Title: Shared contextual knowledge strengthens inter-subject synchrony and pattern similarity in the semantic network

Authors: Petar P. Raykov, James L. Keidel, Jane Oakhill, Chris M. Bird

Affiliation: School of Psychology, University of Sussex, Falmer BN1 9QH, UK

Corresponding author: Petar P. Raykov, School of Psychology, University of Sussex, Falmer BN1 9QH, UK. Email: P.Raykov@sussex.ac.uk

Pages: 27; **Figures:** 4; **Words:** Introduction: 1342, Discussion: 1362

Conflict of Interest: "The authors declare no competing financial interests."

Acknowledgements: This work was supported by the European Research Council (337822 to C.M.B.) and an Economic and Social Research Council studentship.

Abstract

Understanding what is happening during an event can be helped by prior knowledge of the specific context. However, the effect of contextual knowledge on neural representations of events, and relatedly, how shared knowledge impacts on the similarity of neural processing of events across individuals is poorly understood. To investigate this, we manipulated the availability of knowledge about the narrative content of events while keeping all other aspects equally familiar. The presence of narrative knowledge boosted subjective ratings of coherency and objective measures of memory performance. Narrative knowledge had two effects on fMRI markers of neural processing: (1) it strengthened temporal inter-subject correlations in ventromedial prefrontal cortex as well as left angular and inferior frontal gyri, and (2) it also increased spatial inter-subject pattern similarity in the bilateral anterior temporal lobes. We propose that shared narrative knowledge constrains participants' interpretation of the videos and thereby results in greater alignment of neural processing of the events. We propose a division of labour between semantic control brain regions (VMPFC, IFG and AG), which process the moment-by-moment content of the narrative, and the ATL which represents the overarching narrative gist of an event.

Introduction

Understanding events as they unfold requires our attention over multiple timescales (Hasson, Chen, & Honey, 2015). Following a conversation or a complex sequence of actions involves attention to moment-by-moment changes in incoming information. It is also thought that we extract and represent more slowly changing features of the event such as the location and the people present as overarching "event models" (Zacks et al., 2007). These event models provide coherence to our experiences, even in the face of rapidly changing sensory input (Zacks et al., 2007). Our prior knowledge affects both our ability to comprehend conversations and to construct event models. However, the effects of prior knowledge on the brain systems that underpin event processing are poorly understood.

Recently, MRI scanning combined with the use of naturalistic materials, has begun to shed light on the brain regions that support processing of lifelike events (Ben-Yakov et al., 2013; Bird et al., 2015; J. Chen et al., 2017; Zacks et al., 2001). Various studies have described both transitory responses, corresponding to discrete events within a continuous video (Zacks et al., 2001), as well as patterns of activity that persist over time and which may correspond to representations of a currently active event model (Baldassano et al., 2018; Bird et al., 2015; Oedekoven et al., 2017).

Interestingly, effects are not only seen within individual subjects but some aspects of event processing are shared across brains. Hasson and colleagues (2004) found significant correlations in the timecourse of activity elicited by a movie across subjects (inter-subject correlations: ISC). In more recent work, Chen et al. (2017) demonstrated the existence of shared multivoxel representations of events. Participants watched and recalled a full movie in the scanner and the authors analysed multivoxel patterns of activity averaged across sub-events within the movie. The authors found that activity patterns were correlated across participants, both when they watched, or recalled, specific events from the movie (inter-subject pattern similarity: ISPS). Similar effects can been seen when people construct an event in their mind's eye while listening to it being described (Zadbood et al., 2017). The authors suggested that the "units" of information being shared across participants were event models.

Effects in these studies are typically seen across large regions of sensory cortex as well as throughout the brain's so-called default mode network (DMN; Raichle et al., 2001). However, a major challenge is to dissect what are often widespread effects into the processes and representations supported by different brain regions. For example, the complexity of event models makes it difficult to confidently ascribe pattern similarity effects to any particular aspect of the model. Thus, observed effects could reflect very broad schematic representations of the content of the scenes (e.g. 'a conversation in a café', or 'a chase down the street'). While such schematic representations provide an essential scaffold for comprehending and remembering short events, by themselves they provide no means of differentiating similar events (e.g. talking to different people on different days while having a coffee in the same café). Instead, much more specific information is required, such as the specific actions being carried out by the actors or the topic of the conversation.

A number of classic studies have investigated narrative comprehension of prose passages by manipulating prior knowledge (e.g. Bransford & Johnson, 1972; Dooling & Lachman, 1971; Maguire, Frith, & Morris, 1999). In such experiments, ambiguous or confusing narratives can be understood with the provision of specific background contextual information before the narrative is read. Ames and colleagues (2015) used such a manipulation to investigate the effect of prior contextual knowledge on ISCs. The authors found that participants sharing the same background knowledge showed increased synchronization in BOLD response in the ventromedial prefrontal cortex (vmPFC) and posterior cingulate cortex (PCC). However, the key contrast in such paradigms involves comparing meaningful passages with passages that are very difficult to understand and it is therefore unclear whether the effects reflect this more fundamental difference in the conditions (coherent versus incoherent narratives).

In a conceptually similar study, van Kesteren et al. (2010) investigated how prior knowledge affects the encoding of novel naturalistic events depicted in a movie. Participants watched the first part of an extended movie either in a temporally scrambled order or in an intact order and, on the following day, watched the last 15 minutes of the movie in the scanner. In this study, the end of the movie is interpretable by itself, but a deeper understanding is possible if the first part was

viewed in an intact as opposed to a scrambled sequence. The participants who had watched the first part of the movie in the intact condition showed higher ISC in the vmPFC. Taken together, the Ames et al. and van Kesteren et al. studies suggest that providing “common ground” for participants to interpret a complex narrative leads to neural processing within the vmPFC that is more consistent across individuals. However, it would be surprising if this were the only region associated with such effects, since prior knowledge would expect to comprise linked semantic concepts that would be represented throughout the brain’s semantic network (SN) (Binder et al., 2009).

