this session they were presented with a static black screen with a central white dot (but were not asked to, and did not, fixate); there was no interaction between the participant and the experimenter until the scan ended. Functional brain images and audio were recorded during the session. Participants' speech was recorded using a customized MR-compatible recording system (FOMRI II; Optoacoustics Ltd.).

Behavioral Analysis

Scene timestamps—Timestamps were identified that separated the audiovisual movie into 48 "scenes", following major shifts in the narrative (e.g., location, topic, and/or time). These timestamps were selected by an independent coder with no knowledge of the experimental design or results. The scenes ranged from 11 to 180 [s.d. 41.6] seconds long. Each scene was given a descriptive label (e.g., *Press conference*). Together with the two identical cartoon segments, this resulted in 50 total scenes.

Transcripts—Transcripts were written of the audio recording of each participant's spoken recall. Timestamps were then identified that separated each audio recording into the same 50 scenes that had been previously selected for the audiovisual stimulus. A scene was counted as "recalled" if the participant described any part of the scene. Scenes were counted as "out of order" if they were initially skipped and then described later. See Tables S1 and S2.

fMRI Acquisition

MRI data were collected on a 3T full-body scanner (Siemens Skyra) with a 20-channel head coil. Functional images were acquired using a T2*-weighted echo planar imaging (EPI) pulse sequence (TR 1500 ms, TE 28 ms, flip angle 64, whole-brain coverage 27 slices of 4 mm thickness, in-plane resolution 3×3 mm^2, FOV 192×192 mm^2), with ascending interleaved acquisition. Anatomical images were acquired using a T1-weighted MPRAGE pulse sequence (0.89 mm^3 resolution).

fMRI Analysis

Preprocessing—Preprocessing was performed in FSL [http://fsl.fmrib.ox.ac.uk/fsl], including slice time correction, motion correction, linear detrending, high-pass filtering (140 s cutoff), and coregistration and affine transformation of the functional volumes to a template brain (Montreal Neurological Institute [MNI] standard). Functional images were resampled to 3 mm isotropic voxels for all analyses. All calculations were performed in volume space. Projections onto a cortical surface for visualization were performed, as a final step, with NeuroElf (http://neuroelf.net).

Motion was minimized by instructing participants to remain very still while speaking, and stabilizing participants' heads with foam padding. Artifacts generated by speech may introduce some noise, but cannot induce positive results, as our analyses depend on spatial correlations between sessions (movie-recall or recall-recall). Similar procedures regarding speech production during fMRI are described in previous publications from our group[53,54].

Region of Interest (ROI) definition—An anatomical hippocampus ROI was defined based on the probabilistic Harvard-Oxford Subcortical Structural atlas[55], and an ROI for posterior medial cortex (PMC) was taken from an atlas defined from resting-state connectivity[56]: specifically, the posterior medial cluster in the "dorsal default mode network" set (http://findlab.stanford.edu/functional_ROIs.html), see Fig. S3D. A default mode network ROI was created by calculating the correlation between the PMC ROI and every other voxel in the brain (i.e., "functional connectivity") during the movie for each subject, averaging the resulting maps across all subjects, and thresholding at R = 0.4 (Fig. S1). 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[57].

Pattern similarity analyses—The brain data were transformed to standard MNI space. For each participant, data from movie-viewing and spoken recall were each divided into the same 50 scenes as defined for the behavioral analysis. BOLD data were averaged across timepoints within-scene, resulting in one pattern of brain activity for each scene: one "movie pattern" elicited during each movie scene, and one "recollection pattern" elicited during spoken recall of each scene. Recollection patterns were only available for scenes that were successfully recalled, i.e., each participant possessed a different subset of recalled scenes. Each scene-level pattern could then be compared to any other scene-level pattern in any region (e.g., an ROI or a searchlight cube). For cases in which a given scene was described more than once during recall, data were used from the first description only. All such comparisons were made using Pearson correlation.

For searchlight analyses[58], pattern similarity was calculated in $5 \times 5 \times 5$ voxel cubes (i.e., $15 \times 15 \times 15$ mm cubes) centered on every voxel in the brain. Statistical significance was determined by shuffling scene labels to generate a null distribution of the average across participants, i.e., baseline correlations were calculated from all (matching and non-matching) scene pairs[25]. This procedure was performed for each searchlight cube, with one cube centered on each voxel in the brain; cubes with 50% or more of their volume outside

the brain were discarded. The results were corrected for multiple comparisons across the entire brain using FDR[59] (threshold q = 0.05). Importantly, the nature of this analysis ensures that the discovered patterns are content-specific at the scene level, as the correlation between neural patterns during matching scenes must on average exceed an equal-sized random draw of correlations between all (matching and non-matching) scenes in order to be considered statistically significant.

Four types of pattern similarity analyses were conducted: (1) Movie-recall within participant: The movie pattern was compared to the recollection pattern for each scene within each participant (Fig. 2A). (2) Movie-recall between participants: For each participant, the recollection pattern of each scene was compared to the movie pattern for that scene averaged across the remaining participants (Fig. 2D). (3) Recall-recall between-participant: For each participant, the recollection pattern of each scene was compared to the recollection patterns for that scene averaged across the remaining participants (Fig. 3A). The preceding analyses each resulted in a single brain map per participant. The average map was submitted to the shuffling-based statistical analysis described above and p-values were plotted on the brain for every voxel and thresholded using FDR correction over all brain voxels (Figs. 2B, 2E, and 3B). The same pattern comparison was performed in the PMC ROI and the results plotted for each individual participant (Figs. 2C, 2F, and 3C). For the fourth type of analysis, movie-movie between-participant, for each participant the movie pattern of each scene was compared to the movie pattern for that scene averaged across the remaining participants. The average R-value across participants was plotted on the brain for every voxel (Fig. S3B) and the values for the PMC ROI are shown for individual participants in Fig. S3C.

Classification of individual scenes—We computed the discriminability of neural patterns for individual scenes during movie and recall (Fig. 4) in the PMC ROI (same ROI as Figs. 2C, 2F, 3C). Participants were randomly assigned to one of two groups (N=8 and N=9), an average was calculated within each group, and data were extracted for the PMC ROI. Pairwise correlations were calculated between the two group means for all 50 movie scenes. For any given scene (e.g., scene 1, group 1), the classification was labeled "correct" if the correlation with the matching scene in the other group (e.g., scene 1, group 2) was higher than the correlation with any other scene (e.g., scenes 2–50, group 2). Accuracy was then calculated as the proportion of scenes correctly identified out of 50 (chance level = 0.02). Classification rank was calculated for each scene as the rank of the matching scene correlation in the other group among all 50 scene correlations. The entire procedure was repeated using 200 random combinations of two groups sized N=8 and N=9. Statistical significance was assessed using a permutation analysis in which, for each combination of two groups, scene labels were randomized before computing the classification accuracy and rank. Accuracy was then averaged across the 200 combinations, for each scene the mean rank across the 200 combinations was calculated, and this procedure was performed 1000 times to generate null distributions for overall accuracy and for rank of each scene. Classification rank p-values were corrected for multiple comparisons over all scenes using FDR at threshold q = 0.001.

All above analyses were identical for movie and recall except that data were not extant for all 50 scenes for every participant, due to participants recalling 34.4 scenes on average. Thus, group average patterns for each scene were calculated by averaging over the extant data; for 41 scenes there were data available for at least one participant in each group, considering all 200 random combinations of participants into groups of N=8 and N=9.

Dimensionality of the shared neural patterns: The SRM[60] algorithm operates over a series of data vectors (in this case multiple participants' brain data) and finds a common representational space of lower dimensionality. Using SRM, we asked: when the data are reduced to k dimensions, how does this affect scene-level classification across brains? We tested this in the PMC region. The movie data were split randomly into two sets of scenes (25 and 25). SRM was trained on one set of 25 scenes (data not averaged within-scene). For the remaining 25 scenes, the data were split randomly into two groups (8 and 9 participants), and scene-level classification accuracy was calculated for each scene, using one group's low-dimensional pattern in SRM representational space to identify the matching scene in the other group. This entire process was repeated for 10 random splits of scenes, 40 random splits of participants, and for different numbers of dimensions of k, ranging from 3 dimensions to 75 dimensions. The average of these 400 results is plotted in Fig. 5A (chance level 0.04). To examine how many of the movie dimensions were generalizable to recall, we performed the same analysis, training SRM on the movie data (25 scenes at a time, data not averaged within-scene), and performing scene classification across participants during spoken recall (recall-recall) using the identified movie dimensions. See Fig. 5B. Unfortunately, because participants are not temporally aligned during recall, we could not use the SRM to identify additional shared dimensions that might be unique to the recall data.

Visualization of the BOLD signal during movie and recall—In order to visualize the signals underlying our pattern similarity analyses, we randomly split the movie-viewing data into two independent groups of equal size (N=8 each) and averaged BOLD values across participants within each group ("Movie Group 1" and "Movie Group 2"). An average was made in the same manner for the recall data from the same groups of eight participants each ("Recall Group 1" and "Recall Group 2"). These group mean images were then averaged across timepoints and within scene, exactly as in the prior analyses, creating one brain image per group per scene. A midline sagittal view of these average brains during one representative scene (scene 36) of the movie is shown in Fig. 6A. For the posterior medial area outlined by a white box in each panel of Fig. 6A, we show the average activity for fourteen different scenes (scenes that were recalled by all sixteen of the randomly selected subjects) for each group in Fig. 6B–E.

