



Published in final edited form as:

Neuroimage. 2019 January 01; 184: 161–170. doi:10.1016/j.neuroimage.2018.09.010.

Shared understanding of narratives is correlated with shared neural responses

Mai Nguyen^{1,*}, Tamara Vanderwal², and Uri Hasson^{1,3}

¹Department of Psychology, Princeton University, Princeton, NJ 08540

²Yale Child Study Center, Yale University, 230 S Frontage Rd, New Haven, CT 06520

³Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540

Abstract

Humans have a striking ability to infer meaning from even the sparsest and most abstract forms of narratives. At the same time, flexibility in the form of a narrative is matched by inherent ambiguity in its interpretation. How does the brain represent subtle, idiosyncratic differences in the interpretation of abstract and ambiguous narratives? In this fMRI study, subjects were scanned either watching a novel 7-min animation depicting a complex narrative through the movement of geometric shapes, or listening to a narration of the animation's social story. Using an intersubject representational similarity analysis that compared interpretation similarity and neural similarity across subjects, we found that the more similar two people's interpretations of the abstract shapes animation were, the more similar were their neural responses in regions of the default mode network (DMN) and fronto-parietal network. Moreover, these shared responses were modality invariant: the shapes movie and the verbal interpretation of the movie elicited shared responses in linguistic areas and a subset of the DMN when subjects shared interpretations. Together, these results suggest a network of highlevel regions that are not only sensitive to subtle individual differences in narrative interpretation during naturalistic conditions, but also resilient to large differences in the modality of the narrative.

Keywords

cross-modal; fMRI; individual differences; inter-subject correlation; ISC; naturalistic

*Corresponding author: Mai Nguyen, mnguyen@princeton.edu, Princeton Neuroscience Institute, Office 244B, Princeton University, Princeton, NJ 08544.

Declarations of interest: none

Data sharing: Data from this study are archived at the NIMH. Code for reproducing analyses are available at <https://github.com/mnguyen>. The movie stimulus, titled "When Heider Met Simmel," is available for research use by contacting the corresponding author.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Human communication is remarkably flexible, allowing the same narrative to be communicated in forms as varied as words, images, and even the motion of simple shapes (Heider and Simmel, 1944). At the same time, however, in daily life, narratives are often ambiguous, necessitating ongoing interpretation. How are interpretations and meanings of complex, ambiguous narratives—across different communicative forms—represented in the brain?

Previous work has shown that narratives elicit correlated neural responses across subjects in regions of the Default Mode Network (DMN), including temporal parietal junction (TPJ), angular gyrus, temporal poles, posterior medial cortex (PMC), and medial prefrontal cortex (mPFC) (Hasson et al., 2008, 2010; Jaaskelainen et al., 2008; Wilson et al., 2008; Lerner et al., 2011; Ben-Yakov et al., 2012; Simony et al., 2016a; Pollick et al., 2018). This shared response is driven by the content and interpretation of the narrative, rather than its form. For example, the same narrative presented in different modalities (Regev et al., 2013; Baldassano et al., 2017; Zadbood et al., 2017) or languages (Honey et al., 2012) elicits similar time-courses of neural activity throughout the DMN despite major differences in low-level physical properties. Moreover, when the interpretation of a narrative is manipulated using contextualizing information, these shared neural responses are greater among subjects who shared interpretations than between subjects with contradictory interpretations in DMN (Yeshurun et al., 2017b).

In past work, interpretation of the narrative was uniform within groups and was unambiguously imposed on the narrative; however, in daily life, narratives can be ambiguous and interpreted in many different ways. To what extent will the spontaneous and unguided interpretation of a complex, ambiguous, and abstract narrative covary with the degree of shared neural responses across subjects? Based on previous work, we predicted that the more similar the interpretation of an ambiguous movie across individuals, the more similar their neural responses in the DMN will be. In addition, we predicted that because the DMN has been shown to represent narrative content independent of modality, we should observe a similar relationship across forms as long as participant interpretation is similar.

To test these predictions, we scanned subjects in fMRI watching a novel abstract, ambiguous seven-minute animated movie that depicts a narrative through the movement of simple geometric shapes. This movie follows the classic work of Heider & Simmel (Heider and Simmel, 1944), but uses a longer, more complex social plot involving multiple characters and abstract scenes that are open to multiple interpretations. While animated shape movies have been extensively used to investigate theory of mind (Heider and Simmel, 1944; Oatley and Yuill, 1985; Berry et al., 1992; Scholl and Tremoulet, 2000), the movie in the present work is unique in its length, number of characters, social relationships, and high-level narrative arc (SI Movie 1). In addition, we scanned a second group of subjects listening to verbal description of the movie's narrative. Immediately following stimulus presentation, all subjects were asked to freely describe the stimulus. Based on textual analysis of these free recalls, we then compared recall similarity with neural similarity using both intersubject correlation analysis (ISC) and intersubject representational similarity analysis (RSA). These

analyses suggest that regions of the DMN enable the interpretation of narrative under complex, naturalistic conditions, and is at once sensitive to subtle, individual differences in narrative interpretation and resilient to vast differences in the form of narrative communication.

2. Methods

2.1 Subjects

Fifty-seven adult subjects with normal hearing and normal or corrected-to-normal vision participated in the experiment. One subject was excluded for falling asleep and two for excessive motion during scanning (>3mm), resulting in 36 subjects (ages 18–35, mean 22.5 years; 19 female) in the Movie group and 18 subjects (ages 18–32, mean 22.2 years; 14 female) in the Audio group. The sample size for the Audio group was based on reliability analyses of sample size required to measure reliable ISC (Pajula and Tohka, 2016). A larger sample size for the Movie group was selected based on previous work using ISC to detect differences among conditions (Cooper et al., 2011; Lahnakoski et al., 2014; Schmälzle et al., 2015; Yeshurun et al., 2017b, 2017a). All experimental procedures were approved by the Princeton University Internal Review Board, and all subjects provided informed, written consent.

2.2 Stimuli and experimental design

Subjects were split into two separate groups. The “Movie” group was scanned using fMRI while watching a 7-min animated film. The movie depicted a short story using moving geometric shapes in the style of Heider & Simmel (Heider and Simmel, 1944). While there was no spoken dialogue, the animation included an original piano score that communicated mood and was congruent with events in the narrative (“Movie,” Fig. 1, top row; SI Movie 1 for full movie).

A second group of subjects was also scanned in fMRI while listening to a 7-min verbal description of the animation’s story, narrating the movement of the shapes as social characters (“Audio story,” Fig. 1, bottom row; SI Section 2 for full audio text). The Audio story was based on the director’s interpretation of the animation, and this condition had no visual component. In the Audio story, a small child lives in a simple house with a parent. On the first night, the child has a dream that he is flying in the sky with a flock of birds. On the second night, the child returns to the dream but the birds change into a monster that tries to chase the child. On the third night, the child is joined by a new friend in the dream, and together they defeat the monster. The next day, the two friends meet in real life at the playground.

Although the Audio story was a concrete narration of the same actions and events as the original Movie animation, there were slight differences in the timing of different events across the visual and auditory conditions. Thus, in order to time-lock the two stimuli, the Movie animation was segmented into 76 short events or actions (e.g. child bounces balls, parent kisses child goodnight, child and friend make a plan). The raw Audio story was edited to match the onset of each event in the animation, following Regev et al. (2013) and

Honey et al. (2012) (Fig. 1). To remove transient, non-selective responses that occur at the onset of a stimulus, all scans were preceded by the same, unrelated 37-second movie clip. This clip was cropped from all analyses.

Immediately following stimulus presentation, while still in the scanner, subjects were asked to freely recall the stimulus using their own words and in as much detail as possible. Recalls were collected during a second functional run using a customized MR-compatible recording system with online sound cancelling. Data from this functional run were not included in analyses here.

2.3 Stimulus presentation

Stimuli were presented using MATLAB (MathWorks) and Psychtoolbox (Kleiner et al., 2007). Video was presented by LCD projector on a rear-projection screen mounted in the back of the scanner bore and was viewed through a mirror mounted to the head coil. Audio was played through MRI-compatible insert earphones (Sensorimetrics, Model S14).

2.4 MRI acquisition

Subjects were scanned in a 3T Magnetom scanner (Prisma, Siemens) located at the Princeton Neuroscience Institute Scully Center for Neuroimaging using a 64-channel head-neck coil (Siemens). In the Audio and Movie scans, volumes were acquired using a T2*-weighted multiband EPI pulse sequence (TR 1500 ms; TE 39 ms; voxel size 2×2×2mm; flip angle 55°; FOV 192×192 mm², multiband acceleration factor 4, no prescan normalization) with whole-brain coverage. Following functional scans, a fieldmap (mean and phase) was collected (dwell time 0.93 ms; TE diff 2.46 ms). Finally, a high-resolution anatomical image was collected using a T1weighted MPRAGE pulse sequence (voxel size 1×1×1 mm).