A recent study more directly examined how the availability of prior knowledge affects processing of events (Keidel et al., 2017). Participants saw first and second half clips taken from TV situational comedies (sitcoms). The second half clips could either be a direct continuation of the first half clips, or be clips taken from the same show involving the same characters and location, but a different episode of the show (and consequently depicting a different narrative storyline). The provision of narrative context resulted in increased activation in regions of the SN, including the inferior frontal gyrus (IFG), middle temporal gyrus (MTG) and angular gyrus (AG). It was argued that this reflected the reactivation of linked semantic concepts that were created when watching the first half videos. However, because the non-continuation clips clearly showed a different event, the activation differences may have reflected other differences between the conditions, such as the appearance of the characters and the contents of the scene. This study only examined within-subject univariate effects.

In the current study, we examine how prior narrative context affects the cognitive and neural substrates of event processing. We presented participants with video excerpts taken from sitcoms that were divided into first and second halves. In the high context condition, the first-half clips were presented with normal speech, whereas in the low context condition we spectrally rotated the speech to render it unintelligible. Under this design, all the pairs of clips were clearly taken from the same occasion and differed only in the provision of the narrative context. Importantly, all second-half clips, which were the focus of the analyses, were the same for both conditions, allowing us to detect differences in inter-subject synchronization and pattern similarity while participants viewed an identical set of

stimuli. Both temporal synchronisation and spatial pattern analyses require the comparison of activity across more than one event - either two or more events within a subject, or the same event across multiple subjects. Since our manipulation is hypothesised to affect processing only during the second-half clips, these analyses can only be carried out across subjects.

The study aims to test a number of hypotheses. First, we predict that provision of comprehensible narrative contextual information will aid the comprehension and memory for the subsequent event. Second, a narrative context will constrain the interpretation of the second-half clips, resulting in greater inter-subject similarity within semantic processing regions. ISCs examine the temporally varying BOLD signal and are therefore suited to detecting moment-to-moment processing if the narrative content. These effects may therefore by particularly apparent in regions associated with semantic control processing (Noonan et al., 2013). Conversely, ISPS is based on spatial patterns of activity averaged across the video clips and is therefore more likely to detect stable aspects of the narrative (e.g. overarching themes) that contribute to the overall event model. These effects are predicted to be apparent in regions representing abstract and amodal semantic knowledge (Fairhall & Caramazza, 2013; Rice et al., 2015).

Methods

Participants

Twenty-four right-handed native English speakers with normal or corrected to normal vision were included in the study. Four participants were not included in the final analysis due to artefacts in the MRI scans and 1 further participant did not complete the experiment. One participant had corrupted post-scanning behavioural data for the memory questions and was not included in the behavioural analysis. The project was approved by the Brighton and Sussex Medical School Research Governance and Ethics Committee (RGEC) and all participants gave informed consent and were paid £20.

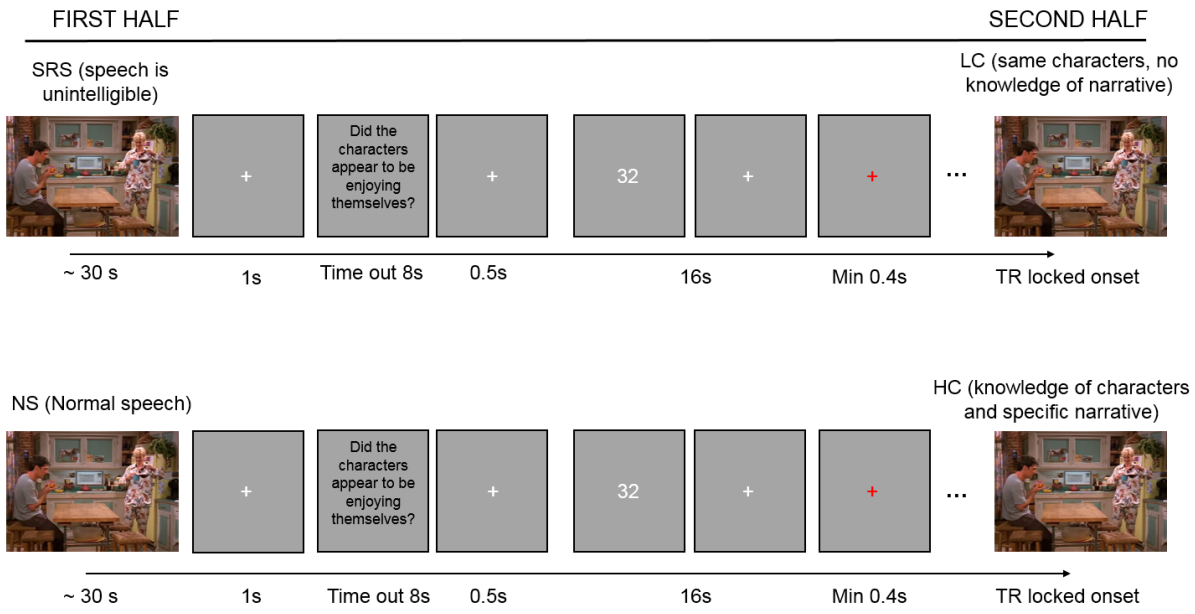

Figure 1 Schematic of study design. Participants viewed videos from unfamiliar TV sitcoms that were divided into two halves. Ten of the first halves had comprehensible speech, whereas the other 10 had unintelligible speech, created by spectrally rotating the audio. Videos were counterbalanced across participants in a within-subjects design. SRS – spectrally rotated speech; NS – normal speech; HC – high context; LC – low context.

Stimuli

Twenty video clips from different US and UK television shows were used in the experiment. The TV shows originally aired between 1970 and 2003 and were selected to be unfamiliar to our sample. Each video was divided into first and second halves. The scene location and characters remained constant across the two halves. For our main experimental manipulation, the speech in 10 of the first half videos was made unintelligible. This was done with Praat (version 6.0.15) (Boersma, 2001) by spectrally rotating the audio of the videos with a sinusoidal function with maximum frequency of 4 KHz. The audio for all videos was scaled to have the same mean decibel intensity. The mean duration of all the excerpts was 32.47(±3.88) seconds. The first half videos (30.57±4.38) were on average shorter than the second half videos (34.37±2.02). Three different video clips, all with spectrally rotated audio, were used in the practice task with total video length of 4.36 minutes. Four lists were created to fully counterbalance the conditions and presentation order.

Procedure

Participants were informed they would see first- and second-half videos in four separate runs (five videos in each). They were told that the speech in some of the first-half videos would be unintelligible, and were asked to watch the clips as they would watch television at home. Participants were also informed that their memory for the clips would be tested after the scanning session. They completed an 8-minute practice session outside the scanner, to familiarise themselves with the task and the sound of spectrally rotated speech.