Alteration analysis: Comparison of movie-recall and recall-recall maps—In this analysis (Fig. 7) we quantitatively compared the similarity strength of recall-recall to the similarity strength of movie-recall. We compared recall-recall (Fig. 3B) correlation values to between-participants movie-vs-recall (Fig. 2E) correlation values using a voxel-by-voxel paired t-test. It was necessary to use between-participants movie-recall comparisons because within-participant pattern similarity was expected to be higher than between-participant similarity merely due to lower anatomical variability. To assess whether the differences were

statistically significant, we performed a resampling analysis wherein the individual participant correlation values for recall-recall and movie-recall were randomly swapped between conditions to produce two surrogate groups of 17 members each, i.e., each surrogate group contained one value from each of the 17 original participants, but the values were randomly selected to be from the recall-recall comparison or from the between-participant movie-recall comparison. These two surrogate groups were compared using a t-test, and the procedure was repeated 100,000 times to produce a null distribution of t values. The veridical t-value was compared to the null distribution to produce a p-value for every voxel. The test was performed for every voxel that showed *either* significant recall-recall similarity (Fig. 3B) *or* significant between-participant movie-recall similarity (Fig. 3B), corrected for multiple comparisons across the entire brain using an FDR threshold of q = 0.05 (see *Methods: Pattern similarity analyses*); voxels p < 0.05 (one-tailed) are plotted on the brain (Fig. 4A). This map shows regions where between-participants recall-recall similarity was significantly greater than between-participants movie-recall similarity, i.e., where the map in Fig. 3B was stronger than the map in Fig. 2E. See also Fig. S8A for scene-by-scene differences and Fig. S8B for map of correlation difference values.

Discriminability of individual scenes was further assessed within the regions shown in Fig. 7B. For every searchlight cube underlying the voxels shown in Fig. 7B, we asked whether each participant's individual scene recollection patterns could be classified better using a) the movie data from other participants, or b) the recall data from other participants. Unlike the classification analysis described in Fig. 4, calculations were performed at the individual participant level (e.g., using each participant's recollection patterns compared to the average patterns across the remaining participants, for either movie or recall). Mean classification rank across scenes was calculated to produce one value per participant for a) the movie data from other participants, and one for b) the recall data from other participants. A t-test between these two sets of values was performed at each searchlight cube (Fig. 7C).

Simulation of movie-to-recall pattern alteration—The logic of the "alteration" analysis is that *only* if the movie patterns change into recall patterns in a systematic manner across subjects is it possible for the recall-recall pattern correlations to be higher than movie-recall pattern correlations. We demonstrate the logic using a simple computer simulation. 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). 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) in Fig. S7.

Fig. S7A shows the 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). In short, if the change from movie to recall was simply adding noise, then the recall patterns could not possibly become more similar to each other then they are to the original movie pattern.

Fig. S7B shows what happens 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 change 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. In this simulation, the movie-movie correlations are presented at two levels, such that the average inter-subject correlation (on the y-axis) is R=1.0 (red lines) or R=0.3 (blue) before movie-recall noise is added (at 0 on the x-axis). In practice, movie-movie pattern correlations were higher overall than recall-recall and movie-recall correlations; compare Fig. S3C to Figs. 2F and 3C.

Reinstatement in individual participants vs. between participants—We performed a second-order similarity analysis: correlation of representational dissimilarity matrices (RDMs)[25,35] within and between participants. An RDM was created from the pairwise correlations of patterns for individual scenes in the movie ("movie-RDM") and a separate RDM created from the pairwise correlations of patterns for individual scenes during recall ("recall-RDM"), for each participant. The movie-RDMs map the relationships between all movie scenes, and the recall-RDMs map the relationships between all recalled scenes. Because the RDMs were always calculated within-brain, we were able to assess the similarity between representational structures within and between participants (by comparing the movie-RDMs to the recall-RDMs within-participant and between-participants [see Fig. 8A–B insets]) in a manner less susceptible to potential anatomical misalignment between participants.

As each participant recalled a different subset of the 50 scenes, comparisons between movie-RDMs and recall-RDMs were always restricted to the extant scenes in the recall data for that participant. Thus, two different comparisons were made for each pair of participants, e.g., S1 movie-RDM vs. S2 recall-RDM, and S2 movie-RDM vs. S1 recall-RDM. In total this procedure yielded 17 within-participant comparisons and 272 between-participant comparisons. We calculated movie-RDM vs. recall-RDM correlations, within-participant, in a searchlight analysis across the brain volume (Fig. 8A). The same analysis was performed between all pairs of participants (Fig. 8B).

Due to the differing amounts of averaging in the within-participant and between-participant maps (i.e., averaging over 17 vs. 272 individual maps respectively), we did not perform significance testing on these maps separately, but instead tested the difference between the

maps in a balanced manner. Statistical significance of the difference between the two was evaluated using a permutation analysis that randomly swapped condition labels for within- and between-participant RDM correlation values, and FDR corrected at a threshold of q = 0.05. Thus, averaging was performed over exactly 17 maps for each permutation.

Control analysis for elapsed time during movie and recall—In order to examine whether the brain regions revealed in the movie-recall pattern similarity analysis (Fig. 2) have a time-varying signal that shifts independent of the stimuli, we compared segments across subjects during recall at the same time elapsed from the start of recall. We extracted a 20-TR (30-second) window from each subject's neural data (PMC ROI) at each minute of the movie and recall, from 1 to 10 minutes. We then averaged across time to create a single voxel pattern for each segment, and calculated the correlation between all movie segments vs. all recall segments within each subject, exactly as in the main analyses of the paper (e.g., Figs. 2A–C). The mean correlation of matched-for-time-elapsed segments was R = 0.0025, and the mean of all comparisons not matched for time elapsed is R = −0.0023. A paired t-test of matched vs. non-matched segment correlation values, across subjects, yields a p-value of 0.22. Performing the same analysis using 1-minute windows (40 TRs) yields a mean diagonal of R = −0.0025, a mean non-diagonal of R = −0.0015, and a p-value of 0.84. This analysis suggests that PMC does not have a time-varying signal that shifts independent of the stimuli.

Was acoustic output correlated between subjects?—We extracted the envelope of each subject's recall audio and linearly interpolated within each scene so that the lengths would be matched for all subjects (100 timepoints per scene). We then calculated the correlation between all audio segments across subjects (each subject vs. the average of all others), exactly as in the main analyses of the paper (e.g., Figure 3). The mean correlation of matching scene audio timecourses was R = 0.0047, and the mean of all non-matching scene audio timecourses R = 0.0062. A paired t-test of matching vs. non-matching values, across subjects, yields a p-value of 0.98. This analysis shows that speech output during recall was not correlated across subjects.

Length of scenes vs. neural pattern similarity—To explore whether the length of movie scenes and/or length of recall of a scene affected inter-subject neural similarity, we calculated the correlation of scene length vs. inter-subject pattern similarity in posterior medial cortex. During the movie, all subjects have the same length scenes, and thus we averaged across subjects to get the strongest possible signal. However, the correlation between movie scene length and inter-subject similarity of movie patterns in PMC was not significant: R = 0.17, p > 0.2. During recall, subjects all have different length scenes, and thus we performed the correlation separately for each subject. The correlation between recall scene length and inter-subject similarity of recall patterns was R = 0.006 on average, with no p values less than 0.05 for any subject.

Encoding model—Detailed semantic labels (1000 time segments for each of 10 labels) were used to construct an encoding model to predict neural activity patterns from semantic content[61,62]. A score was derived for each of the 50 scenes for each of the 10 labels (e.g.,

proportion of time during a scene that Speaking was true; proportion of time that was Indoor; average Arousal; number of Locations visited within a scene; etc.). The scene-level labels/predictors are displayed in Fig. S6A.

Encoding Model: Semantic labels: The semantic and affective features of the stimulus were labeled by dividing the stimulus into 1000 time segments and scoring each of 10 features in each segment. First, the movie was split into 1000 time segments by a human rater (mean duration 3.0 seconds, s.d. 2.2 seconds) who was blind to the neural analyses. The splits were placed at shifts in the narrative (e.g., location, topic, and/or time), in a procedure similar to, but much more fine-grained, than for the original 50 scenes. Each of the 1000 segments was then labeled for the following content:

NumberPersons: How many people are present onscreen

Location: What specific location is being shown (e.g., "Phone box on Brixton Road")

Indoor/Outdoor: Whether the location is indoor or outdoor

Speaking: Whether or not anyone is speaking

Arousal: Excitement/engagement/activity level

Valence: Positive or negative mood

Music: Whether or not there is music playing

WrittenWords: Whether or not there are written words onscreen

Sherlock: Whether or not the character Sherlock Holmes is onscreen

John: Whether or not the character John Watson is onscreen

For the labels Arousal and Valence, assessments were collected from four different raters (Arousal: Cronbach's alpha = 0.75; Valence: Cronbach's alpha = 0.81) and the average across raters used for model prediction. The other eight labels were deemed objective and not requiring multiple raters.