2.5 Behavioral data analysis

An independent coder blind to the aim of the study segmented the narrative into 22 separate events and then coded each subject's recall for a description of each event. A subject was rated as having recalled an event if they described any part of the event.

Free recalls were then lightly edited to remove non-stimulus related utterances (e.g. "I don't remember," "I'm done," etc.). The edited recalls were then assessed for similarity to each other within and across stimulus groups using Latent Semantic Analysis (LSA), a statistical method for representing the similarity of texts in semantic space (Fig. 2A, left). In brief, LSA derives a semantic space via singular vector decomposition (SVD) on the word frequency matrix of a large corpus of text. Semantic similarity is defined as the cosine distance between (words, phrases, paragraphs or longer) in this space. Semantic similarity measured by LSA has been shown to have human-like performance (Landauer et al., 1998). Here, we used LSA to measure similarity of subject recalls within and across the two stimuli. The semantic space was derived from the Touchstone Applied Science Associates (TASA) college reading-level corpus with 300 factors, as implemented on lsa.colorado.edu. Finally, to order subjects by similarity to each other for visualization purposes, we then conducted agglomerative hierarchical clustering with complete-linkage on the LSA similarity matrices.

2.6 MRI data analysis

2.6.1 Preprocessing.—MRI data were preprocessed using FSL 5.0 (FMRIB, Oxford) including 3D motion correction, fieldmap correction, linear trend removal, high-pass filtering (140 Hz), and spatial smoothing with a Gaussian kernel (FWHM 4 mm). Motion correction was performed using FSL’s MCFLIRT with 6 degrees of freedom, and estimates of both relative, frame-wise movement and absolute movement were extracted. Subjects with excessive head motion (>3mm absolute movement) were removed from the sample. All data was aligned to standard 2-mm MNI space. Following preprocessing, the first 60 TRs were cropped to remove the introductory videos and transitory changes at the start of the stimulus. Voxels with low mean signal (2 std below average) were also removed. These voxels typically were located on edges of the brain or areas with typical signal loss, including fronto-orbital regions and anterior medial temporal regions. On average across subjects, 19% (std = 2.6%) of voxels were removed. Data was z-scored over time. All analyses were conducted in volume space using custom Matlab scripts and then visualized using FSLview.

2.6.2 Audio correlations between stimuli—Because the Movie and Audio were aligned in time such that the start of each event occurs at the same time across, the audio envelopes (audio amplitudes) may be correlated. Following Honey et al. (2012), for between-condition analyses, we thus projected out the audio envelope from each subject’s neural response. The audio envelope for each stimulus was calculated using a Hilbert transform and then down-sampled to the 1.5-second TR using an anti-aliasing, lowpass finite impulse response filter. The resulting envelopes were then convolved with a hemodynamic response function (Glover, 1999). The envelopes were entered into a linear regression model for each voxel in each subject in the corresponding condition. For betweencondition analyses, the BOLD response timecourse was then replaced with residuals of the regression.

2.6.3 Intersubject correlation (ISC) across modalities—To test the hypothesis that subjects who share interpretations of the stimuli show greater neural similarity, we compared the intersubject correlation (ISC) between groups of subjects who differed in how similarly they interpreted the stimulus. ISC was used to measure neural similarity among subjects as it has been extensively demonstrated to capture shared neural responses to naturalistic stimuli across subjects (Hasson et al., 2008; Lerner et al., 2011; BenYakov et al., 2012; Simony et al., 2016).

For the analysis of neural similarity within the Movie group, we split the Movie subjects into two groups based on how similar their interpretations of the Movie were to each other, as measured by the mean LSA similarity of each Movie subject to every other Movie subject. The resulting “high interpretation similarity” group consisted of the 18 Movie subjects with the most similar recalls to each other, while the “low interpretation similarity group” consisted of the 18 Movie subjects with the least similar recalls. We then separately calculated ISC for each sub-group by correlating each movie subject’s response time course to the average of others in the same subgroup (N subjects in subgroup - 1) in the same voxel. The mean of the resulting N correlations is taken as ISC (Hasson et al., 2004; Lerner et al., 2011; Honey et al., 2012; Regev et al., 2013). Finally, to identify areas that showed greater

neural similarity among subjects who agreed on the interpretation compared to subjects with different interpretations, we directly compared ISC between the two sub-groups using a t-test (two-tailed, $\alpha=.05$) in every voxel that had significant ISC, as calculated over all subjects ($N=36$).

We then repeated the same analyses to compare similarity of neural responses across Movie and Audio subjects. For this cross-modal analysis, the Movie subjects were again divided into two groups, this time based on their average similarity to the Audio group. This resulted in a “high interpretation similarity to the Audio story” group and a “low interpretation similarity to the Audio story” group. Thus the more “high similarity” group contained Movie subjects who shared similar interpretation of the narrative with the Audio group, while the “low similarity” group contained Movie subjects who did not.

Statistical significance of ISC was assessed using a permutation test. Each voxel’s time course was phase-scrambled by taking the Fast Fourier Transform of the signal, randomizing the phase of each Fourier component, and then inverting the Fourier transformation. This randomization procedure thus only scrambles the phase of the signal, leaving its power spectrum intact. Using the phase-scrambled surrogate dataset, the ISC was again calculated for all voxels as described above, creating a null distribution of average correlation values for each voxel. This bootstrapping procedure was repeated 1000 times, producing 1000 bootstrapped correlation maps (Regev et al, 2012; Honey et al., 2012).

To correct for multiple comparisons, following Regev et al. (2012), the largest ISC value across the brain for each bootstrap was selected, resulting in a null distribution of the maximum noise correlation and representing the chance level of calculating high correlation values across voxels in each bootstrap. The family-wise error rate of the measured maps was controlled at $q = .05$ by selecting a correlation threshold (R^*) such that only 5% of the null distribution of maximum correlation values exceeded R^* . In other words, only voxels with mean correlation value (R) above the threshold derived from the boot-strapping procedure (R^*) were considered significant after correction for multiple-comparisons.

2.6.5 Intersubject representational similarity analysis (RSA)—To further test these findings using a more fine-grained analysis, we then identified regions of the brain where greater recall similarity between pairs of subjects predicts greater neural similarity by conducting a voxel-wise intersubject representation similarity analysis (RSA) (Kriegeskorte et al., 2006, 2008) between LSA recall similarity and intersubject neural correlations. Intersubject RSA follows the same logic of the classic RSA as described by Kriegeskorte et al. (2008): in RSA, a similarity matrix is built by comparing the similarity of every pair of experimental conditions (e.g. different categories of images). This similarity matrix is then compared with a neural similarity matrix, built by comparing the similarity of patterns of neural activations between each pair of experimental conditions. Larger correlations between the two matrices in a given brain suggest that the brain region represents the information in the experimental condition matrix.

In the present work, rather than building similarity matrices that compare different experimental conditions, we build similarity matrices that compare the extent of similarity in

the interpretation of the movie across different subjects. We first constructed a recall similarity matrix using LSA cosine similarity (see Section 2.5), which represents how similarly each pair of subjects interpreted the stimulus. For each gray matter voxel in the brain, we then constructed a neural similarity matrix by correlating each subject's response timecourse with every other subject's response timecourse in the same voxel. We then calculated Spearman's r between the matrix of neural similarity and the matrix of recall similarity (Fig. 2B). This analysis therefore identifies regions where subjects who have more similar interpretation also have more similar neural responses, suggesting that these areas represent idiosyncratic representations of the stimulus. RSA was conducted both within the Movie group and between the Movie group and Audio group. For the within-group Movie RSA ($n=36$ Movie subjects), the neural and recalls similarity matrices are symmetrical, so only the lower triangles are correlated. For the between-group Movie-Audio RSAs ($n=36$ Movie subjects, $n=18$ Audio subjects), the entire matrix is correlated.

Following Kriegeskorte et al. (2008), statistical significance for RSA was assessed using a permutation test. For each voxel, the rows and columns of the neural similarity matrix were randomly shuffled, and the resulting shuffled matrix was correlated with the LSA similarity matrix as described above. This shuffling procedure was repeated 1000 times, resulting in a null distribution of 1000 values for the null hypothesis that there is no relationship between recall similarity and neural similarity. Following (Chen et al., 2016; Baldassano et al., 2017), the mean and standard deviation of the null distributions were used to fit a normal distribution and calculate p-values. We corrected for multiple comparisons by controlling the False Discovery Rate (FDR) (Benjamini and Hochberg, 1995) of the RSA map using q criterion = 0.05.