The scanning session started with a 4-minute resting state scan (data not reported here). Following this, the task was presented in four functional runs, each approximately 10 minutes long. Each run consisted of 5 first half and 5 second half clips and every clip was followed by a question and an active baseline task. In each run there were either two or three clips with rotated speech. Movie onsets were time-locked to the repetition time (TR). After each clip, a white fixation cross was presented for 1 second. Each video was followed by a question about the general relationship between the characters in the scene (e.g. Did the characters appear to be enjoying themselves?). Behavioural piloting confirmed that participants could

answer the questions regardless of whether speech was rotated or not. The questions were presented for 8 seconds or until participants made a Yes/No response and were followed by a 500 milliseconds fixation. Participants made an odd/even number judgement during an active baseline task (Stark & Squire, 2001), which was presented in 16-second blocks between clips. Six randomly chosen numbers, between 1 and 98, were each presented for 2 seconds, followed by a fixation cross presented for 667 milliseconds. After the baseline task a red fixation cross was presented for a minimum of 400 milliseconds to signal that the next video was about to begin.

Outside of the scanner, participants completed a questionnaire about their familiarity with the shows. Only 6 participants reported any familiarity with 1-3 of the shows and none were familiar with the particular scenes used in the experiment. Afterwards, participants carried out a computer-based memory task, in which the first 4-6 seconds of the second half videos were presented as a memory cue. Participants were then asked to rate from 1 to 10: (1) the vividness of their memory of the video, (2) how coherent they found the story in the videos, (3) how engaging they found the video. Participants were also asked an open-ended memory question specific to each second half video (e.g. What was the address on his chest written in?). All memory questions concerned only material presented in the second half of the videos.

MRI acquisition

Data were acquired on a 1.5 T Siemens Avanto MRI scanner. Functional T2* weighted BOLD-sensitive images were acquired with EPI sequence with the following parameters: FOV = 192mm, TR = 2.62 seconds, TE = 42 milliseconds, 90 degree flip angle, slice thickness = 3mm, 35 interleaved ascending slices with .6 mm gap, and 3.0x3.0x3.0 mm voxels. A high resolution T1-weighted image was acquired with the following parameters: FOV 256mm, TR = 2.73 seconds, TE = 3.57ms, 1.0x1.0x1.0mm voxel size.

Image pre-processing

All EPI images were pre-processed using SPM 12 (Wellcome Department of Imaging Neuroscience, London, UK). Field maps were used to correct for image distortions and susceptibility-by-movement effects using the Realign and Unwarp

option (Andersson, Hutton, Ashburner, Turner, & Friston, 2001; Hutton et al., 2002). All EPI images were aligned to the first image of the first session. The anatomical image of each subject was co-registered to their mean realigned EPI image. The anatomical images were then segmented into grey and white matter maps. Anatomical and EPI images were normalized to the MNI space using DARTEL (Ashburner, 2007) and smoothed with an 8 mm FWHM kernel. Images for the inter-subject pattern similarity were pre-processed as above with the exception that a 6 mm FWHM smoothing kernel was applied to the normalised images, as previously used by Chen et al. (2017).

Data Analysis

Data were analysed with SPM 12, the CoSMoMVPA toolbox (Oosterhof, Connolly, & Haxby, 2016) and custom scripts in MATLAB (Version 2016b, The MathWorks, Inc., Natick, MA, USA). All analyses were conducted on MNI normalised images within a grey matter mask. Segmentation of the high-resolution structural images provided us with grey matter tissue probability map for each subject. These probability maps were normalised to MNI, averaged across participants. The averaged mask was smoothed with an 8mm FWHM kernel. We selected all voxels within this average probability map higher than a threshold of 0.3. To describe and visualise our data we used the Bspmview toolbox (www.bobspunt.com/bspmview), which implements the MNI coordinates from the Anatomical Automatic Labelling 2 toolbox for SPM 12. Significance was tested with a one-sample random effects *t* test against zero. Unless otherwise stated, images were whole brain cluster corrected for FWE $p < 0.05$ at voxel height-defining threshold of $p < 0.001$.

GLM analysis

A single task regressor for each of the four conditions (SRS, NS, LC and HC: see Figure 1) was included in the first-level models. For all GLM first level models the questions after each video were modelled with a single regressor of no interest and the odd/even number judgment task was left unmodelled to represent the implicit baseline. A block design first level analysis was conducted to replicate previous findings. In this analysis, all video stimuli were modelled with boxcar functions whose durations matched the stimulus duration. The models also included

the six motion parameters, a regressor for the mean session effects, and a high-pass filter with a cut-off of 1/128 Hz. We also ran an analysis identical to the above, but modelled only the onset of the videos with a gamma function, rather than including the whole duration of the video. This was done to replicate previous findings by Keidel et al. (2017). First-level models for the inter-subject pattern analysis included the same nuisance regressors as the previous analysis. However, each video was represented with its own block regressor that covered the whole duration of the video. This allowed us to examine video specific patterns.

Inter-Subject Analyses

Inter-subject correlation (ISC)

The ISC allowed us to examine how dynamic processing of the second half videos was modulated by previous knowledge. ISCs were computed voxelwise over second half videos (HC and LC videos, which contained coherent speech). To examine the similarity across participants under the same condition we constructed 2 condition lists. To construct 2 condition lists from 4 counterbalancing lists we combined the lists in which videos were seen under the same condition, but in a different order. There were 9 and 10 people in the two condition lists. It has been shown that averaging over at least 4 people's time courses provides reliable ISC estimation (Hasson et al., 2004). The first 2 TRs (5.24 seconds) of each video were removed in order to remove transient onset effects that can lead to artificially high ISC (Ames et al., 2015).

The raw time course for each video and each subject were extracted. These time-courses were used to compute the Fisher-transformed correlations across subjects for each video. For a given subject we computed the correlation between the subject's specific video time-course and the average time-course for the rest of the participants watching the same video in the same condition. This resulted in 20 time-course correlations for each subject, which represented the time-course similarity across participants watching the same videos (diagonal values in Fig. 3 A & C). The time-course correlations across participants for different videos were also computed for the general ISC analysis (see Fig. 3 A).

First, a general ISC analysis allowed us to examine ISCs while watching the same videos irrespective of whether the context was familiar or not (see Fig. 3 A).