Encoding Model: Feature Selection and Fitting: To create the encoding model, we split the data randomly into two groups of participants (8 and 9) and created an average pattern for each scene in each group. Next, for one label (e.g., Arousal), we regressed each voxel's activity values from 48 scenes on the label values for those 48 scenes, constituting a prediction of the relationship between label value (e.g., Arousal level for a scene) and voxel activity. By calculating this fit separately for every voxel in PMC, we created a predicted pattern for each of the two held-out scenes. These predicted patterns were then compared to the true patterns for those scenes in Group 1, iterating across all possible pairs of held-out scenes, enabling calculation of classification accuracy (with chance level 50%). The label with the highest accuracy (out of the 10 labels) was ranked #1. Next, the same procedure was repeated using the rank #1 label and each of the remaining 9 labels in a multiple

regression (i.e., we generated every possible 2-predictor model including the rank #1 label; for each model, we predicted the held-out patterns and then computed classification accuracy). The label that yielded the highest accuracy in conjunction with the rank #1 label was ranked #2. This hierarchical label selection was repeated to rank all 10 labels.

Encoding Model: Testing in held out data: Using the ranking order determined from the Group 1 data, we then calculated classification accuracy using the heretofore untouched Group 2 data: as above, we created a predicted pattern for two held-out scenes, and these predicted patterns were compared to the true patterns for those scenes in Group 2, iterating across all possible pairs of held-out scenes, enabling calculation of classification accuracy (with chance level 50%). We first did this for a model incorporating the rank #1 label, then a model incorporating the rank #1 and #2 labels, and so on. The entire procedure was performed for 100 random unique splits of participants into two groups (N=8 and N=9). The classification accuracy and label rankings across the 100 combinations are plotted in Fig. S6B and Fig. S6C respectively. The predictor confusion matrix is displayed in Fig. S6D.

Comparison between semantic similarity and neural pattern similarity—To examine if shared spatial patterns during a recalled scene reflect shared content in verbal recalls, we tested if neural pattern similarity correlates with semantic similarity in the content of participants' verbal recall. We performed latent semantic similarity analysis using the LSA package in R (Wild, 2005). Transcripts from all participants were used as the corpus, with each recalled scene constituting a *document*. The corpus was preprocessed to convert all characters to lower-case, and to remove punctuation and stop words. Words were then reduced to their word-stems, which then constitute the *terms* that enter into the computation of the latent semantic space. The corpus was then converted to a term×document matrix, and singular-value decomposition was applied to compute the latent semantic space.

Using the latent semantic space, we computed the semantic similarity between recalled scenes. For each pair of participants, we computed the semantic similarity between mutually recalled scenes, and correlated this vector of semantic similarities with the corresponding vector of neural pattern similarities in PMC. The correlation was then averaged across every pair of participants. To assess statistical significance, we ran a non-parametric permutation test. For each pair of participants, we shuffled the neural pattern similarity vector prior to correlating it with the semantic similarity vector. This was iterated 1000 times to generate a null distribution of average r-values. The true r-value was compared to the null-distribution to determine the p-value.

The correlation was weak but significant (R=0.08, p < 0.001), suggesting that the shared neural patterns might reflect shared content in recalled memories.

Spatial resolution of neural signals—How refined was the spatial alignment of recollection patterns across brains? The alignment had to be robust enough to overcome imperfect registration (due to anatomical variation) across different brains. Our data indicate that the pattern alignment was strong enough to survive spatial transformation of brain data to a standard anatomical space. Recently, it was argued that relatively coarse organized

patterns (e.g., small eccentricity-related biases toward horizontal or vertical orientations within primary visual cortex) can underlie spatial pattern correlations in the brain[63]. Importantly, while the current results reveal a relatively coarse spatial structure that is shared across people (Fig. 6), they do not preclude the existence of finer spatial structure in the neural signal that may be captured when comparisons are made within-participant[35,64] or using more sensitive methods such as hyperalignment[35]. Furthermore, when we calculated movie-RDM vs. recall-RDM second-order correlations within and between brains, we found a region in the temporoparietal junction in which individual-unique aspects of neural patterns contributed reliably to reinstatement strength above and beyond the shared representation (though the fact that this region is small relative to the total area in which shared patterns were identified suggests that the between-brain comparison captured a substantial portion of the movie-recall pattern similarity). Using a similar RSA approach, Charest et al. (2014)[4] found individual-unique neural responses in inferior temporal cortex during visual object viewing. There were numerous differences between our study and the Charest et al. study, in stimulus content (Charest et al. used objects selected for personal significance), paradigm, and regions of interest; further work is needed to understand the factors that influence the balance of idiosyncratic and shared signals between brains.

Visual imagery—To what extent did spoken recollection in this study engage visual imagery? We observed extraordinarily rich recollection behavior (Fig. 1B, Table S2), in which participants managed to recount, in largely correct order, the details of most of the movie scenes; this suggested that participants were able to mentally replay virtually the entire movie, despite the absence of any external cues. However, movie-recall reinstatement effects were not found in low level visual areas, but instead were located in high level visual areas (Fig. S4[66,67]), and extensively in higher order brain regions outside of the visual system (Fig. 2). Our observation of reinstatement in high level visual areas is compatible with studies showing reinstatement in these regions during cued visual imagery[24,68,69]. 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. It has been suggested that the requirement to note high-resolution details is a key factor in eliciting activity in early visual cortex during visual imagery[70]. Thus, our findings do not conflict with studies showing that activity patterns in early visual cortex can be used to decode a simple image held in mind during a delay, in tasks that required vivid imagery of low-level visual features[71,72].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Isola, P.; Xiao, J.; Torralba, A.; Oliva, A. What makes an image memorable?; 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2011. p. 145-152.

2. Halbwachs, M. The Collective Memory. Harper & Row Colophon Books; 1980.

3. Sperber, D. Explaining Culture: A Naturalistic Approach. Blackwell Publishers; 1996.

4. Coman A, Hirst W. Cognition through a social network: The propagation of induced forgetting and practice effects. J. Exp. Psychol. Gen. 2012; 141:321–336. [PubMed: 21910558]

5. Roediger HL, Abel M. Collective memory: a new arena of cognitive study. Trends Cogn. Sci. 2015; 19:359–361. [PubMed: 25953047]

6. Raichle ME, et al. A default mode of brain function. Proc Natl Acad Sci U A. 2001; 98:676–682.

7. Hasson U. Intersubject synchronization of cortical activity during natural vision. Science. 2004; 303:1634–1640. [PubMed: 15016991]

8. Jaaskelainen PI, et al. Inter-Subject Synchronization of Prefrontal Cortex Hemodynamic Activity During Natural Viewing. Open Neuroimaging J. 2008; 2:14–19.

9. Wilson SM, Molnar-Szakacs I, Iacoboni M. Beyond Superior Temporal Cortex: Intersubject Correlations in Narrative Speech Comprehension. Cereb. Cortex. 2008; 18:230–242. [PubMed: 17504783]

10. Lerner Y, Honey CJ, Silbert LJ, Hasson U. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. J. Neurosci. 2011; 31:2906–2915. [PubMed: 21414912]

11. Honey CJ, Thompson CR, Lerner Y, Hasson U. Not Lost in Translation: Neural Responses Shared Across Languages. J. Neurosci. 2012; 32:15277–15283. [PubMed: 23115166]

12. Lahnakoski JM, et al. Synchronous brain activity across individuals underlies shared psychological perspectives. NeuroImage. 2014; 100:316–324. [PubMed: 24936687]

13. Simony E, et al. Dynamic reconfiguration of the default mode network during narrative comprehension. Nat. Commun. 2016

14. Regev M, Honey CJ, Simony E, Hasson U. Selective and Invariant Neural Responses to Spoken and Written Narratives. J. Neurosci. 2013; 33:15978–15988. [PubMed: 24089502]

15. Wang M, He BJ. A cross-modal investigation of the neural substrates for ongoing cognition. Front. Psychol. 2014; 5

16. Borges JL. Funes the Memorious. La Nacion. 1942

17. Wheeler ME, Petersen SE, Buckner RL. Memory's echo: Vivid remembering reactivates sensory-specific cortex. Proc. Natl. Acad. Sci. 2000; 97:11125–11129. [PubMed: 11005879]

18. Danker JF, Anderson JR. The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. Psychol. Bull. 2010; 136:87–102. [PubMed: 20063927]

19. Polyn SM. Category-Specific Cortical Activity Precedes Retrieval During Memory Search. Science. 2005; 310:1963–1966. [PubMed: 16373577]

20. Johnson JD, McDuff SGR, Rugg MD, Norman KA. Recollection, Familiarity, and Cortical Reinstatement: A Multivoxel Pattern Analysis. Neuron. 2009; 63:697–708. [PubMed: 19755111]

21. Kuhl BA, Rissman J, Chun MM, Wagner AD. Fidelity of neural reactivation reveals competition between memories. Proc. Natl. Acad. Sci. 2011; 108:5903–5908. [PubMed: 21436044]

22. Buchsbaum BR, Lemire-Rodger S, Fang C, Abdi H. The neural basis of vivid memory is patterned on perception. J. Cogn. Neurosci. 2012; 24:1867–1883. [PubMed: 22640391]