2.6.6 ROI analysis—Because our hypotheses were focused on the DMN, we also conducted the ISC and RSA analyses on independently-defined DMN ROIs (Fig. 2C). These ROIs were defined using functional connectivity on previously published, independent data (Chen et al., 2016) where subjects were scanned in fMRI watching a movie. A seed ROI for posterior medial cortex was taken from a resting state-state connectivity atlas (posterior medial cluster functional ROI in “dorsal DMN” set) (Shirer et al., 2012). Following Chen et al. (2016), the DMN ROIs were then defined by correlating the average response in the PMC ROI to every other voxel in the brain during the movie for each of 17 subjects, averaging the resulting connectivity map, and thresholding at $R = 0.5$ (Fig. 1C). Although the DMN is typically defined using resting-state data, recent work has shown that the same network is activated during temporally extended stimuli (Simony et al., 2016).

2.6.7 Control analyses accounting for head motion—To verify that these findings were not due to correlated subject motion, we extracted a timecourse of each subject's estimated framewise displacement and then correlated each pair of subject's motion time course. We correlated each pair of subject's motion correlations with similarity of interpretation, as well as used t-tests to compare the level of correlated motion between high- and low-similarity groups. Finally, we regressed out each subject's motion parameters (3 translation, 3 rotation) from their neural response. The residuals from this regression were then used to repeat the above neural analyses.

3. Results

We compared the behavioral and neural responses within and between two different groups to identify areas of the brain that represent shared understanding of ambiguous narratives over time. The “Movie” group was scanned in fMRI while they watched a novel 7-min ambiguous animation that told a complex social narrative using only the movement of simple geometric shapes. The “Audio” group was scanned while listening to a verbal description of the social interactions in the animation as interpreted by the animator (e.g. the father tucks his son into bed).

3.1 Behavioral results

3.1.1 Variance in shared interpretation across subjects—Subject recalls of the Movie and Audio varied substantially in length and content. In the Movie group, the average spoken recall was 177.6 sec (std = 86.6 sec) and described on average 14.4 events (std = 4.8), as rated by an independent coder. In the Audio group, the average recall was 221.3 sec (std = 68.8 sec) and described an average 18.5 events (std = 3.1). While recall length did not differ between groups ($t(52) = 1.82, p=.073$), Audio subjects recalled more events than Movie subjects ($t(52, 3.19, p=.0024)$).

The LSA (Landauer et al., 1998) results were as expected: we found that there was significantly more variance in recall similarity among the Movie subjects than the Audio subjects ($t(781) = 22.7, p<.001$; mean LSA in Movie group=0.619, std=.125; mean in Audio group=0.852, std=0.045; Fig 3A). For example, Movie subjects differed in their interpretation of a small circle as an inanimate object (e.g. a ball pushed around by the triangle child) or an animate character (e.g. a dog or pet or other animate being). This spread in behavioral outcomes was ecologically derived, meaning it was not prompted. Each subject freely came to their interpretation on his or her own (for examples, Fig. 3B). This enabled us to separate subjects based on differing levels of similarity when interpreting the exact same stimulus.

Based on the average recall similarity to each other, the Movie subjects were split into two equal sized groups (N=18 each group): a “high similarity” group (mean similarity = 0.745, std = 0.071, range = 0.58–0.89) and a “low similarity” group (mean similarity = 0.528, std = 0.123, range = .26-.81). There were no demographic differences between the two groups (age: $t(34)=.11, p>.05$; gender: $\chi^2(1)=1.87, p>.05$; race: $p>.05$ all races, max $\chi^2=.36$).

Between stimuli, the average LSA similarity between the Movie group and the Audio group was 0.545 (std=0.125, range=0.26–0.8). Based on average similarity to the Audio group, the subjects in the Movie group were split into a “high interpretation similarity to Audio” group (mean similarity = 0.604, std = 0.059, range=0.48–0.8) and a “low interpretation similarity to Audio” group (mean similarity = 0.395, std = 0.082, range = 0.26–0.64). There were no demographic differences between the two groups (age: $t(34)=.25, p>.05$; gender: $\chi^2(1)=0, p=1$; race: $p>.05$ all races, max $\chi^2=.36$).

There was some overlap in participants among the groups: 13 of the Movie subjects were in both the “Movie high similarity” group and the “high similarity to Audio” group, while

another 13 subjects were in both the “Movie low similarity” group and the “low similarity to Audio.” The remaining 10 Movie subjects did not overlap.

3.2 Neural results

3.2.1 ISC: greater shared response among subjects with shared interpretations—To test that groups of subjects with more similar interpretations of the stimuli showed greater neural similarity than subjects that had differing interpretations, we measured neural similarity among subjects with similar/dissimilar recalls using ISC, and then tested for differences in ISC between the two groups using t-tests.

ISC within Movie subjects.: The 18 Movie subjects with the most similar recalls to each other showed significantly correlated neural responses throughout early visual areas, including much of occipital cortex, and auditory areas including A1+, superior temporal gyrus (STG), and middle temporal gyrus (MTG). Significant ISC was also observed in high-level regions including bilateral angular gyrus, bilateral supramarginal gyrus (SMG), bilateral inferior frontal gyrus (IFG), PMC, and anterior paracingulate cortex (PCC) ($p < .05$, FWER corrected; Fig. S1, top row). The 18 “low recall similarity” Movie subjects also showed significant ISC in sensory regions, extending into linguistic areas of the superior temporal lobe (Fig. S1, bottom row). However, a t-test contrasting ISC maps between the two groups revealed significantly greater ISC among the “similar” Movie subjects throughout V1+, PMC, right angular gyrus, left SMG, and bilateral superior frontal gyrus (SFG) (Fig. 3A, left). Moreover, the subjects with more similar recalls had significantly greater ISC in all DMN ROIs except mPFC ($p < .05$; Fig 3A, top right) relative to the dissimilar group.

ISC between Movie and Audio subjects.: The Audio subjects and the 18 Movie subjects with the most similar recalls to the Audio subjects showed significantly correlated neural responses in linguistic areas, including posterior superior temporal sulcus (pSTS) and inferior temporal gyrus (ITG). In contrast, the Audio subjects and the 18 Movie subjects with recalls that were the least similar to the Audio subjects only showed significant neural similarity in a small cluster of voxels in right angular gyrus ($p < .05$, FWER corrected; Fig. S1B, bottom row). A t-test revealed significant differences between these two ISC maps in bilateral angular gyrus, PMC, and left MTG (Fig 3B, left). In ROI analyses, the 18 subjects with more similar recalls, across modalities, had greater ISC in right dorsolateral prefrontal cortex (DLPFC) and PMC ($p < .05$; Fig 3B, upper right).

3.2.2 RSA: Interpretation similarity is correlated with neural similarity—To test the hypothesis that greater similarity in the interpretation of a narrative will be reflected in greater neural similarity across subjects, we conducted an intersubject RSA over the entire brain, which directly compares neural similarity with recall similarity in every gray matter voxel in the brain. (Fig. 2B, see Section 2.6.5 above for details).

RSA in Movie group.: We first compared the neural and recall similarity among all subjects in the Movie group ($n=36$). We found that the level of recall similarity was correlated with the level of neural similarity in PMC, right angular gyrus, right SMG, bilateral anterior STG,

bilateral DMPFC, and bilateral DLPFC ($q < .05$, FDR corrected; Fig. 4). The same correlations were also measured in six independently defined ROIs of the DMN (Fig. 5A).

RSA across Movie and Audio groups.: We additionally searched for a relationship between interpretation and neural similarity between all subjects in the Movie group and all subjects in the Audio group. In ROI analyses, we found that interpretation similarity across modalities was significantly correlated with neural similarity across modalities in right DLPFC, PMC, and right angular gyrus ($p < .05$) with a trending correlation in left DLPC ($p = .068$) (Fig 5B). However, in whole-brain analyses, no voxels passed significance testing in the Movie- Audio comparison, although using a lower threshold revealed largely the same voxels as in the Movie, within-group RSA.