We compared the time-course correlations across participants watching the same video versus different videos.

Next, we examined how narrative themes affected ISCs. We contrasted the similarity of a participants' time-course with that of other participants for HC versus LC videos (see Fig. 3 B), in order to examine whether participants were more synchronised when they watched the HC clips versus when they watched the LC clips. This contrast was performed for each subject by averaging over the 10 Fisher transformed correlations across HC videos and subtracting the average of the 10 LC coefficients.

Inter-Subject Pattern Similarity (ISPS)

This analysis was performed to examine spatial pattern similarity across participants experiencing the same events. The inter-subject pattern (ISPS) analysis was conducted on the normalised t-maps generated for each subject and each 2nd half video. For each subject, a searchlight map was generated by centring a spherical searchlight with radius of 3 voxels (mean size 110 voxels) at each voxel. For a given participant the activity patterns for a specific video were correlated with the average activity pattern across participants watching the same video in the same condition. This resulted in 20 correlations for each subject representing the video specific pattern similarity across subjects (see diagonal Fig. 4 A). The across participants pattern similarity was also calculated for non-matching videos resulting in 380 correlation coefficients for each searchlight (see off-diagonal Fig. 4 A). Contrast matrixes were used to weight the resulting Fisher-transformed correlations for each searchlight. The summed correlations were assigned to the central voxel of each searchlight. Two types of ISPS analyses were conducted by specifying different contrast matrixes

First, we conducted a "General Show" ISPS analysis to examined the similarity across participants watching the same clip versus when they were watching non-matching clips, irrespective of their prior knowledge (see Fig. 4).

Then we examined whether participants showed more highly correlated patterns for HC videos than for LC videos (see Fig. 4). Due to stimulus counterbalancing, there were two groups (n = 9 and n = 10) of people who saw the same videos in the same condition. Therefore, for the ISPS we compared the

similarity between a subject's patterns and the average patterns of the other subjects in their group.

Results

Behavioural results

Participants had an overall accuracy of 72.2% for the memory questions, which is high level of performance given the fact that the questions were open ended. Participants responded more accurately to the same questions when in the HC condition (80%) than in the LC condition (64.4%: $t_{17} = 3.39$, $p = 0.003$). The memory questions were specific to the 2nd half videos and did not require information from the 1st half videos. HC videos were also rated as being remembered more vividly ($t_{17} = 7.15$, $p < 0.001$), more coherent ($t_{17} = 6.88$, $p < 0.001$) and more engaging ($t_{17} = 6.23$, $p < 0.001$).

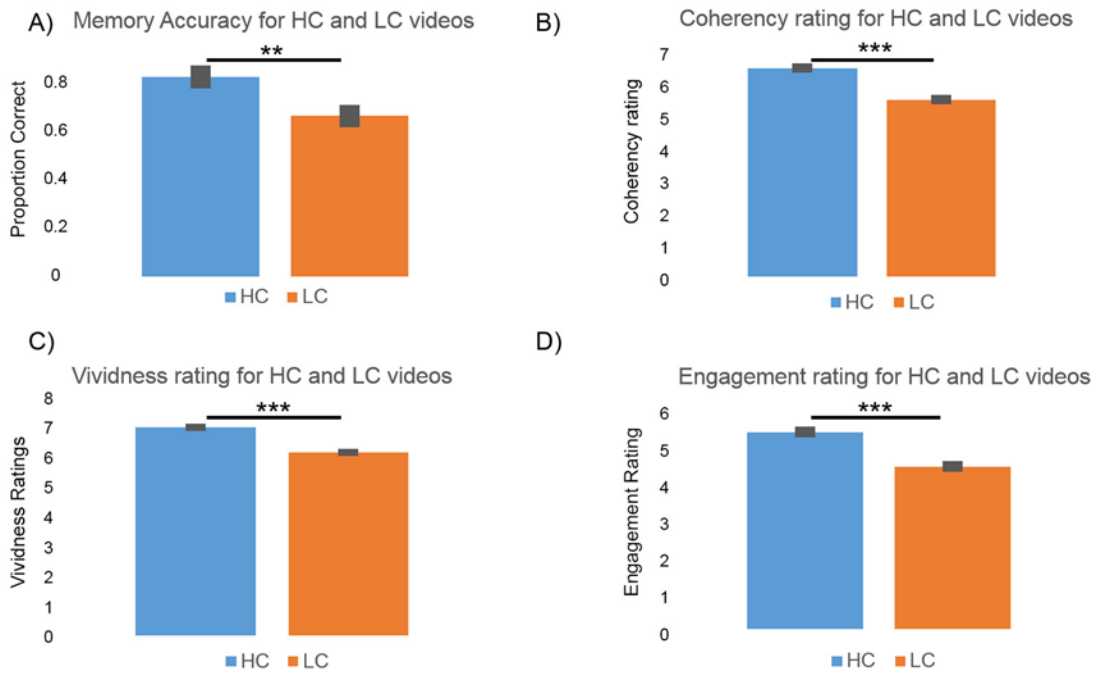

Figure 2. *Behavioural results.* A) Shows average memory performance for questions on HC versus LC videos. B) Shows that participants on average found HC videos more coherent. C) Higher vividness ratings were observed for HC videos. D) Participants reported HC videos as more engaging. ** $p < .01$; *** $p < .001$

fMRI Results

GLM

The contrast of watching videos versus the active baseline task revealed higher activation in visual, auditory, medial and anterior temporal cortices (see Supplementary Fig. 1). This is consistent with previous studies using videos (e.g. Bartels & Zeki, 2004). Brain areas that showed higher activation for the onset of the videos versus baseline are shown in Supplementary Fig. 3. These results replicate Keidel et al.'s (2017) findings and indicate that retrosplenial cortex (RSC) and parahippocampal cortex (PHG) extending into the ventral precuneus showed transient activation associated with the onset of the videos. The results of the time-course analysis was also consistent with this previous study in showing higher activation in the middle temporal gyrus (MTG), supramarginal gyrus (SMG) and angular gyrus (AG) for videos depicting a continuation of a previous narrative (HC clips). See Supplementary Material for further discussion.

Inter-Subject Correlation

The ISC analysis allowed us to examine synchronization of the BOLD response across participants. ISCs for watching videos irrespective of the context manipulation were found in extensive regions of the occipital and superior temporal lobes, encompassing visual and auditory processing regions. Higher synchronization was also observed in bilateral IFG, medial prefrontal cortex, AG and precuneus (see Fig. 3).