23. Wing EA, Ritchey M, Cabeza R. Reinstatement of Individual Past Events Revealed by the Similarity of Distributed Activation Patterns during Encoding and Retrieval. J. Cogn. Neurosci. 2014:1–13. [PubMed: 24047384]

24. Bird CM, Keidel JL, Ing LP, Horner AJ, Burgess N. Consolidation of Complex Events via Reinstatement in Posterior Cingulate Cortex. J. Neurosci. 2015; 35:14426–14434. [PubMed: 26511235]

25. Kriegeskorte N. Representational similarity analysis – connecting the branches of systems neuroscience. Front. Syst. Neurosci. 2008

26. Buckner RL, Andrews-Hanna JR, Schacter DL. The brain's default network: anatomy, function, and relevance to disease. Ann. N. Y. Acad. Sci. 2008; 1124:1–38. [PubMed: 18400922]

27. Rugg MD, Vilberg KL. Brain networks underlying episodic memory retrieval. Curr. Opin. Neurobiol. 2013; 23:255–260. [PubMed: 23206590]

28. Honey CJ, et al. Slow cortical dynamics and the accumulation of information over long timescales. Neuron. 2012; 76:423–434. [PubMed: 23083743]

29. Hasson U, Malach R, Heeger DJ. Reliability of cortical activity during natural stimulation. Trends Cogn. Sci. 2010; 14:40–48. [PubMed: 20004608]

30. Mitchell TM, et al. Learning to Decode Cognitive States from Brain Images. Mach. Learn. 2004; 57:145–175.

31. Poldrack RA, Halchenko YO, Hanson SJ. Decoding the Large-Scale Structure of Brain Function by Classifying Mental States Across Individuals. Psychol. Sci. 2009; 20:1364–1372. [PubMed: 19883493]

32. Shinkareva SV, Malave VL, Mason RA, Mitchell TM, Just MA. Commonality of neural representations of words and pictures. NeuroImage. 2011; 54:2418–2425. [PubMed: 20974270]

33. Kaplan JT, Meyer K. Multivariate pattern analysis reveals common neural patterns across individuals during touch observation. Neuroimage. 2012; 60:204–212. [PubMed: 22227887]

34. Rice GE, Watson DM, Hartley T, Andrews TJ. Low-Level Image Properties of Visual Objects Predict Patterns of Neural Response across Category-Selective Regions of the Ventral Visual Pathway. J. Neurosci. 2014; 34:8837–8844. [PubMed: 24966383]

35. Charest I, Kievit RA, Schmitz TW, Deca D, Kriegeskorte N. Unique semantic space in the brain of each beholder predicts perceived similarity. Proc. Natl. Acad. Sci. 2014 201402594.

36. Wandell BA, Dumoulin SO, Brewer AA. Visual Field Maps in Human Cortex. Neuron. 2007; 56:366–383. [PubMed: 17964252]

37. Formisano E, et al. Mirror-Symmetric Tonotopic Maps in Human Primary Auditory Cortex. Neuron. 2003; 40:859–869. [PubMed: 14622588]

38. Benson NC, et al. The Retinotopic Organization of Striate Cortex Is Well Predicted by Surface Topology. Curr. Biol. 2012; 22:2081–2085. [PubMed: 23041195]

39. Moser EI, Kropff E, Moser M-B. Place cells, grid cells, and the brain's spatial representation system. Annu Rev Neurosci. 2008; 31:69–89. [PubMed: 18284371]

40. O'Keefe J, Conway DH. Hippocampal place units in the freely moving rat: why they fire where they fire. Exp. Brain Res. 1978; 31:573–590. [PubMed: 658182]

41. Kriegeskorte N, et al. Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. Neuron. 2008; 60:1126–1141. [PubMed: 19109916]

42. Hassabis D, Maguire EA. Deconstructing episodic memory with construction. Trends Cogn. Sci. 2007; 11:299–306. [PubMed: 17548229]

43. Ranganath C, Ritchey M. Two cortical systems for memory-guided behaviour. Nat. Rev. Neurosci. 2012; 13:713–726. [PubMed: 22992647]

44. Ames DL, Honey CJ, Chow MA, Todorov A, Hasson U. Contextual Alignment of Cognitive and Neural Dynamics. J. Cogn. Neurosci. 2014; 27:655–664. [PubMed: 25244122]

45. Alba JW, Hasher L. Is memory schematic? Psychol. Bull. 1983; 93:203–231.

46. Kurby CA, Zacks JM. Segmentation in the perception and memory of events. Trends Cogn. Sci. 2008; 12:72–79. [PubMed: 18178125]

47. Baldassano C, et al. Discovering event structure in continuous narrative perception and memory. bioRxiv 81018. 2016

48. Buzsáki G, Moser EI. Memory, navigation and theta rhythm in the hippocampal-entorhinal system. Nat. Neurosci. 2013; 16:130–138. [PubMed: 23354386]

49. Hasson U, Ghazanfar AA, Galantucci B, Garrod S, Keysers C. Brain-to-brain coupling: a mechanism for creating and sharing a social world. Trends Cogn. Sci. 2012; 16:114–121. [PubMed: 22221820]

50. Zadbood A, Chen J, Leong YC, Norman KA, Hasson U. How we transmit memories to other brains: constructing shared neural representations via communication. bioRxiv 81208. 2016

Methods-Only References

51. McGuigan, P. A Study in Pink. Sherlock; 2010.

52. Fleischer, D. Let's All Go to the Lobby. Filmack; 1957.

53. Stephens GJ, Silbert LJ, Hasson U. Speaker�listener neural coupling underlies successful communication. Proc. Natl. Acad. Sci. 2010; 107:14425–14430. [PubMed: 20660768]

54. Silbert LJ, Honey CJ, Simony E, Poeppel D, Hasson U. Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. Proc. Natl. Acad. Sci. 2014; 111:E4687–E4696. [PubMed: 25267658]

55. Desikan RS, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage. 2006; 31:968–980. [PubMed: 16530430]

56. Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD. Decoding subject-driven cognitive states with whole-brain connectivity patterns. Cereb. Cortex. 2012; 22:158–165. [PubMed: 21616982]

57. Simony E, et al. Dynamic reconfiguration of the default mode network during narrative comprehension. Nat. Commun. 2016; 7:12141. [PubMed: 27424918]

58. Kriegeskorte N, Goebel R, Bandettini P. Information-based functional brain mapping. Proc. Natl. Acad. Sci. U. S. A. 2006; 103:3863–3868. [PubMed: 16537458]

59. Benjamini Y, Hochberg Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B-Methodol. 1995; 57:289–300.

60. Chen, P-H (Cameron), et al. Advances in Neural Information Processing Systems. Cortes, C.; Lawrence, ND.; Lee, DD.; Sugiyama, M.; Garnett, R., editors. Vol. 28. Curran Associates, Inc.; 2015. p. 460-468.

61. Naselaris T, Kay KN, Nishimoto S, Gallant JL. Encoding and decoding in fMRI. NeuroImage. 2011; 56:400–410. [PubMed: 20691790]

62. Mitchell TM, et al. Predicting Human Brain Activity Associated with the Meanings of Nouns. Science. 2008; 320:1191–1195. [PubMed: 18511683]

63. Freeman J, Heeger DJ, Merriam EP. Coarse-Scale Biases for Spirals and Orientation in Human Visual Cortex. J. Neurosci. 2013; 33:19695–19703. [PubMed: 24336733]

64. Kamitani Y, Tong F. Decoding the visual and subjective contents of the human brain. Nat. Neurosci. 2005; 8:679–685. [PubMed: 15852014]

65. Haxby JV, et al. A Common, High-Dimensional Model of the Representational Space in Human Ventral Temporal Cortex. Neuron. 2011; 72:404–416. [PubMed: 22017997]

66. Wang L, Mruczek REB, Arcaro MJ, Kastner S. Probabilistic Maps of Visual Topography in Human Cortex. Cereb. Cortex bhu277. 2014

67. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. Large-scale automated synthesis of human functional neuroimaging data. Nat. Methods. 2011; 8:665–670. [PubMed: 21706013]

68. Cichy RM, Heinzle J, Haynes J-D. Imagery and Perception Share Cortical Representations of Content and Location. Cereb. Cortex. 2012; 22:372–380. [PubMed: 21666128]

69. St-Laurent M, Abdi H, Buchsbaum BR. Distributed Patterns of Reactivation Predict Vividness of Recollection. J. Cogn. Neurosci. 2015:1–19.