3.2.3 Shared responses cannot be explained by correlated motion—The mean frame-wise displacement during the Movie was 0.083 mm (std=0.03) and for the Audio was 0.072mm (std= 0.31), suggesting excellent participant compliance. Motion was not significantly correlated between most subjects (Movie: mean $r = 0.071$, std=0.08; Movie-Audio: mean $r = 0.063$, std=0.086), though a subset of subjects did show correlated motion (Movie: 84 of 630 pairwise comparisons; Movie-Audio: 74 or 648 pairwise comparisons; all $p < .05$, uncorrected for multiple comparisons). However, correlated motion did not vary systematically among subjects as a function of recall similarity: the correlation between motion correlations and recall similarity in the Movie group was -0.049 ($p > .05$) and in the Movie-Audio group was -0.009 ($p > .05$). There was also no difference in the level of correlated motion between the “similar” and “dissimilar” groups used in the ISC analyses (Movie: $t(304) = 0.69$, $p > .05$; Movie-to-Audio: $t(646) = 0.08$, $p > .05$). In addition, we regressed out motion estimates for each subject from their neural response, and repeated the above whole-brain analyses on the residuals of this regression, identifying the same regions as above (Fig. S3).

4. Discussion

Narratives form an important part of daily life, and interpreting unclear or ambiguous narratives is essential to surviving in both the physical and social worlds. Further, humans must be able to interpret complex and dynamic narratives across different communicative forms, including written or spoken word, sign language, and even abstract moving physical forms, as in the present work.

We found that the more similarly two people interpreted the social events depicted in an ambiguous animation, the more similar their neural responses were in a subset of DMN regions, including PMC, right angular gyrus, bilateral STG, and DMPFC. In addition, we observed this relationship in regions outside the DMN, including more anterior right SMG and bilateral DLPFC. Despite vast differences in the physical properties of moving geometric shapes and spoken words, this relationship persisted across modalities: we found that subjects who watched the shapes animation and subjects who listened to the audio story had significantly correlated neural responses only when they shared interpretations of the narrative. Moreover, the more similarly someone who watched the abstract shapes animation interpreted the narrative to someone who listened to the audio version of the story, the more

similar their neural responses in the ROI analysis of PMC, right DLPFC, and right angular gyrus.

This work is the first to identify high-level regions of the brain that can discriminate between idiosyncratic, spontaneous differences in the interpretation of an ambiguous narrative. These results are consistent with previous studies that directly manipulated interpretation by directing attention to different aspects of a spoken narrative (Cooper et al., 2011), changing perspective (Lahnakoski et al., 2014), or biasing with contextual information (Yeshurun et al., 2017b). Unlike these previous results, however, in the present work, we do not manipulate the interpretation of the animation into discrete experimental groups. Rather, we let subjects freely attribute intentions to the motion of simple geometric shapes, which spontaneously led to the creation of complex social narratives in subjects' minds as expressed in their post-viewing descriptions of the animation. We then show that the subtle individual differences in rich narrative interpretation are reflected in individual differences in the neural responses of high-level regions.

The present work also directly demonstrates the modality invariance of these regions: The shared neural representations of a moving shapes animation and a verbal description of the same narrative in left pSTS, left ITG, PMC, and bilateral angular gyrus. This finding extends previous work showing that these regions respond similarly to slightly different pairs of stimuli, for example, individually presented words versus images of the same item (e.g. Chee et al., 2000; Marinkovic et al., 2003; Bruffaerts et al., 2013); spoken versus written sentences, paragraphs or narratives (Spitsyna et al., 2006; Jobard et al., 2007; Lindenberg and Scheef, 2007; Regev et al., 2013), and audio-visual versus spoken narratives (Baldassano et al., 2017; Zadbood et al., 2017). However, the present work is the first to show that despite vast differences in stimulus properties, sparse and abstract stimuli (like triangles and squares) can elicit similar neural responses to explicit verbal storytelling as long as both stimuli induce similar interpretations of the stimuli. This line of work underscores the flexibility of the default mode network, and/or how potent the interpretation of narrative is in the human brain.

4.1 Processing in the default mode network

Many of the regions in which we identify a relationship between interpretation similarity and neural similarity overlap substantially with the DMN (Raichle et al., 2001; Buckner et al., 2008). While the DMN was originally conceptualized as task-negative network showing decreased activity during externally-directed tasks (Raichle et al., 2001; Fox & Raichle, 2007; Buckner et al., 2008), later work observed robust DMN activity to a variety of tasks, including episodic memory (Svoboda et al., 2006; Spreng et al., 2009; Andrews-Hanna et al., 2014), working memory (Vatansever et al., 2015, 2017b), forecasting (Spreng et al., 2009), semantic processing (Krieger-Redwood et al., 2016; Vatansever et al., 2017a), and emotional processing (Barrett and Satpute, 2013). Notably, the DMN largely overlaps with the mentalizing network (Schilbach et al., 2008; Mars et al., 2012), regions of the brain that show increased activity while thinking about other minds. In measuring mentalizing in the brain, one common task contrasts shape animations that have animate motion (e.g. chasing, kicking) and non-animate motion (e.g. random motion). Greater activation was found during

animate than inanimate movies in regions of DMN (Castelli et al., 2002; Vanderwal et al., 2008; Schurz et al., 2014). In contrast to these studies, we use a novel shapes animation that is unique for its length (7 mins vs the typical 10–30 seconds) and complex narrative arc, as well as the large number of interacting characters with different relationships (parent and child, friends, antagonists). We therefore show that these areas not only respond preferentially to animate films, but can discriminate between subtle differences in interpretations of dynamically occurring social interactions.

The DMN is also implicated in the processing of complex narratives. Across many studies, researchers have observed robust correlations among subjects in the DMN in response to the high-level features of narratives. These shared responses across subjects only occur in the DMN with temporally coherent narratives (Hasson et al., 2008; Lerner et al., 2011; Simony et al., 2016), are insensitive to low-level stimulus features such as modality or language (Regev et al., 2013; Honey et al., 2012; Zadbood et al., 2017; Chen et al., 2016), and are locked to narrative interpretation (Yeshurun et al., 2017a, 2017b). Based on these findings, we have previously suggested that the DMN is situated at the top of a timescale processing hierarchy, integrating high-level information over minutes or more (Hasson et al., 2008; Lerner et al., 2011; Hasson & Honey, 2012). Other work has additionally supported a role of the DMN as a high-level network that flexibly integrates information from many lower-level networks (Vatansever et al., 2015; Margulies et al., 2016; Vidaurre et al., 2017). The present findings are therefore consistent with DMN as a high-level global integrator, and additionally demonstrates that the temporal patterns within the DMN are sensitive to idiosyncratic differences in narrative interpretation.

4.2 Processing outside the DMN

We also found a significant relationship between recall similarity and neural similarity with regions outside of the DMN, including right posterior SMG and bilateral DLPFC centered in MFG. Both of these regions are part of the frontoparietal control network (FPCN), which is widely implicated in cognitive control and decision-making processes (Vincent et al., 2008). The FPCN is highly, albeit heterogeneously, interconnected with the DMN (Spreng et al., 2010, 2012). In particular, both SMG and DLPFC are part of a proposed FPCN subnetwork, FPCN-A, that has been shown to be more highly correlated with DMN in a variety of tasks ranging from classic cognitive control tasks, resting state, social-cognitive tasks, and movie watching (Dixon et al., 2018). Moreover, Dixon et al. found in a meta-analysis that FPCN-A is more involved in mentalizing, emotional processing, and complex social reasoning than other regions of the FPCN, which is consistent with a role for SMG and DLPFC in representing idiosyncratic interpretations of an animated shapes movie.

4.3 Limitations and alternate interpretations

In addition to high-level regions, we also found greater neural similarity in visual areas among subjects who shared interpretations than among subjects who had did not. This difference may arise from differences in fixation patterns as a function of interpretation, which would be consistent with a study showing differences in fixation patterns during the same audiovisual movie when subjects take psychological perspectives. These researchers also observed significant differences in neural synchrony in early visual cortex as a result of

different perspectives (Lahnakoski et al., 2014). However, future work should use eye-tracking during scanning in order to test this hypothesis.

Another possible contribution to neural differences among subjects may be due to different but systematic changes in arousal or attention that track with the stimuli in the groups. For example, we observe significant correlations between neural and recall similarity in a subset of fronto-parietal control network regions, raising the possibility that cognitive control or attentional processes may contribute to our findings. To better address this possible confound, subsequent studies could take additional in-scanner physiological measures as well as utilize post-scan questionnaires to measure engagement or emotional arousal.