The ISC comparison between HC and LC videos showed stronger coupling in the ventromedial portion of the prefrontal cortex as well as the left AG, IFG and superior parietal lobule (SPL) (see Fig. 3). We also observed a cluster in the left inferior parietal lobule (IPL), which showed greater synchronization for LC clips.

Inter-Subject Pattern Similarity

The general ISPS allowed us to examine where spatial patterns of activity were more similar across participants watching the same videos versus watching different videos irrespective of the context manipulation. The results are shown in Fig. 4. This analysis revealed significant video-specific similarity across participants in the primary sensory areas and areas associated with higher cognitive processes,

such as the bilateral IFG, precuneus, the MTG, the inferior portion of the anterior temporal pole (ATL) and the medial prefrontal cortex. The video-specific pattern similarity results reported here strongly resemble the scene-specific cross-subject similarity reported by Chen et al. (2017).

However, our main interest was to examine whether narrative theme modulates this effect. We expected higher spatial pattern similarity for people watching HC videos as they had shared narrative knowledge and were, potentially, better able to interpret the videos. Indeed, for the HC > LC contrast we found that left and right anterior temporal lobe (ATL) to show higher pattern similarity across participants. Surprisingly, the reverse contrast showed that the precentral gyrus exhibited higher pattern similarity across subjects for LC clips.

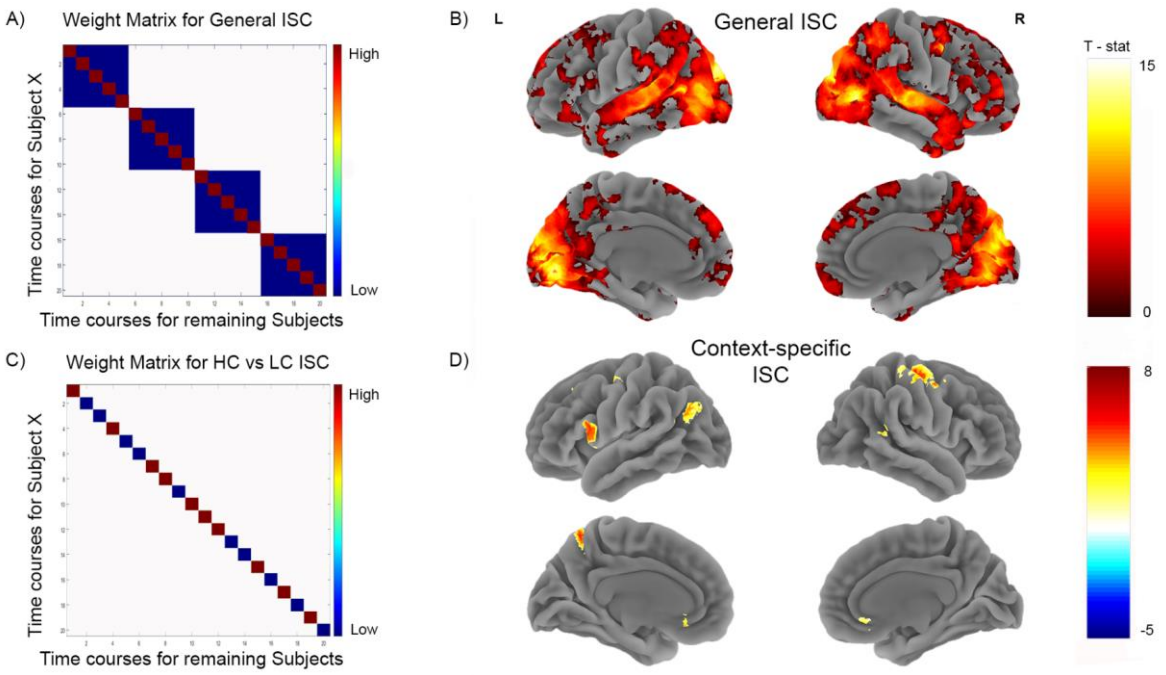

Figure 3 *Inter-subject correlations*. A) The weight matrix (General ISC) tests for video specific time course similarity across participants. Each cell represents the correlation between subjects' time course for a particular video with the average time course of all remaining participants for a particular video. The diagonal represents correlations between time courses for the same videos. B) Brain map from video specific analysis, which shows extended synchronization across the brain for people watching the same videos. C) Weight matrix that tests for the time-course similarity across the same videos, depending on the prior knowledge provided for them. D) Brain map showing how time-course synchronicity was modulated by prior knowledge.