70. Kosslyn SM, Thompson WL. When is early visual cortex activated during visual mental imagery? Psychol. Bull. 2003; 129:723–746. [PubMed: 12956541]

71. Harrison SA, Tong F. Decoding reveals the contents of visual working memory in early visual areas. Nature. 2009; 458:632–635. [PubMed: 19225460]

72. Serences JT, Ester EF, Vogel EK, Awh E. Stimulus-specific delay activity in human primary visual cortex. Psychol Sci. 2009; 20:207–214. [PubMed: 19170936]
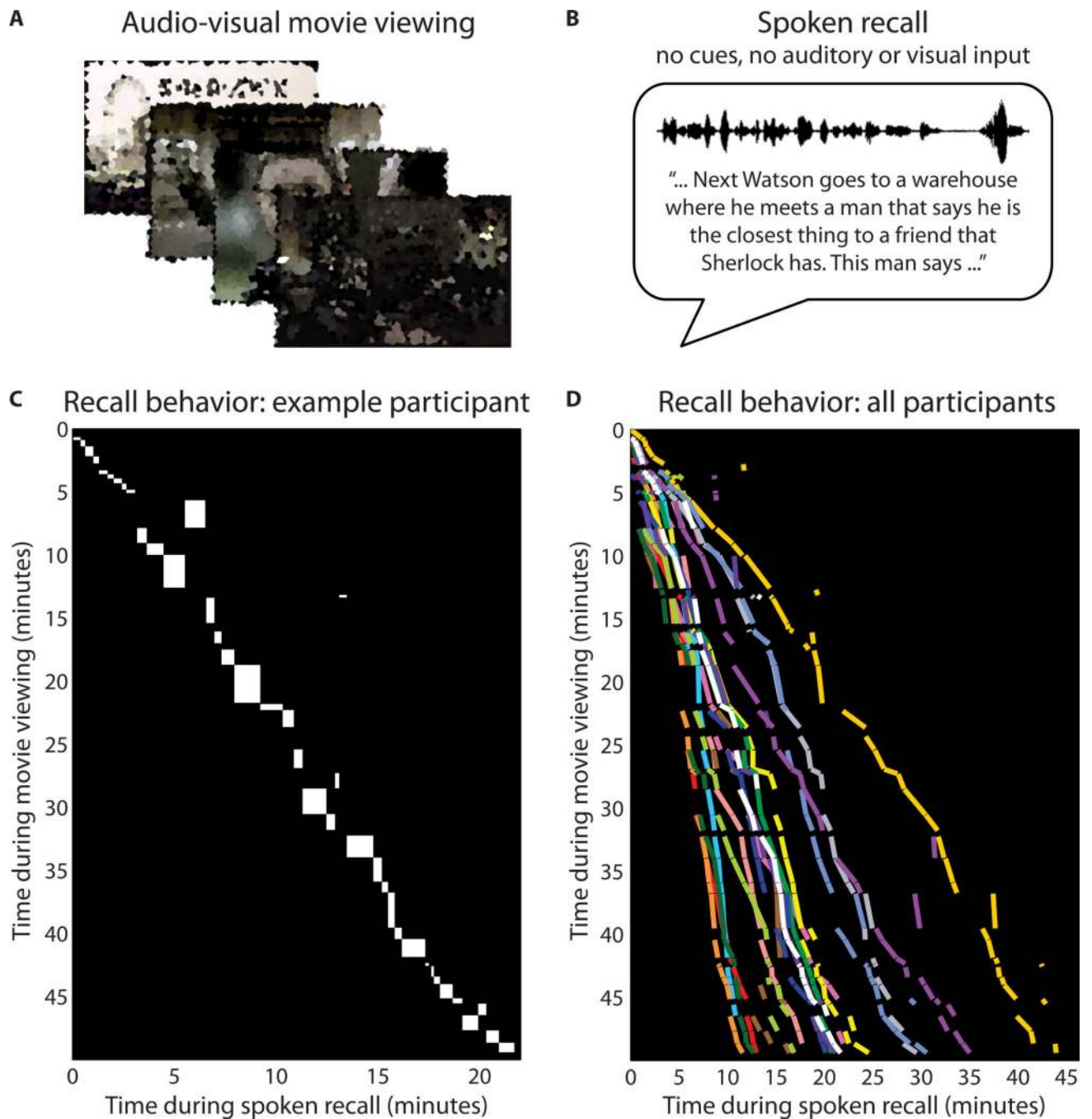
A) Audio-visual movie viewing

B) Spoken recall
no cues, no auditory or visual input

"... Next Watson goes to a warehouse where he meets a man that says he is the closest thing to a friend that Sherlock has. This man says ..."

C) Recall behavior: example participant

D) Recall behavior: all participants

Figure 1. Experiment design and behavior

A) In Run 1, participants viewed a 50-minute movie, BBC's Sherlock (episode 1). Images in the figure are blurred for copyright reasons; in the experiment, movies were shown at full clarity. **B)** In the immediately following Run 2, participants verbally recounted aloud what they recalled from the movie. Instructions to "retell what you remember in as much detail as you can" were provided before the start of the run. No form of memory cues, time cues, or any auditory/visual input were provided during the recall session. Speech was recorded via microphone. **C)** Diagram of scene durations and order for movie viewing and spoken recall

in a representative participant. Each rectangle shows, for a given scene, the temporal position (location on y-axis) and duration (height) during movie viewing, and the temporal position (location on x-axis) and duration (width) during recall. **D)** Summary of durations and order for scene viewing and recall in all participants. Each line segment shows, for a given scene, the temporal position and duration during movie viewing and during recall; i.e., a line segment in [D] corresponds to the diagonal of a rectangle in [C]. Each color indicates a different participant (N=17). See also Tables S1, S2.
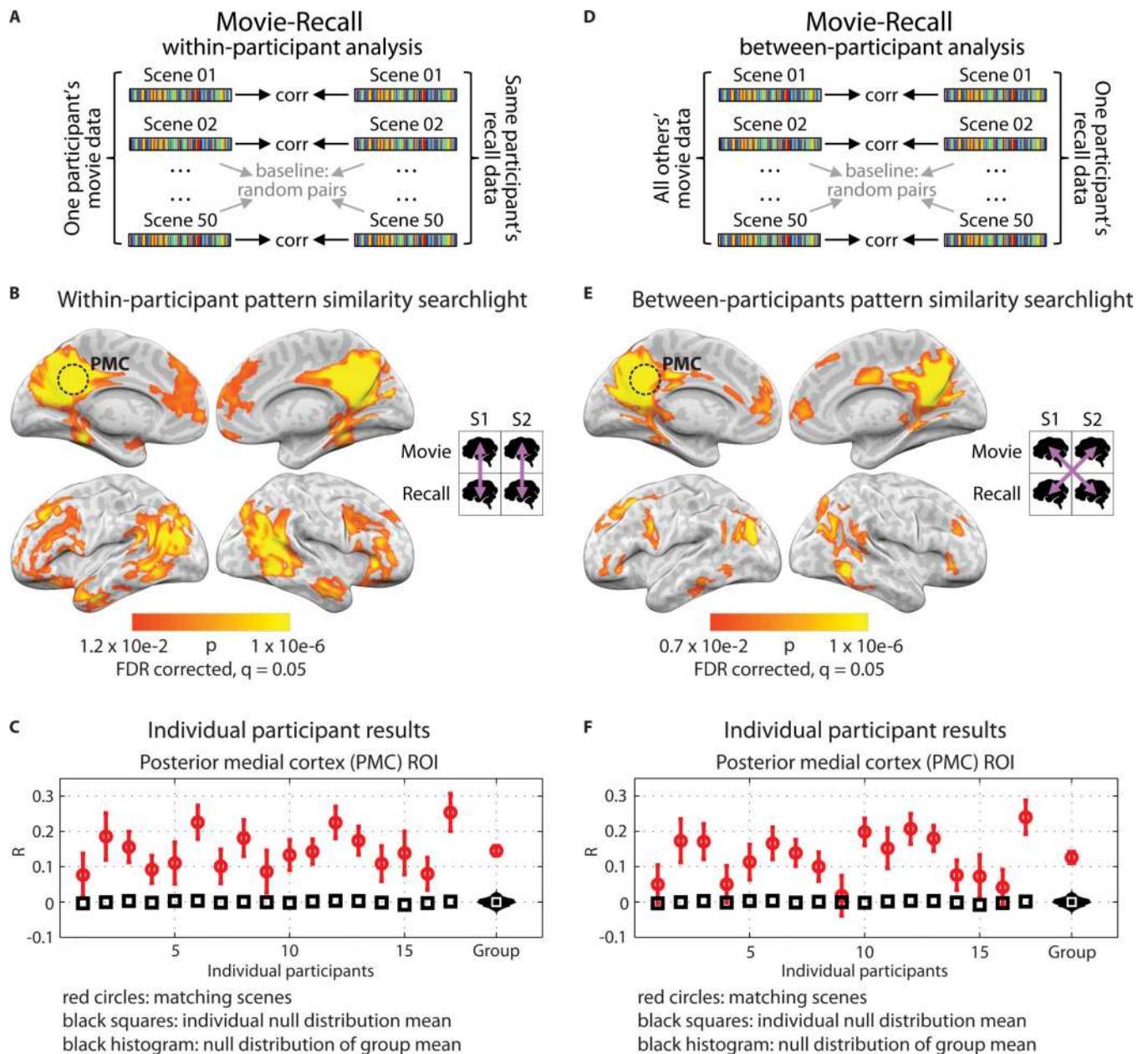
Figure 2. Pattern similarity between movie and recall