Finally, it is possible that differences among the high- and low-recall similarity groups may better reflect differences in memory and performance on the recall task, rather than differences in interpretation. Such effects of memory are difficult to control for in this particular experiment, as the ambiguity of the stimulus makes assessing absolute recall performance (rather than recall similarity to others) difficult. All the analyses in this study were performed during the encoding phase, and as such are related to online processing and encoding of the narratives (for further discussion of the tight relationship between memory and online processing see Hasson et al., 2015).

4.4 Intersubject RSA

Finally, this work introduces a novel analytic method, intersubject RSA, for measuring individual differences in neural responses using complex, naturalistic stimuli. This method and type of stimuli can provide important insights into the shared processing of complex social information across subjects that leads to the creation of a shared reality and facilitates social communication (Hasson et al., 2012; Hasson and Frith, 2016). Future applications of this approach could enable us to delineate the development of high-level social cognitive abilities in the DMN during childhood, as well as to understand the development of cross-modal representations in the DMN. This method may also enable the detection of abnormalities during complex naturalistic perception and narrative interpretation relevant to psychotic disorders. For example, previous work has shown that individuals with autism spectrum disorder or schizophrenia show atypical interpretations of more simple shape-based animations (Castelli et al., 2002; Salter et al., 2008; Horan et al., 2009; Bell et al., 2010), but differences in interpretation under naturalistic conditions have not yet been linked to individual differences in neural responses. Finally, the intersubject RSA method may be used to explore neural representations of individual trait differences, such as differences in cultural background, creativity or mentalizing, or political view.

5. Conclusions

In conclusion, this work provides evidence that shared understanding results in shared neural responses within and across forms of communication. The similarity between neural patterns elicited by similar interpretation of the same narrative communicated in different forms (shapes versus words) demonstrates the remarkable modality invariance and strong social nature of the default mode network. This work invokes the role of the default mode network in representing subtle differences in interpretation of complex narratives.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We would like to thank Adele Goldberg, Janice Chen, and Chris Baldassano for helpful advice on the analysis, Shanie Hsieh for behavioral coding and timestamping, and Amy Price for helpful comments on the paper. The animation was created by visual artist Tobias Hoffman and TV, the original score was written by Jodi S. van der Woude, and the audios were recorded by Max Rosmarin.

Funding: This work was supported by NIH Grant 5DP1HD091948–02 (UH), NIH Grant 1R01MH112566–01 (MN), and The Allison Family Foundation (TV).

References

- Andrews-Hanna JR, Saxe R, Yarkoni T (2014) Contributions of episodic retrieval and mentalizing to autobiographical thought: Evidence from functional neuroimaging, resting-state connectivity, and fMRI meta-analyses. *NeuroImage* 91:324–335. [PubMed: 24486981]
- Baldassano C, Chen J, Zadbood A, Pillow JW, Hasson U, Norman KA (2017) Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron* 95:709–721.e5. [PubMed: 28772125]
- Barrett LF, Satpute AB (2013) Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. *Curr Opin Neurobiol* 23:361–372. [PubMed: 23352202]
- Bell MD, Fiszdon JM, Greig TC, Wexler BE (2010) Social attribution test — multiple choice (SATMC) in schizophrenia: Comparison with community sample and relationship to neurocognitive, social cognitive and symptom measures. *Schizophr Res* 122:164–171. [PubMed: 20400268]
- Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol* 57:289–300.
- Ben-Yakov A, Honey CJ, Lerner Y, Hasson U (2012) Loss of reliable temporal structure in event-related averaging of naturalistic stimuli. *NeuroImage* 63:501–506. [PubMed: 22813575]
- Berry DS, Misovich SJ, Kean KJ, Baron RM (1992) Effects of Disruption of Structure and Motion on Perceptions of Social Causality. *Pers Soc Psychol Bull* 18:237–244.
- Buffaerts R, Dupont P, Peeters R, De Deyne S, Storms G, Vandenberghe R (2013) Similarity of fMRI Activity Patterns in Left Perirhinal Cortex Reflects Semantic Similarity between Words. *J Neurosci* 33:18597–18607. [PubMed: 24259581]
- Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The Brain's Default Network: Anatomy, Function, and Relevance to Disease. *Ann N Y Acad Sci* 1124:1–38. [PubMed: 18400922]
- Castelli F, Frith C, Happé F, Frith U (2002) Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain* 125:1839–1849. [PubMed: 12135974]
- Chee MW, Weekes B, Lee KM, Soon CS, Schreiber A, Hoon JJ, Chee M (2000) Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: evidence from fMRI. *NeuroImage* 12:392–403. [PubMed: 10988033]
- Chen J, Leong YC, Honey CJ, Yong CH, Norman KA, Hasson U (2016) Shared memories reveal shared structure in neural activity across individuals. *Nat Neurosci* 20:115–125. [PubMed: 27918531]
- Cooper EA, Hasson U, Small SL (2011) Interpretation-mediated changes in neural activity during language comprehension. *NeuroImage* 55:1314–1323. [PubMed: 21232614]
- Dixon ML, Vega ADL, Mills C, Andrews-Hanna J, Spreng RN, Cole MW, Christoff K (2018) Heterogeneity within the frontoparietal control network and its relationship to the default and dorsal attention networks. *Proc Natl Acad Sci*:201715766.

- Glover GH (1999) Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage* 9:416–429. [PubMed: 10191170]
- Hasson U, Chen J, Honey CJ (2015) Hierarchical process memory: memory as an integral component of information processing. *Trends Cogn Sci* 19:304–313. [PubMed: 25980649]
- Hasson U, Frith CD (2016) Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philos Trans R Soc B Biol Sci* 371:20150366.
- Hasson U, Ghazanfar AA, Galantucci B, Garrod S, Keysers C (2012) Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends Cogn Sci* 16:114–121. [PubMed: 22221820]
- Hasson U, Malach R, Heeger DJ (2010) Reliability of cortical activity during natural stimulation. *Trends Cogn Sci* 14:40. [PubMed: 20004608]
- Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R (2004) Intersubject Synchronization of Cortical Activity During Natural Vision. *Science* 303:1634–1640. [PubMed: 15016991]
- Hasson U, Yang E, Vallines I, Heeger DJ, Rubin N (2008) A Hierarchy of Temporal Receptive Windows in Human Cortex. *J Neurosci* 28:2539–2550. [PubMed: 18322098]
- Heider F, Simmel M (1944) An experimental study of apparent behavior. *Am J Psychol* 57:243–259.
- Honey CJ, Thompson CR, Lerner Y, Hasson U (2012) Not Lost in Translation: Neural Responses Shared Across Languages. *J Neurosci* 32:15277–15283. [PubMed: 23115166]
- Horan WP, Nuechterlein KH, Wynn JK, Lee J, Castelli F, Green MF (2009) Disturbances in the spontaneous attribution of social meaning in schizophrenia. *Psychol Med* 39:635–643. [PubMed: 18606048]
- Jaaskelainen PI, Koskentalo K, Balk HM, Autti T, Kauramaki J, Pomren C, Sams M (2008) Inter-Subject Synchronization of Prefrontal Cortex Hemodynamic Activity During Natural Viewing. *Open Neuroimaging J* 2:14–19.
- Jobard G, Vigneau M, Mazoyer B, Tzourio-Mazoyer N (2007) Impact of modality and linguistic complexity during reading and listening tasks. *NeuroImage* 34:784–800. [PubMed: 17110132]
- Kleiner M, Brainard D, Pelli D, Ingling A, Murray R, Broussard C, others (2007) What's new in Psychtoolbox-3. *Perception* 36:1.
- Krieger-Redwood K, Jefferies E, Karapanagiotidis T, Seymour R, Nunes A, Ang JWA, Majernikova V, Mollo G, Smallwood J (2016) Down but not out in posterior cingulate cortex: Deactivation yet functional coupling with prefrontal cortex during demanding semantic cognition. *NeuroImage* 141:366–377. [PubMed: 27485753]
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868. [PubMed: 16537458]
- Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis – connecting the branches of systems neuroscience. *Front Syst Neurosci* Available at: <http://journal.frontiersin.org/article/10.3389/neuro.06.004.2008/abstract> [Accessed December 14, 2016].
- Lahnakoski JM, Glerean E, Jääskeläinen IP, Hyönä J, Hari R, Sams M, Nummenmaa L (2014) Synchronous brain activity across individuals underlies shared psychological perspectives. *NeuroImage* 100:316–324. [PubMed: 24936687]
- Landauer TK, Foltz PW, Laham D (1998) An introduction to latent semantic analysis. *Discourse Process* 25:259–284.
- Lerner Y, Honey CJ, Silbert LJ, Hasson U (2011) Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J Neurosci Off J Soc Neurosci* 31:2906–2915.
- Lindenberg R, Scheef L (2007) Supramodal language comprehension: Role of the left temporal lobe for listening and reading. *Neuropsychologia* 45:2407–2415. [PubMed: 17451759]
- Margulies DS, Ghosh SS, Goulas A, Falkiewicz M, Huntenburg JM, Langs G, Bezgin G, Eickhoff SB, Castellanos FX, Petrides M, Jefferies E, Smallwood J (2016) Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc Natl Acad Sci* 113:12574–12579. [PubMed: 27791099]
- Marinkovic K, Dhond RP, Dale AM, Glessner M, Carr V, Halgren E (2003) Spatiotemporal Dynamics of Modality-Specific and Supramodal Word Processing. *Neuron* 38:487–497. [PubMed: 12741994]