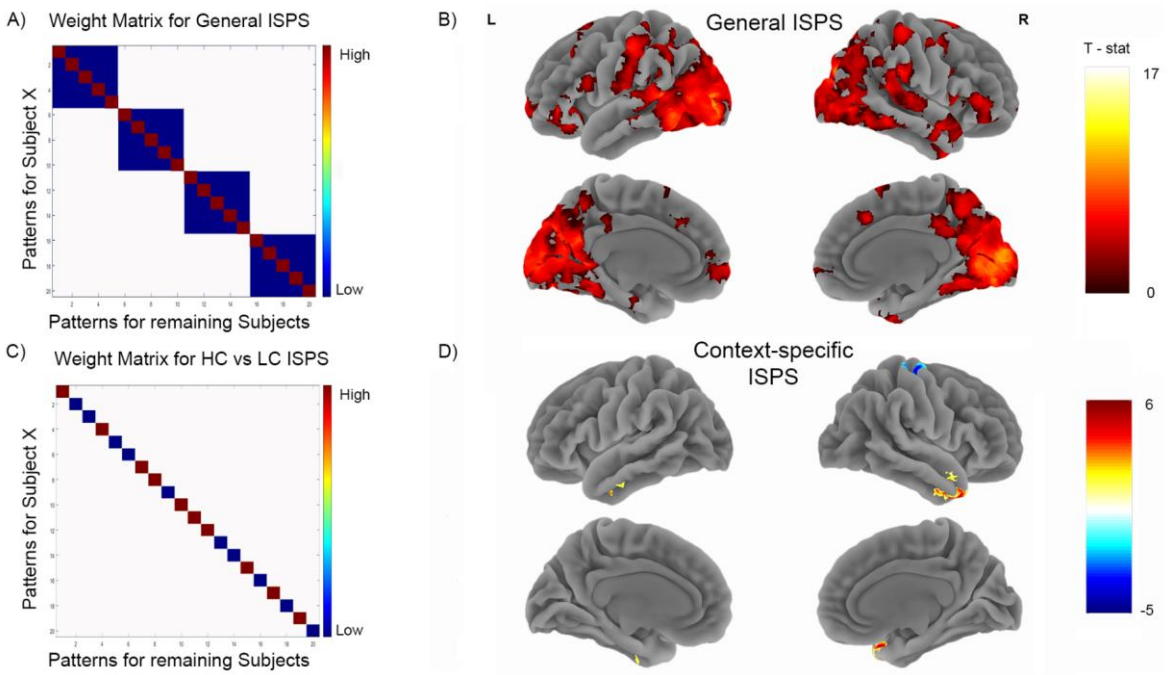

Figure 4 Inter-subject Pattern Similarity. A) The weight matrix (General ISPS) tests for video specific spatial pattern similarity across participants. The diagonal represents spatial similarity for the same videos. B) Brain map from video specific analysis, which shows extended pattern similarity across the brain for people watching the same videos. C) Weight matrix that tests for the spatial similarity across the same videos, depending on the prior knowledge provided for them. D) Brain map showing how spatial pattern similarity was modulated by prior knowledge.

Discussion

This study examined the cognitive and neural effects of prior knowledge of a narrative storyline on the processing of novel events. Participants viewed videos for which they either were, or were not, provided with knowledge of the preceding narrative (HC and LC respectively). Prior knowledge increased participants' memory performance and subjective comprehension and vividness ratings. Several regions showed higher coupling in the moment-to-moment BOLD response across participants in the HC condition, particularly in regions associated with the brain's semantic network (including the left superior and inferior frontal gyrus, left AG, vmPFC and right posterior MTG). In addition, we observed greater cross-participant multivoxel pattern similarity for HC videos in another region strongly associated with semantic knowledge, the bilateral anterior temporal lobes. The results show that increasing the amount of "common ground" with which to interpret a new event increases the degree to which neural responses in regions of the semantic network are shared across participants. More generally these results highlight the role of the semantic network in processing narrative content, a central element of naturalistic events.

Prior knowledge about a particular situation enables subsequent new information to be comprehended more easily and remembered better (Bransford & Johnson, 1972). It also constrains how the new information is interpreted. For example, in one of the second half videos in our experiment (taken from the show "Just Shoot Me") two of the main characters (Jack and Maya) discuss the fact that penguins cannot fly. Jack claims that "penguins once filled the skies", but one day their confidence was shattered and they never flew again. A participant viewing this clip in the HC condition can relate this discussion back to the first half of the video, in which Jack explains how Maya hurt her colleague's confidence and the negative effect this has had on his work. However, because participants viewing this clip in the LC condition were familiar with the general setting (Jack and Maya in an office), but not the specifics of the narrative, they would be likely to have formed more idiosyncratic interpretations of what the scene is really about. Thus, while participants in both conditions shared the general "scaffold" of the event model, only in the HC condition could participants bring the more specific information gained from the first half of the video to their understanding of the second half.

Previous studies using ISC and pattern similarity measures to investigate event processing have generally compared participants watching similar, or the same events to different events (Bird et al., 2015; J. Chen et al., 2016; J. Chen et al., 2017; Oedekoven et al., 2017; St-Laurent et al., 2015). These studies have found widespread effects throughout sensory brain regions and the DMN. We found similar effects when we compared the similarity across participants watching the same show, regardless of prior knowledge. These effects are likely to be driven by the common elements of the same shows, including the location and people present. Indeed, a recent study showed that the neural representation of events in the DMN, indexed by multivoxel patterns of activity, is dominated by the spatial context (Robin et al., 2018). Our findings comparing the HC with the LC condition reveal a much more restricted set of regions are involved in integrating prior knowledge of the narrative. These effects were found predominately within the semantic network. Furthermore, different regions were highlighted with the ISC analysis to the ISPS analysis, which may reflect the different processing roles of these regions.

ISC effects were seen in many regions which have been associated with semantic processing, particularly in “semantic control” (Binder et al., 2009; Noonan et al., 2013). These were the superior and inferior frontal gyri, the AG, posterior MTG, vmPFC and premotor cortex. The fact that ISCs are based on temporal correlations in activity, suggest that these regions are involved in moment-by moment integration of knowledge gained by watching the intact first half of each video (for participants watching in the HC condition) with the incoming information provided by the second half of the video. Of note, the MTG, IFG and AG were recently highlighted in a recent study of narrative context processing (Keidel et al., 2017). Activity in these regions was higher, and in the case of the MTG, peaked earlier, when a narrative context was provided. The authors suggested that these regions processed the “on-the-fly” links between semantic concepts necessary to comprehend the storyline.

The results of our ISC analyses also extend previous findings examining how prior knowledge affects event processing. Ames et al. (2015) and van Kesteren et al. (2010) found higher inter-subject correlation in vmPFC when participants were

processing stimuli for which they had prior knowledge, and suggested that vmPFC activity supports the integration of incoming information with previously established representations (see also Gilboa & Marlatte, 2017; van Kesteren et al., 2012). We have here both replicated this vmPFC finding and additionally showed increased synchronization in several other regions of the semantic network when participants share knowledge of the narrative. Whether these regions process any previously acquired information (such as people or locations), or they more specifically process the abstract concepts necessary to understand a narrative, is a question for future research.

We additionally carried out ISPS analyses to see where the spatial patterns of BOLD activity were more consistent between participants sharing narrative contextual information. These effects were localised to the anterior temporal lobes bilaterally. Although these results are due to the same contrast as the ISC results discussed above (greater similarity for the HC compared to the LC clips), they are likely to be driven by subtly different mechanisms. While ISC effects reflect temporal variations in activity during the clips that are more synchronised across participants, the ISPS effects are driven by spatial variation in the activity patterns for each clip, where the patterns reflect time-averaged responses. Therefore, these effects are likely to represent shared representations of the overarching narrative themes rather than more dynamic processing of concepts relating to the ongoing narrative. This information, together with information about the location and people present, is a core element of the event model describing a particular situation. The importance of the ATL for semantic memory and conceptual knowledge is well-established (L. Chen et al., 2017; Patterson et al., 2007; Rice et al., 2015). Within the ATL, the anterior and ventral regions identified in our study have been argued to act as a semantic “hub”, supporting amodal conceptual representations that are independent of specific sensory input (Murphy et al., 2017; Patterson, et al., 2007). Our results extend this theory, suggesting that the region supports abstract representations of information as complex as narrative themes.