A) Schematic for within-participant movie-recall (reinstatement) analysis. BOLD data from the movie and from the recall sessions were divided into scenes, then averaged across time within-scene, resulting in one vector of voxel values for each movie scene and one for each recalled scene. Correlations were computed between matching pairs of movie/recalled scenes *within* participant. Statistical significance was determined by shuffling scene labels to generate a null distribution of the participant average. **B)** Searchlight map showing where significant reinstatement was observed; FDR correction q = 0.05, p = 0.012. Searchlight was a 5×5×5 voxel cube. **C)** Reinstatement values for all 17 participants in independently-defined PMC. 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)** Schematic for between-participants movie-recall analysis. Same as [A], except that correlations were computed between every matching pair of movie/recall scenes *between* participants. **E)** Searchlight map showing regions where significant between-participants movie-recall similarity was observed; FDR correction q = 0.05, p = 0.007. **F)** Reinstatement values in PMC for each participant in the between-participants analysis, same notation as [C].
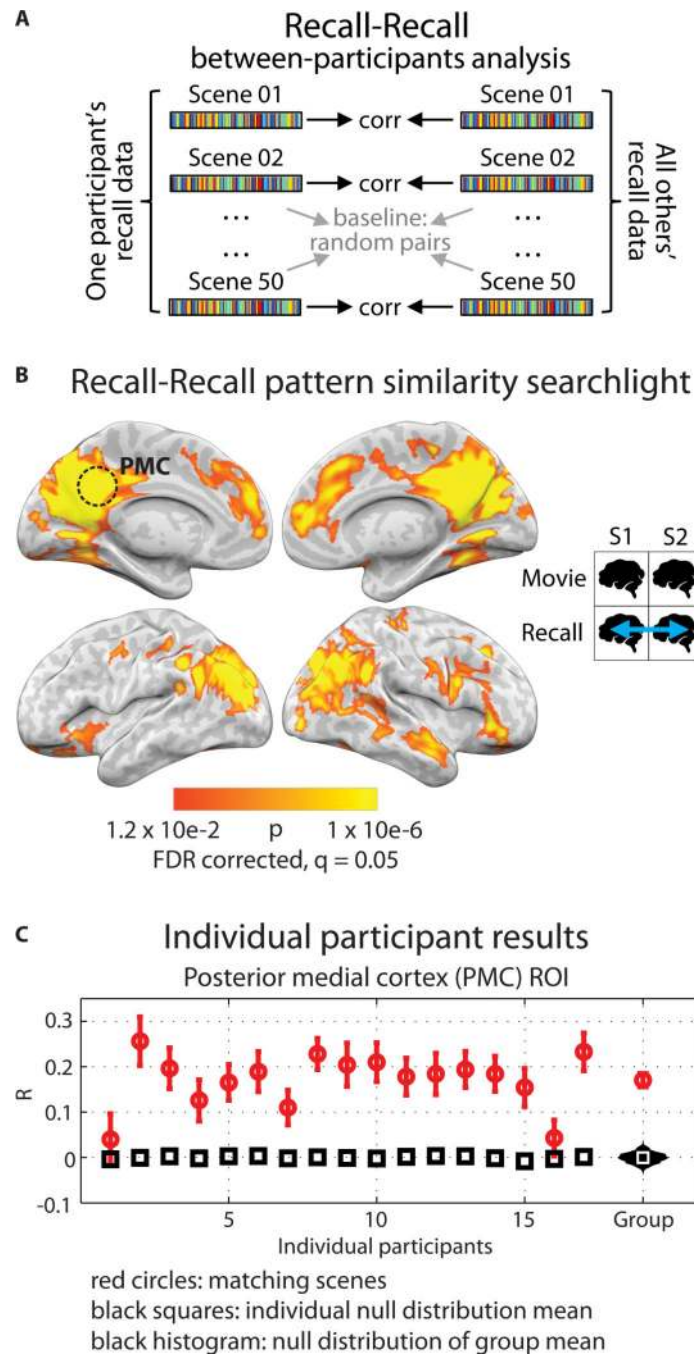
Figure 3. Between-participants pattern similarity during spoken recall
A) Schematic for between-participants recall-recall analysis. BOLD data from the recall sessions were divided into matching scenes, then averaged across time within each voxel, resulting in one vector of voxel values for each recalled scene. Correlations were computed between every matching pair of recalled scenes. Statistical significance was determined by shuffling scene labels to generate a null distribution of the participant average. **B)** Searchlight map showing regions where significant recall-recall similarity was observed; FDR correction at q = 0.05, p = 0.012. Searchlight was a 5×5×5 voxel cube. **C)** Recall-recall

correlation values for all 17 participants in independently-defined PMC. 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.
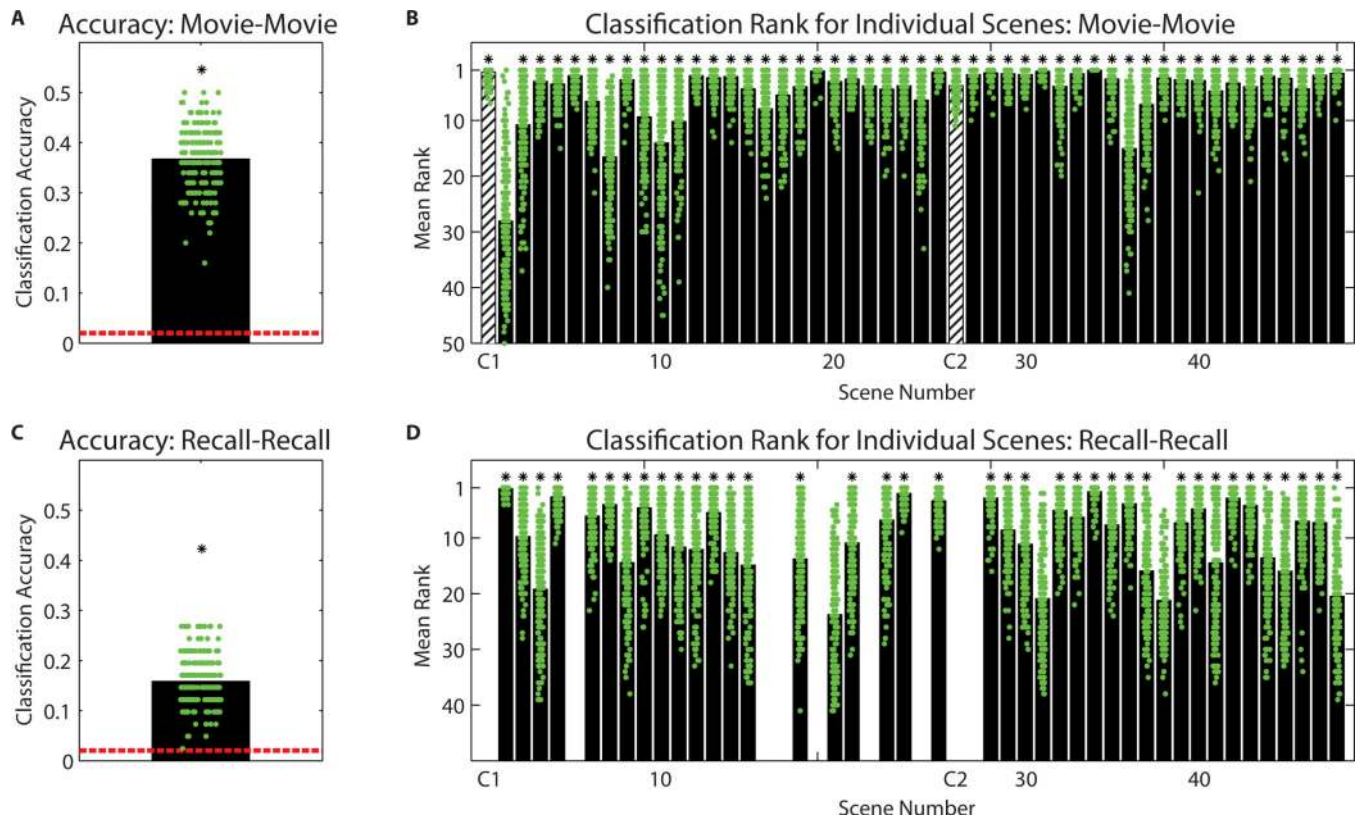
Figure 4. Classification accuracy