- Mars RB, Neubert X, Noonan MP, Sallet J, Toni I, Rushworth MFS (2012) On the relationship between the “default mode network” and the “social brain.” *Front Hum Neurosci* 6 Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3380415/> [Accessed May 29, 2016].
- Nichols TE, Holmes AP (2002) Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp* 15:1–25. [PubMed: 11747097]
- Oatley K, Yuill N (1985) Perception of personal and interpersonal action in a cartoon film. *Br J Soc Psychol* 24:115–124.
- Pajula J, Tohka J (2016) How Many Is Enough? Effect of Sample Size in Inter-Subject Correlation Analysis of fMRI. *Comput Intell Neurosci* 2016:1–10.
- Pollick FE, Vicary S, Noble K, Kim N, Jang S, Stevens CJ (2018) Chapter 17 - Exploring collective experience in watching dance through intersubject correlation and functional connectivity of fMRI brain activity In: *Progress in Brain Research* (Christensen JF, Gomila A, eds), pp 373–397 *The Arts and The Brain*. Elsevier Available at: <http://www.sciencedirect.com/science/article/pii/S0079612318300165> [Accessed July 5, 2018].
- Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. *Proc Natl Acad Sci* 98:676–682. [PubMed: 11209064]
- Regev M, Honey CJ, Simony E, Hasson U (2013) Selective and Invariant Neural Responses to Spoken and Written Narratives. *J Neurosci* 33:15978–15988. [PubMed: 24089502]
- Salter G, Seigal A, Claxton M, Lawrence K, Skuse D (2008) Can autistic children read the mind of an animated triangle? *Autism* 12:349–371. [PubMed: 18579644]
- Schilbach L, Eickhoff SB, Rotarska-Jagiela A, Fink GR, Vogeley K (2008) Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the “default system” of the brain. *Conscious Cogn* 17:457–467. [PubMed: 18434197]
- Schmälzle R, Häcker FEK, Honey CJ, Hasson U (2015) Engaged listeners: shared neural processing of powerful political speeches. *Soc Cogn Affect Neurosci* 10:1137–1143. [PubMed: 25653012]
- Scholl BJ, Tremoulet PD (2000) Perceptual causality and animacy. *Trends Cogn Sci* 4:299–309. [PubMed: 10904254]
- Schurz M, Radua J, Aichhorn M, Richlan F, Perner J (2014) Fractionating theory of mind: A metaanalysis of functional brain imaging studies. *Neurosci Biobehav Rev* 42:9–34. [PubMed: 24486722]
- Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD (2012) Decoding Subject-Driven Cognitive States with Whole-Brain Connectivity Patterns. *Cereb Cortex* 22:158–165. [PubMed: 21616982]
- Simony E, Honey CJ, Chen J, Lositsky O, Yeshurun Y, Wiesel A, Hasson U (2016a) Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat Commun* 7:12141. [PubMed: 27424918]
- Simony E, Honey CJ, Chen J, Lositsky O, Yeshurun Y, Wiesel A, Hasson U (2016b) Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat Commun* 7:12141. [PubMed: 27424918]
- Spitsyna G, Warren JE, Scott SK, Turkheimer FE, Wise RJS (2006) Converging Language Streams in the Human Temporal Lobe. *J Neurosci* 26:7328–7336. [PubMed: 16837579]
- Spreng RN, Mar RA, Kim ASN (2009) The Common Neural Basis of Autobiographical Memory, Prospection, Navigation, Theory of Mind, and the Default Mode: A Quantitative Metaanalysis. *J Cogn Neurosci* 21:489–510. [PubMed: 18510452]
- Spreng RN, Sepulcre J, Turner GR, Stevens WD, Schacter DL (2012) Intrinsic Architecture Underlying the Relations among the Default, Dorsal Attention, and Frontoparietal Control Networks of the Human Brain. *J Cogn Neurosci* 25:74–86. [PubMed: 22905821]
- Spreng RN, Stevens WD, Chamberlain JP, Gilmore AW, Schacter DL (2010) Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *NeuroImage* 53:303–317. [PubMed: 20600998]
- Svoboda E, McKinnon MC, Levine B (2006) The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia* 44:2189–2208. [PubMed: 16806314]
- Vanderwal T, Hunyadi E, Grupe DW, Connors CM, Schultz RT (2008) Self, mother and abstract other: An fMRI study of reflective social processing. *NeuroImage* 41:1437–1446. [PubMed: 18486489]

- Vatansever D, Bzdok D, Wang H-T, Mollo G, Sormaz M, Murphy C, Karapanagiotidis T, Smallwood J, Jefferies E (2017a) Varieties of semantic cognition revealed through simultaneous decomposition of intrinsic brain connectivity and behaviour. *NeuroImage* 158:1–11. [PubMed: 28655631]
- Vatansever D, Manktelow AE, Sahakian BJ, Menon DK, Stamatakis EA (2017b) Angular default mode network connectivity across working memory load. *Hum Brain Mapp* 38:41–52. [PubMed: 27489137]
- Vatansever D, Menon DK, Manktelow AE, Sahakian BJ, Stamatakis EA (2015) Default Mode Dynamics for Global Functional Integration. *J Neurosci* 35:15254–15262. [PubMed: 26586814]
- Vidaurre D, Smith SM, Woolrich MW (2017) Brain network dynamics are hierarchically organized in time. *Proc Natl Acad Sci* 114:12827–12832. [PubMed: 29087305]
- Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL (2008) Evidence for a Frontoparietal Control System Revealed by Intrinsic Functional Connectivity. *J Neurophysiol* 100:3328–3342. [PubMed: 18799601]
- Wilson SM, Molnar-Szakacs I, Iacoboni M (2008) Beyond Superior Temporal Cortex: Intersubject Correlations in Narrative Speech Comprehension. *Cereb Cortex* 18:230–242. [PubMed: 17504783]
- Yeshurun Y, Nguyen M, Hasson U (2017a) Amplification of local changes along the timescale processing hierarchy. *Proc Natl Acad Sci* 114:9475–9480. [PubMed: 28811367]
- Yeshurun Y, Swanson S, Simony E, Chen J, Lazaridi C, Honey CJ, Hasson U (2017b) Same story, different story: the neural representation of interpretive frameworks. *Psychol Sci* 28:307–319. [PubMed: 28099068]
- Zadbood A, Chen J, Leong YC, Norman KA, Hasson U (2017) How We Transmit Memories to Other Brains: Constructing Shared Neural Representations Via Communication. *Cereb Cortex*:1–13.

Highlights

- Shared narrative interpretation is correlated with neural similarity in DMN and FCPN
- Neural responses in DMN depend on shared interpretation, not stimulus modality
- Intersubject representation similarity analysis can detect individual differences

A. Experimental design

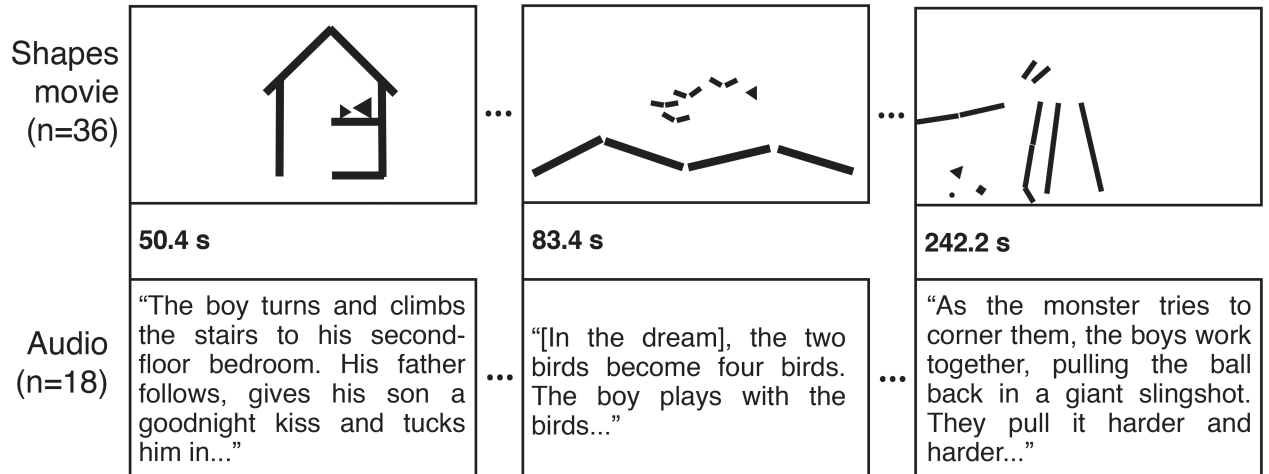


Figure 1. Experimental Design. While being scanned in fMRI, subjects watched a short animation made of moving shapes (Movie) or listened to an audio version of the movie narrative (Audio). The Movie and Audio were time-locked such that each event began at the same time.