Of course, we would expect multimodal events to be represented across many cortical regions. If we take together these previous findings and the results of the current study, it appears that there is a division of labour in the instantiation of event models in the brain. Regions of the DMN, notably posterior midline regions,

represent the stable core of the model, such as the location and identity of the people present (J. Chen et al., 2017; Robin et al., 2018). By contrast, the ATL may play a central role in representing the overarching narrative gist. Finally, other regions of the semantic network, particularly the vmPFC, IFG and AG, may more tightly track the changing aspects of the associated narrative and link then to prior knowledge. This framework is similar to that proposed by Ranganath and Ritchey (2012).

In summary, we have shown how manipulation of knowledge about a narrative storyline impacts on the cognitive and neural processing of events. The availability of narrative information led to increases in intersubject synchronization and spatial pattern similarity in regions associated with semantic processing. Moment-by-moment tracking of narrative information and linking it to prior knowledge was associated with several regions of the semantic control network. Conversely, global updating of the event model with the central narrative themes was linked to representations supported by the ATLs. These results provide important new insights into how the brain represents and updates narrative information as well as highlights an important case of functional specialisation within the wider network of brain regions that are involved with event processing.

References

Ames, D. L., Honey, C. J., Chow, M. A., Todorov, A., & Hasson, U. (2015). Contextual alignment of cognitive and neural dynamics. *Journal of Cognitive Neuroscience, 27*(4), 655-664. doi: 10.1162/jocn_a_00728

Andersson, J. L. R., Hutton, C., Ashburner, J., Turner, R., & Friston, K. (2001). Modeling Geometric Deformations in EPI Time Series. *NeuroImage, 13*(5), 903-919. doi: 10.1006/nimg.2001.0746

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage, 38*(1), 95-113. doi: 10.1016/j.neuroimage.2007.07.007

Baldassano, C., Hasson, U., & Norman, K. A. (2018). Representation of real-world event schemas during narrative perception. *bioRxiv*, 252718. doi: 10.1101/252718

Bartels, A., & Zeki, S. (2004). Functional brain mapping during free viewing of natural scenes. *Human Brain Mapping, 21*(2), 75-85.

Ben-Yakov, A., Eshel, N., & Dudai, Y. (2013). Hippocampal immediate poststimulus activity in the encoding of consecutive naturalistic episodes. *Journal of Experimental Psychology: General, 142*(4), 1255-1263. doi: 10.1037/a0033558

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex, 19*(12), 2767-2796. doi: 10.1093/cercor/bhp055

Bird, C. M., Keidel, J. L., Ing, L. P., Horner, A. J., & Burgess, N. (2015). Consolidation of Complex Events via Reinstatement in Posterior Cingulate Cortex. *The Journal of Neuroscience, 35*(43), 14426-14434. doi: 10.1523/JNEUROSCI.1774-15.2015

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*(9/10), 341-345. doi: citeulike-article-id:4016976

Bransford, J. D., & Johnson, M. K. (1972). Contextual Prerequisites for Understanding - Some Investigations of Comprehension and Recall. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 717-726. doi: 10.1016/S0022-5371(72)80006-9

Chen, J., Honey, C. J., Simony, E., Arcaro, M. J., Norman, K. A., & Hasson, U. (2016). Accessing Real-Life Episodic Information from Minutes versus Hours Earlier Modulates Hippocampal and High-Order Cortical Dynamics. *Cerebral Cortex, 26*(8), 3428-3441. doi: 10.1093/cercor/bhv155

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience, 20*(1), 115-125. doi: 10.1038/nn.4450

Chen, L., Lambon Ralph, M. A., & Rogers, T. T. (2017). A unified model of human semantic knowledge and its disorders. *Nature Human Behaviour, 1*(3), 0039. doi: 10.1038/s41562-016-0039

Dooling, D. J., & Lachman, R. (1971). Effects of Comprehension on Retention of Prose. *Journal of Experimental Psychology, 88*(2), 216-222. doi: 10.1037/H0030904

Fairhall, S. L., & Caramazza, A. (2013). Brain regions that represent amodal conceptual knowledge. *Journal of Neuroscience, 33*(25), 10552-10558.

Gilboa, A., & Marlatte, H. (2017). Neurobiology of Schemas and Schema-Mediated Memory. *Trends in Cognitive Sciences, 21*(8), 618-631. doi: 10.1016/j.tics.2017.04.013

Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: memory as an integral component of information processing. *Trends in cognitive sciences, 19*(6), 304-313. doi: 10.1016/j.tics.2015.04.006

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science, 303*(5664), 1634-1640. doi: 10.1126/science.1089506

Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., & Turner, R. (2002). Image distortion correction in fMRI: A quantitative evaluation. *NeuroImage, 16*(1), 217-240. doi: 10.1006/nimg.2001.1054

Keidel, J. L., Oedekoven, C. S. H., Tut, A. C., & Bird, C. M. (2017). Multiscale Integration of Contextual Information During a Naturalistic Task. *Cerebral Cortex*, 1-9. doi: 10.1093/cercor/bhx218

Maguire, E. A., Frith, C. D., & Morris, R. G. (1999). The functional neuroanatomy of comprehension and memory: the importance of prior knowledge. *Brain, 122*(10), 1839-1850. doi: 10.1093/brain/122.10.1839

Murphy, C., Rueschemeyer, S.-A., Watson, D., Karapanagiotidis, T., Smallwood, J., & Jefferies, E. (2017). Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality. *NeuroImage, 147*, 19-31. doi: 10.1016/j.neuroimage.2016.11.067

Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (2013). Going beyond inferior prefrontal involvement in semantic control: evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. *Journal of Cognitive Neuroscience, 25*(11), 1824-1850. doi: 10.1162/jocn_a_00442

Oedekoven, C. S. H., Keidel, J. L., Berens, S. C., & Bird, C. M. (2017). Reinstatement of memory representations for lifelike events over the course of a week. *Scientific Reports, 7*(1), 14305. doi: 10.1038/s41598-017-13938-4

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Frontiers in Neuroinformatics, 10*, 27. doi: 10.3389/fninf.2016.00027

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8*(12), 976-987. doi: 10.1038/nrn2277

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences, 98*(2), 676-682.