A) Classification of movie scenes between brains. Participants were randomly assigned to one of two groups (N=8 and N=9), an average was calculated within each group, and data were extracted for the PMC ROI. Pairwise correlations were calculated between the two group means for all 50 movie scenes. Accuracy was calculated as the proportion of scenes correctly identified out of 50. The entire procedure was repeated using 200 random combination of two groups sized N=8 and N=9 (green markers), and an overall average calculated (36.7%, black bar; chance level [2.0%] plotted in red). **B)** Classification rank for individual movie scenes (i.e., the rank of the matching scene correlation in the other group among all 50 scene correlations). Green markers show the results from each combination of two groups sized N=8 and N=9; black bars show the average over all group combinations, 4.8 on average. (* indicates values passing the FDR-corrected threshold of q = 0.001.) Striped bars indicate introductory video clips at the beginning of each functional scan (see Methods). **C)** Classification of recalled scenes between brains. Same analysis as in (A) except that sufficient data were extant for 41 scenes. Overall classification accuracy was 15.8% (black bar, chance level 2.4%). **D)** Classification rank for individual recalled scenes, 9.5 on average. (* indicates values passing the FDR-corrected threshold of q = 0.001.)
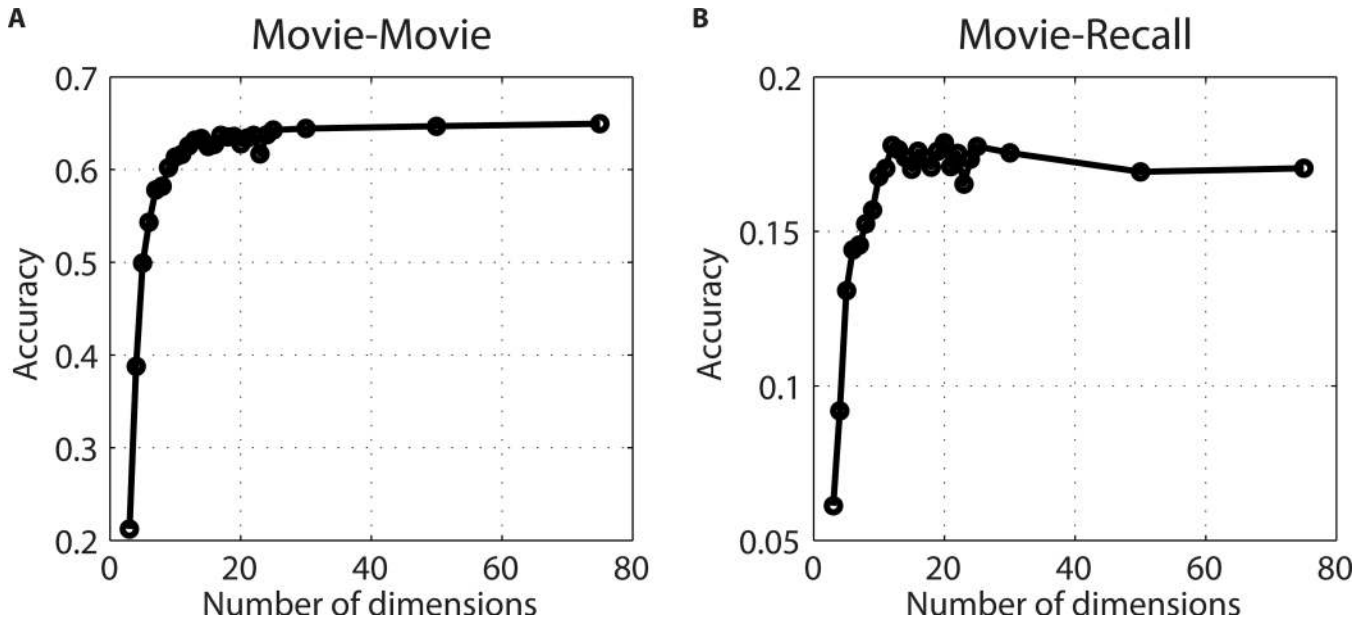
Figure 5. Dimensionality of shared patterns

In order to quantify the number of distinct dimensions of the spatial patterns that are shared across brains and can contribute to the classification of neural responses, we used the Shared Response Model (SRM). This algorithm operates over a series of data vectors (in this case, multiple participants' brain data) and finds a common representational space of lower dimensionality. Using SRM in the PMC region, we asked: when the data are reduced to k dimensions, how does this affect scene-level classification across brains? How many dimensions generalize from movie to recall? **A)** Results when using the movie data in the PMC region (movie-movie). Classification accuracy improves as the number of dimensions k increases, starting to plateau around 15, but still rising at 50 dimensions. (Chance level 0.04.) **B)** Results when training SRM on the movie data and then classifying recall scenes across participants in the PMC region (recall-recall). Classification accuracy improves as the number of dimensions increases, with maximum accuracy being reached at 12 dimensions. Note that there could be additional shared dimensions, unique to the recall data, that would not be accessible via these analyses.
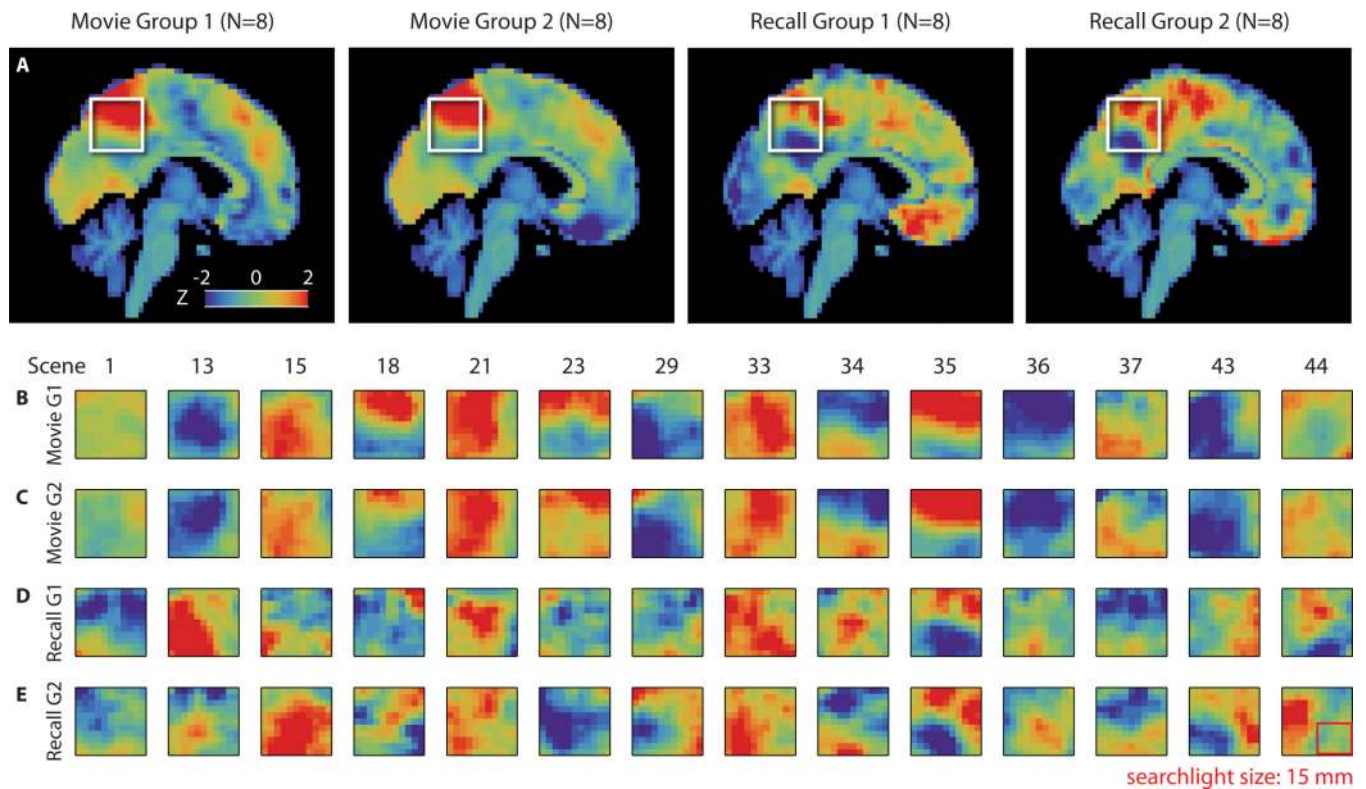
Figure 6. Scene-level pattern similarity between individuals
Visualization of the signal underlying pattern similarity between individuals, for fourteen scenes that were recalled by all sixteen of the participants in these groups, are presented in [B–E]. See Fig. S5 for correlation values for all scenes. **A)** In order to visualize the underlying signal, we randomly split the movie-viewing data into two independent groups of equal size (N=8 each) and averaged BOLD values across participants within each group. An average was made in the same manner for the recall data using the same two groups of eight. These group mean images were then averaged across timepoints and within scene, exactly as in the prior analyses, creating one brain image per group per scene. Sagittal view of these average brains during one representative scene (36) of the movie is shown for each group. Average activity in a posterior medial area (white box in [A]) on the same slice for the fourteen different scenes for Movie Group 1 **(B)**, Movie Group 2 **(C)**, Recall Group 1 **(D)**, and Recall Group 2 **(E)**. Searchlight size shown as a red outline. Our data indicate that cross-participant pattern alignment was strong enough to survive spatial transformation of brain data to a standard anatomical space. While the current results reveal a relatively coarse spatial structure that is shared across people, they do not preclude the existence of finer spatial structure in the neural signal that may be captured when comparisons are made within-participant or using more sensitive methods such as hyperalignment. Further work is needed to understand the factors that influence the balance of idiosyncratic and shared signals between brains. See *Methods: Spatial resolution of neural signals*.
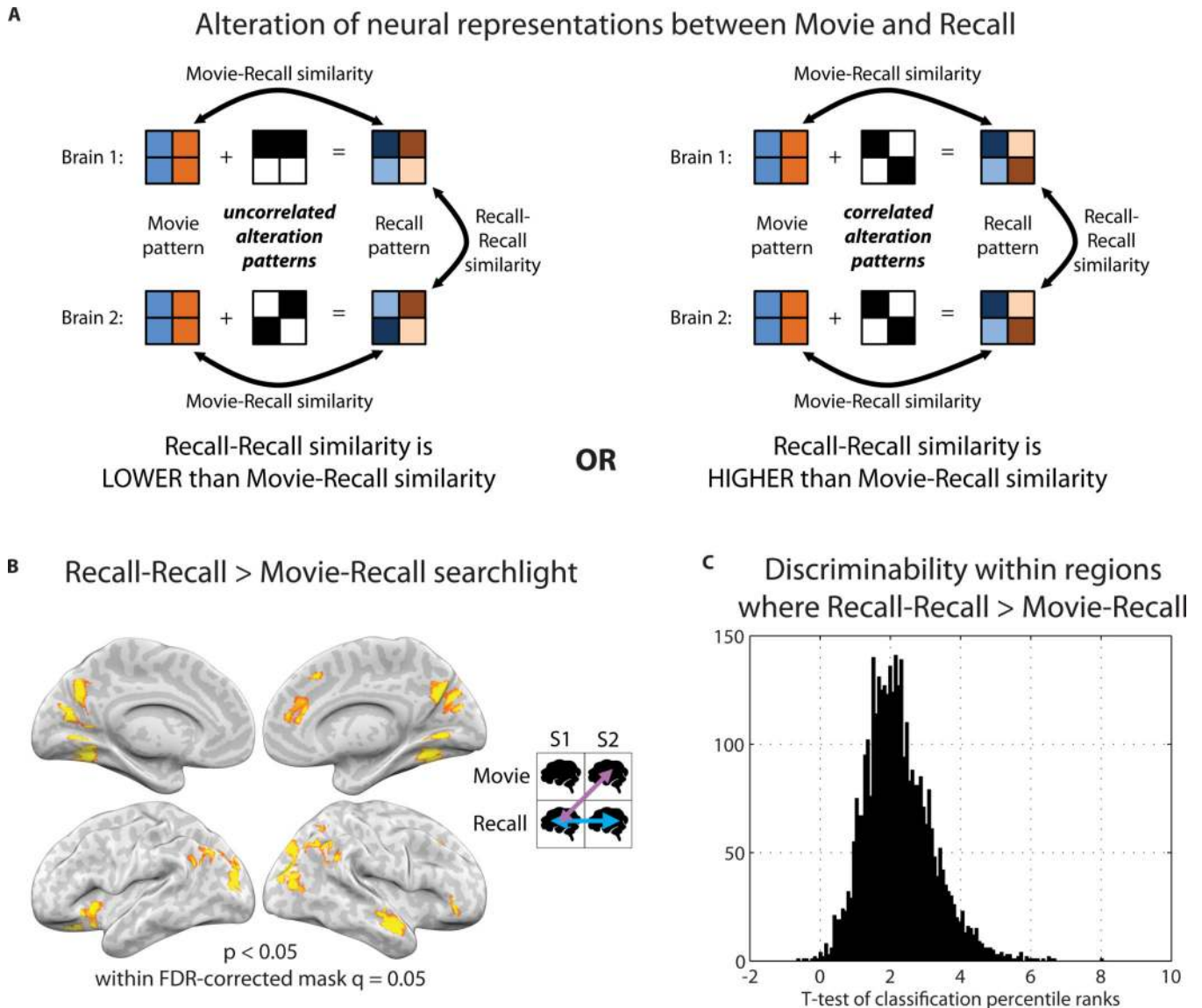
Figure 7. Alteration of neural patterns from perception to recollection