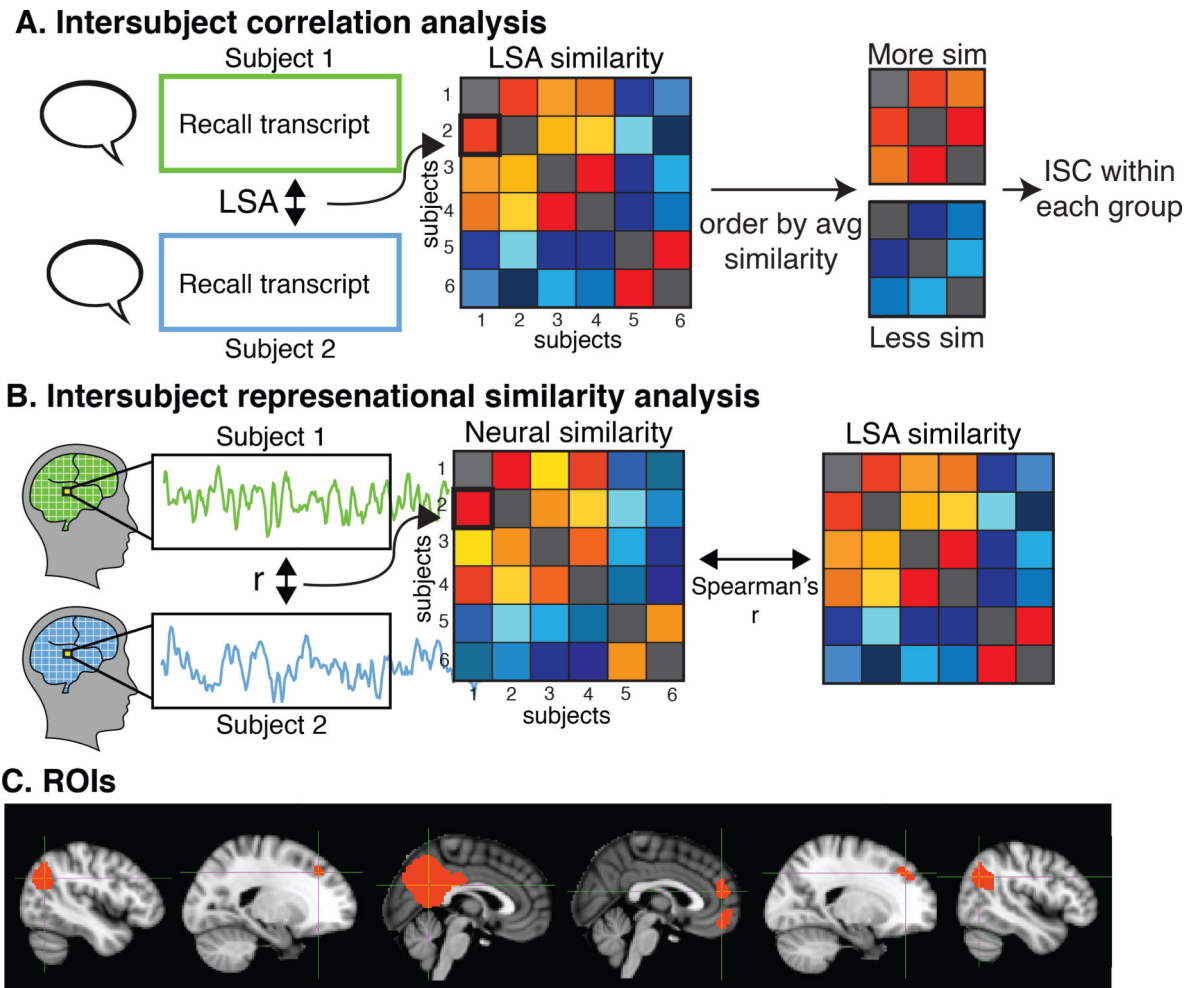


Figure 2. Analysis procedure. **(A)** ISC analysis. Interpretation similarity between every pair of subjects was measured using Latent Semantic Analysis (LSA). Subjects were divided into equal sized groups based on average similarity to each other. Intersubject correlation (ISC) was then compared between subjects who had similar recalls and subjects who did not. **(B)** RSA. An intersubject representational similarity analysis (RSA) was conducted to identify regions of the brain where greater similarity in narrative interpretation was correlated with greater neural similarity. Neural similarity was measured by correlating each subject's response timecourse with every other subject's timecourse in every voxel. The resulting neural similarity matrix was correlated with the recall similarity matrix in all voxels. **(C)** Default mode network ROIs. ISC = intersubject correlation, LSA = latent semantic analysis, RSA = representational similarity analysis, ang = angular gyrus, DMFPC = dorsomedial prefrontal cortex, PMC = posterior medial cortex, MPFC = medial prefrontal cortex.

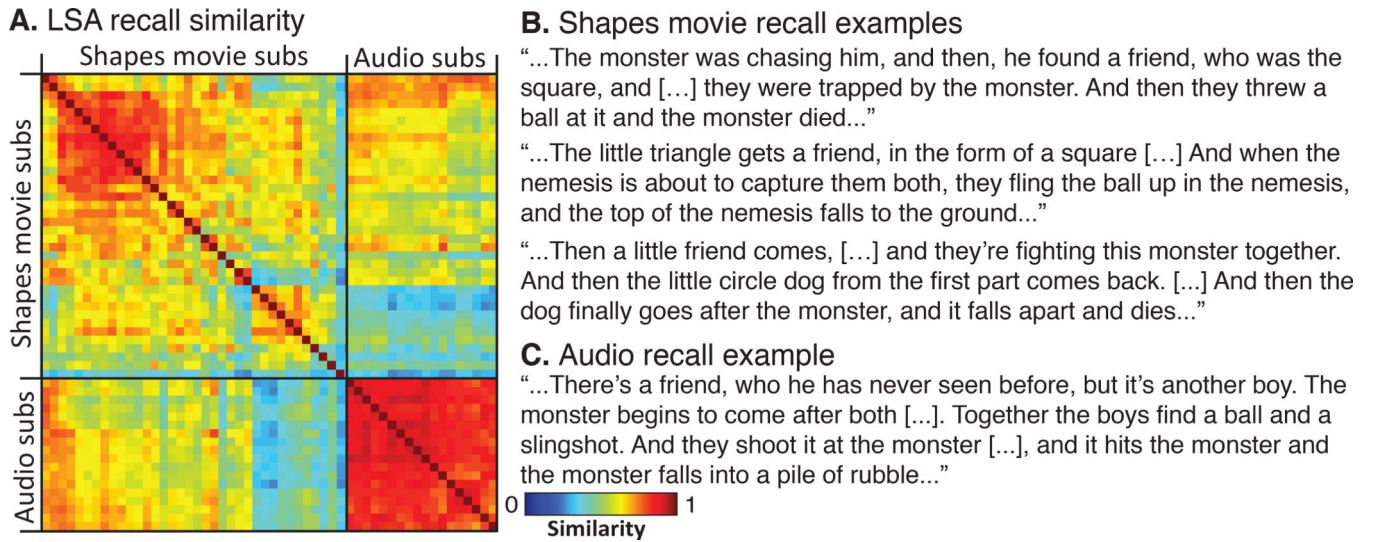
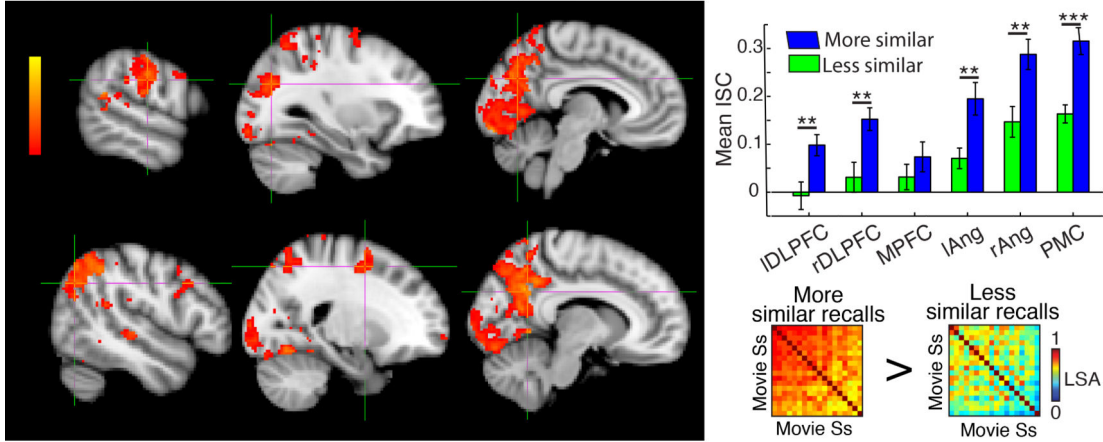