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience, 13*(10), 713-726. doi: 10.1038/nrn3338

Rice, G. E., Hoffman, P., & Lambon Ralph, M. A. (2015). Graded specialization within and between the anterior temporal lobes. *Ann NY Acad Sci*, 26. doi: 10.1111/nyas.12951

Robin, J., Buchsbaum, B. R., & Moscovitch, M. (2018). The primacy of spatial context in the neural representation of events. *The Journal of Neuroscience*, 1638-1617. doi: 10.1523/JNEUROSCI.1638-17.2018

St-Laurent, M., Abdi, H., & Buchsbaum, B. R. (2015). Distributed Patterns of Reactivation Predict Vividness of Recollection. *Journal of Cognitive Neuroscience, 27*(10), 2000-2018. doi: 10.1162/jocn_a_00839

Stark, C. E. L., & Squire, L. R. (2001). When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences, 98*(22), 12760. doi: 10.1073/pnas.221462998

van Kesteren, M. T., Fernandez, G., Norris, D. G., & Hermans, E. J. (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Sciences, 107*(16), 7550-7555. doi: 10.1073/pnas.0914892107

van Kesteren, M. T., Ruiter, D. J., Fernandez, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences, 35*(4), 211-219. doi: 10.1016/j.tins.2012.02.001

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., . . . Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience, 4*, 651. doi: 10.1038/88486

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: a mind-brain perspective. *Psychological bulletin, 133*(2), 273-293. doi: 10.1037/0033-2909.133.2.273

Zadbood, A., Chen, J., Leong, Y. C., Norman, K. A., & Hasson, U. (2017). How We Transmit Memories to Other Brains: Constructing Shared Neural Representations Via Communication. *Cerebral Cortex, 27*(10), 4988-5000. doi: 10.1093/cercor/bhx202.

Table 1. Inter-Subject Correlation. Significant clusters showing scene-specific cross-participant correlations.

Region	x	y	z	Size (voxels)	T
L Superior Occipital Gyrus	-15	-90	27	16039	15.70
L Lingual gyrus	-6	-81	-3	16039*	15.52
R Middle Occipital Gyrus	42	-69	3	16039*	15.19
R Calcarine Sulcus	18	-60	12	16039*	3.62
L Middle Occipital Gyrus	-42	-78	24	16039*	3.84
L Temporal Middle Gyrus	-42	-63	9	16039*	4.63
L Anterior Cingulate Gyrus	-9	33	18	34	7.03
R Medial Orbitofrontal Cortex	12	60	-18	240	6.96
R Caudate nucleus	12	0	15	114	6.68
L Superior Frontal Gyrus	-15	69	6	74	6.56
L Middle Frontal Gyrus	-30	51	9	74*	5.08
R Putamen	24	15	-12	29	4.89
L Caudate	-9	3	12	36	6.15
L Middle Frontal Gyrus	-27	6	54	39	5.59
L Cerebellum Crus II	-42	-57	-45	44	5.54
L Cerebellum Crus II	-33	-78	-51	44*	4.16
L Inferior Temporal Gyrus	-36	0	-42	37	5.43

Table 2. Significant clusters identified for the context-specific ISC contrasting HC > LC. Clusters showing higher cross-participant coupling for HC vs. LC videos.

Region (ISC)	x	y	z	Size (voxels)	T
L Superior Frontal Gyrus	-24	27	42	36	8.31
L Superior Parietal Lobule	-18	-60	60	40	6.86
L Cerebellum IX	-9	-60	-51	72	6.64
L Cerebellum crus II	-12	-81	-45	24	5.13
L Inferior Frontal Gyrus	-57	18	15	65	5.86
L Angular Gyrus	-48	-69	27	72	5.46
L Medial Orbitofrontal Cortex	-3	27	-12	24	5.18
L Precentral gyrus	-39	-3	54	32	5.00
R Middle Frontal Gyrus	51	-6	51	126	6.49
R Middle Temporal Gyrus	51	-48	9	28	5.25
Negative					
L Angular Gyrus	-33	-57	36	24	-5.43

Table 3. Inter-Subject Pattern Similarity. Significant clusters showing scene-specific cross-participant pattern similarity for second half videos regardless of condition.

Region	x	y	z	Size (voxels)	T
L Middle Occipital Gyrus	-36	-90	0	10826	17.37
L Cuneus	-12	-60	24	10826*	12.49
R Calcarine sulcus	9	-93	9	10826*	12.31
R Middle Occipital Gyrus	30	-72	27	10826*	4.65
R Precuneus	6	-72	33	10826*	3.77
L Cerebellum IX	-6	-57	-51	114	10.22
L Inferior Frontal Gyrus	-39	39	-6	311	8.80
R Middle Temporal Gyrus	54	3	-21	285	7.31
R Inferior Temporal Gyrus	33	3	-42	285*	6.58
R Middle Frontal Gyrus	33	6	45	122	7.10
R Precentral Gyrus	54	-3	48	122*	4.24
R Inferior Frontal Gyrus	57	12	21	403	7.00
R Superior Middle Frontal Gyrus	6	21	42	85	6.91
R Supplementary Motor Area	15	12	66	88	6.65
L Medial Orbitofrontal Cortex	-9	66	-3	200	6.48
L Parahippocampal Gyrus	-30	-12	-27	70	5.72
L Supplementary Motor Area	-12	12	60	111	5.98
L Middle Frontal Gyrus	-36	30	45	111*	5.13
L Inferior Frontal Gyrus	-54	12	21	112	5.68

Table 4. Inter-Subject Pattern Similarity HC vs LC. Significant clusters showing higher pattern similarity across participants for HC vs. LC videos.

Region	x	y	z	Size (voxels)	T
R Medial Temporal Pole	39	15	-39	138	6.66
R Medial Temporal Pole	57	6	-21	138*	6.16
R Inferior Temporal Gyrus	51	-5	-36	138*	5.04
L Inferior Temporal Gyrus	-42	-3	-36	56	5.48
Negative					
R Cerebellum Crus I	48	-51	-39	156	-5.85
R Precentral Gyrus	33	-27	69	88	-5.11