A) Schematic of neural activity patterns during a movie scene being modified into activity patterns at recall. For each brain, the neural patterns while viewing a given movie scene are expressed as a common underlying pattern. Each of these Movie patterns is then altered in some manner to produce the Recall pattern. *Left panel*: If patterns are changed in a unique way within each person's brain, then each person's movie pattern is altered by adding an "alteration" pattern that is *uncorrelated* with the "alteration" patterns of other people. In this scenario, Recall patterns necessarily become *more dissimilar* to the Recall patterns of other people than to the Movie pattern. *Right panel*: Alternatively, if a systematic change is occurring across people, each Movie pattern is altered by adding an "alteration" pattern that is *correlated* with the "alteration" patterns of other people. Thus, Recall patterns for a given scene may become *more similar* to the Recall patterns of other people than to the Movie pattern. **B)** Searchlight map showing regions where recall-recall similarity was significantly greater than between-participants movie-recall similarity, i.e., where the map from Fig. 3B

was stronger than the map from Fig. 2E. The analysis revealed regions in which neural representations changed in a systematic way across individuals between perception and recollection. **C)** We tested whether each participant's individual scene recollection patterns could be classified better using 1) the movie data from other participants, or 2) the recall data from other participants. A t-test of classification rank was performed between these two sets of values at each searchlight shown in (B). Classification rank was higher when using the recall data as opposed to the movie data in 99% of such searchlights. Histogram of t-values is plotted.
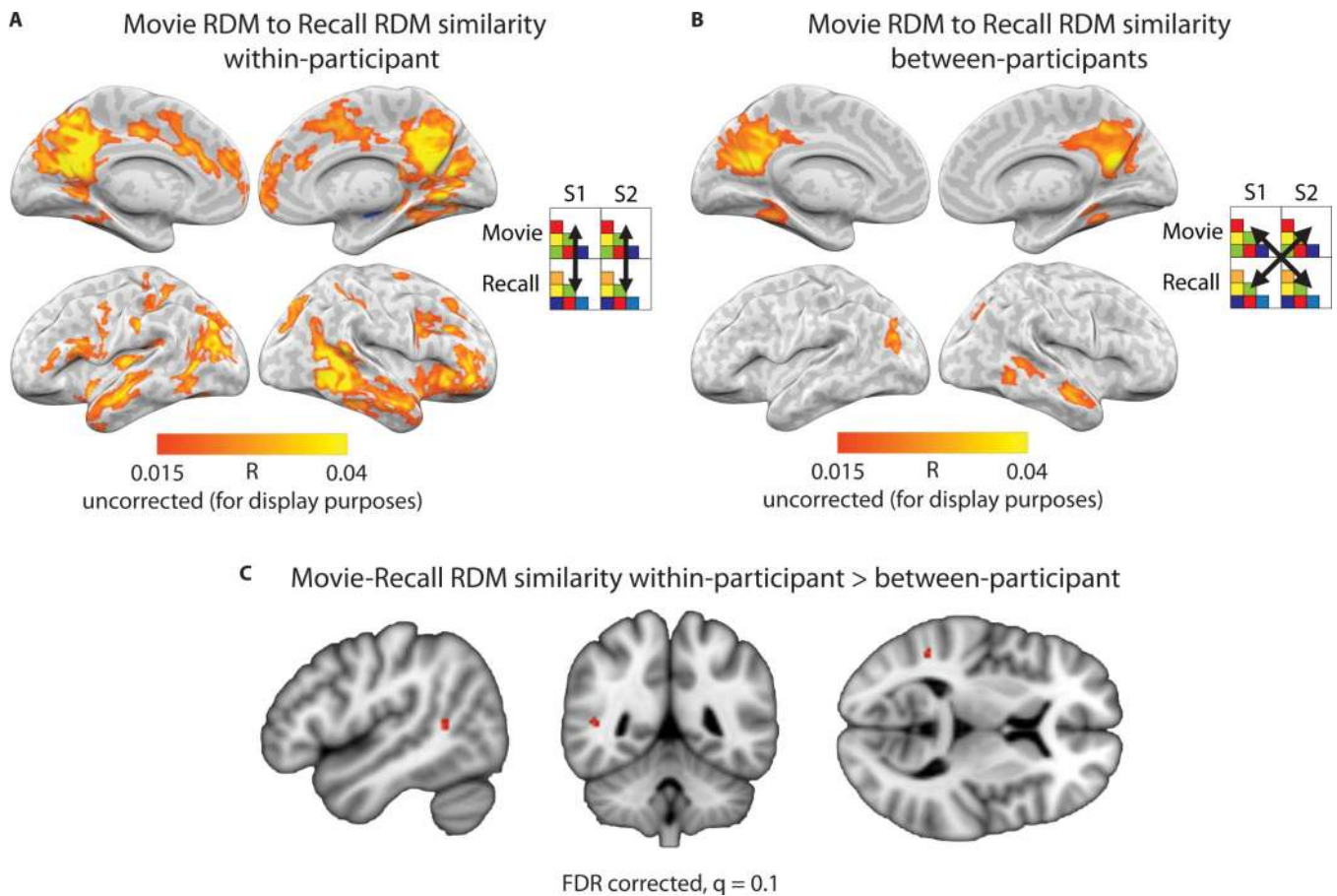
Figure 8. Reinstatement in individual participants vs. between participants
A) Searchlight analysis showing similarity of representational dissimilarity matrices (RDMs) within-participant across the brain. Each RDM was composed of the pairwise correlations of patterns for individual scenes in the movie ("movie-RDM") and separately during recall ("recall-RDM"). Each participant's movie-RDM was then compared to his or her own recall-RDM (i.e., within-participant) using Pearson correlation. The average searchlight map across 17 participants is displayed. **B)** Searchlight analysis showing movie-RDM vs. recall-RDM correlations between participants. The average searchlight map across 272 pairwise combinations of participants is displayed. **C)** The difference was computed between the within-participant and between-participant maps. Statistical significance of the difference was evaluated using a permutation analysis and FDR corrected at a threshold of q=0.05. A cluster of two voxels located in the temporo-parietal junction survived correction (map shown at q=0.10 for visualization purposes, 5-voxel cluster).