Figure 3. Behavioral results. (A) Recall similarity between each pair of subjects in each group was assessed using Latent Semantic Analysis (LSA). While Movie subjects were varied substantially in their interpretation of the animation, there were far less differences in recall similarity among the Audio subjects ($t=22.7$, $p<.001$). (B) Example Shapes Movie recall excerpts. (C) Example Audio recall excerpt.

A. Within modality: larger ISC among Movie subjects with similar recalls



B. Cross-modality: larger ISC between Movie and Audio subs with similar recalls

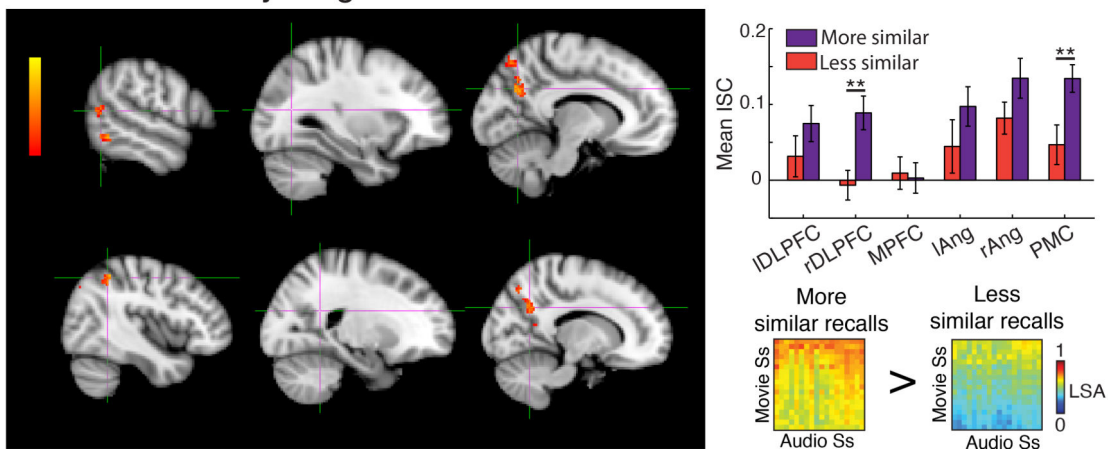


Figure 4. ISC difference. **(A)** Movie subjects with more similar recalls showed significantly greater ISC throughout V1+, PMC, right angular gyrus, left supramarginal gyrus, and bilateral superior frontal gyrus compared to subjects who did not have similar recalls. In addition, the subjects with similar recalls had significantly greater ISC in all DMN ROIs except MPFC. **(B)** Across modalities, Movie and Audio subjects with more similar recalls showed significantly greater ISC in bilateral angular gyrus, PMC, and left middle temporal gyrus. In ROI analyses, subjects, across modalities, with similar recalls had greater ISC in right DMPFC and PMC. DLPFC = dorsolateral prefrontal cortex; MPFC = medial prefrontal cortex; ang = angular gyrus, PMC = posterior medial cortex. * $p < .05$, ** $p < .01$, *** $p < .001$. Error bars are SEM.

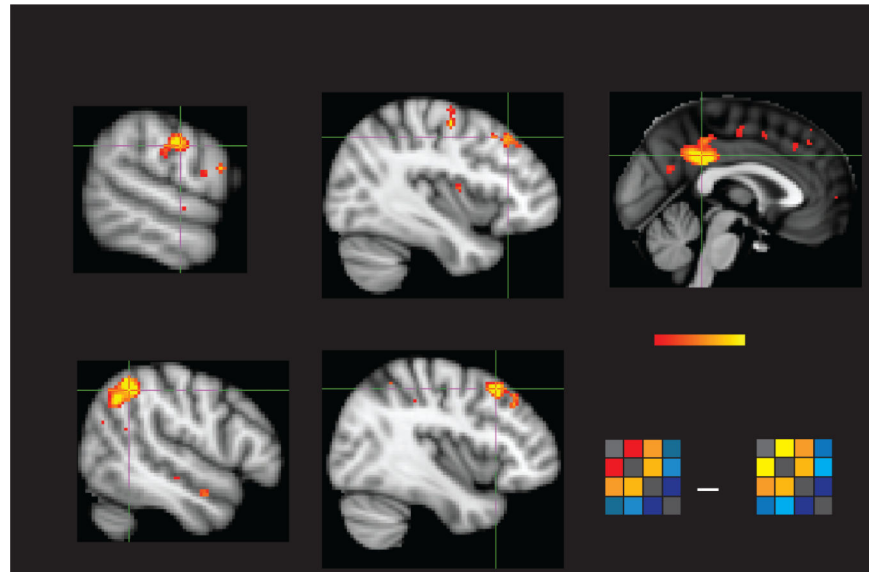
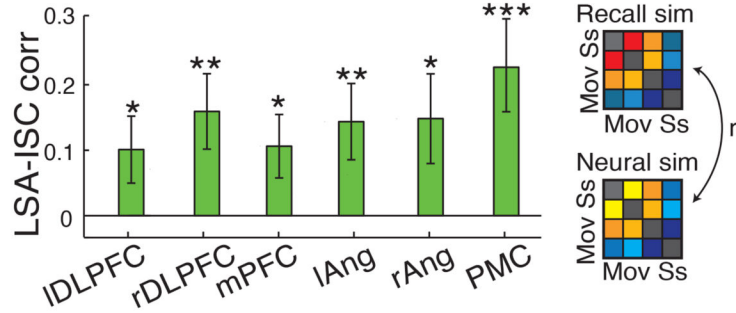


Figure 5. Whole-brain RSA in Movie group. Among subjects who watched the shapes animation, neural similarity and recall similarity were significantly correlated with each other in posterior medial cortex (PMC), bilateral middle frontal gyrus (MFG), right angular gyrus, right superior temporal sulcus (STS), and left inferior parietal lobule (IPL) ($q < .05$, FDR corrected).

A. Within-modality RSA: movie



B. Cross-modal RSA: movie-audio

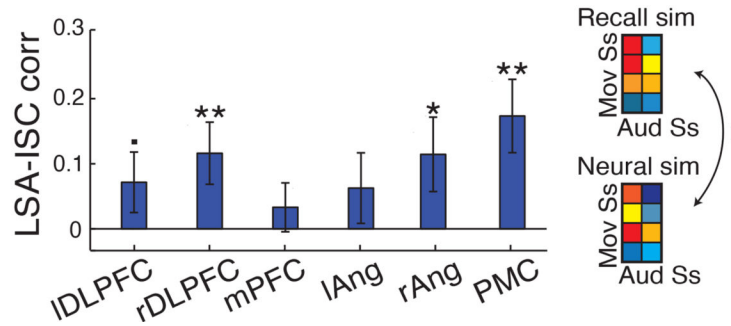


Figure 6. RSA within ROIs. (A) Among Movie subjects, greater recall similarity was significantly correlated with neural similarity in all DMN ROIs. (B) Across Movie and Audio subjects, this relationship was observed only in the PMC and DLPFC ROIs. DLPFC = dorsolateral prefrontal cortex, mPFC = medial prefrontal cortex, ang = angular gyrus, PMC = posterior medial cortex, * $p < .05$, ** $p < .01$, *** $p < .001$. Error bars are SEM.