

**Functional brain network architecture  
supporting the learning of social networks in humans**

Steven H. Tompson<sup>1,2</sup>, Ari E. Kahn<sup>2,3</sup>, Emily B. Falk<sup>4,5,6</sup>,  
Jean M. Vettel<sup>1,2,7</sup>, Danielle S. Bassett<sup>2,8,9,10,11,12</sup>

Author Note

<sup>1</sup>Human Sciences Campaign, U.S. Combat Capabilities Development Center Army Research Laboratory, Aberdeen, MD, 21005 USA

<sup>2</sup>Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, 19104 USA

<sup>3</sup>Department of Neuroscience, University of Pennsylvania, Philadelphia, PA, 19104 USA

<sup>4</sup>Annenberg School of Communication, University of Pennsylvania, Philadelphia, PA 19104 USA

<sup>5</sup>Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104 USA

<sup>6</sup>Marketing Department, Wharton School, University of Pennsylvania, Philadelphia, PA 19104 USA

<sup>7</sup>Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA 93106 USA

<sup>8</sup>Department of Electrical & Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104 USA

<sup>9</sup>Department of Neurology, University of Pennsylvania, Philadelphia, PA 19104 USA

<sup>10</sup>Department of Physics & Astronomy, University of Pennsylvania, Philadelphia, PA 19104 USA

<sup>11</sup>Department of Psychiatry, University of Pennsylvania, Philadelphia, PA 19104 USA

<sup>12</sup>Santa Fe Institute, Santa Fe, NM 87501 USA

Correspondence concerning this article should be addressed to Danielle S. Bassett, Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104

Contact: [dsb@seas.upenn.edu](mailto:dsb@seas.upenn.edu)

## **Abstract**

Most humans have the good fortune to live their lives embedded in richly structured social groups. Yet, it remains unclear how humans acquire knowledge about these social structures to successfully navigate social relationships. Here we address this knowledge gap with an interdisciplinary neuroimaging study drawing on recent advances in network science and statistical learning. Specifically, we collected BOLD MRI data while participants learned the community structure of both social and non-social networks, in order to examine whether the learning of these two types of networks was differentially associated with functional brain network topology. We found that participants learned the community structure of the networks, as evidenced by a slower reaction time when a trial moved between communities than when a trial moved within a community. Learning the community structure of social networks was also characterized by significantly greater functional connectivity of the hippocampus and temporoparietal junction when transitioning between communities than when transitioning within a community. Furthermore, temporoparietal regions of the default mode were more strongly connected to hippocampus, somatomotor, and visual regions for social networks than for non-social networks. Collectively, our results identify neurophysiological underpinnings of social versus non-social network learning, extending our knowledge about the impact of social context on learning processes. More broadly, this work offers an empirical approach to study the learning of social network structures, which could be fruitfully extended to other participant populations, various graph architectures, and a diversity of social contexts in future studies.

Keywords: social network learning, statistical learning, social cognition, functional brain networks

## 1. Introduction

A defining feature of modern social life is the agency that individuals have in constructing their own social environment (Rainie and Wellman, 2012). People may change jobs every few years (Labor, 2017), attend college far from home (Niche, 2014; Strayer, 2016), move cities for work and relationships (Bureau, 2018), and use social media to join clubs and interest groups (Lenhart et al., 2015). In order to navigate these novel social environments, individuals must learn about the complex web of social relationships and cliques that characterize each new social context. For example, integration into friendship networks in the first year of college predicts future success (Stadtfeld et al., 2019) and more diverse social networks in immigrants predict better psychological well-being and cultural adjustment (Repke and Benet-Martínez, 2018). Understanding how people learn relational information about social networks, including which individuals each person is friends with and which communities each person belongs to, may provide key insights into how individuals adapt to novel social contexts.

Efforts to probe the learning of social relations and their complex architectures have traditionally been stymied by the lack of a formal approach for the study of network learning in general. Yet recent advances have met this challenge by conceptualizing relational information as a network in which nodes represent objects or concepts, and in which edges represent shared content or conditional probabilities (Halford et al., 2010; Karuza et al., 2016). Using tasks that are constructed based on this formal graphical conceptualization, evidence suggests that people learning non-social information are sensitive to the individual edges in a relational network of that non-social information. Perhaps even more strikingly, humans learners are also sensitive to the network's meso-scale structure, which can take on many forms including compositions of clusters or communities (Kahn, Karuza, Vettel, & Bassett, 2018; Schapiro et al., 2013). Importantly, some evidence suggests that such meso-scale structure facilitates more efficient information processing (Fine, Jaeger, Farmer, & Qian, 2013; Karuza, Farmer, Smith, Fine, & Jaeger, 2014; Turk-Browne, Scholl, Johnson, & Chun, 2010). For example, the degree to which words are clustered together into communities is associated with how easily a particular word is learned (Goldstein & Vitevitch, 2014). Moreover, individuals that perform a basic perceptual learning task tend to process stimuli more slowly if the stimuli lie in different communities (Karuza et al., 2017; Schapiro et al., 2013). Yet, while there is extensive literature on how people learn relational information and update mental representations of language (Goldstein & Vitevitch, 2014; Karuza et al., 2013; Lieberman, Chang, Chiao, Bookheimer, & Knowlton, 2004), motor sequences (Hung et al., 2018; Kahn et al., 2018; Robertson, 2007), and temporal associations of visual patterns (Fiser & Aslin, 2002, 2005; Garvert, Dolan, & Behrens, 2017; Karuza et al., 2017; Schapiro et al., 2013), little is known about the cognitive and neural processes supporting the learning of relational information in social networks.

To address this gap, we apply graph theory to directly compare functional brain network architecture supporting learning social network structure versus non-social network structure. We designed a novel social network learning task in which social images comprised the nodes of a modular graph, and in which relations among social images are encoded as edges linking the nodes (Tompson et al., 2018). During the task, participants were exposed to a continuous stream of social images defined by a random walk on the graph, thereby holding the transition probabilities implicit in the graph constant (Fiser and Aslin, 2005; Turk-Browne et al., 2005). To evaluate the specificity of our results, we compared our main findings to those obtained during the learning of relational information in non-social networks. For the social portion of the experiment, participants were told

that each image in a set of visual stimuli represented a person, whereas for the non-social control portion of the experiment, participants were told that each image in an equivalent set of visual stimuli represented a rock formation. Importantly, images for the social and non-social task were randomly assigned to each condition for each participant, and thus the only difference between the two tasks was the meaning ascribed to the stimuli.

The task was performed in an MRI scanner during continuous acquisition of BOLD signal. To ensure sensitivity to distributed processing and coordination across brain regions and systems, we capitalized on recently developed tools in the field of network neuroscience (Bassett and Sporns, 2017) to characterize the brain networks supporting the learning of relational information about social networks. Specifically, we identified cortical and subcortical areas that displayed strong connectivity to the rest of the brain; in the parlance of network science, these areas are referred to as hubs, and are thought to be central to coordinating communication between different brain systems and to integrating information from different systems (van den Heuvel and Sporns, 2013). We used whole-brain psychophysiological interaction (WB-PPI; Gerchen et al., 2014; see Figure 1) to assess changes in brain networks during the task and to identify the hubs and cognitive systems that are involved in social network learning versus non-social network learning. Generally, we hypothesized that the manner in which people learn relational information about social networks, including which communities individuals belong to, should share some overlapping processing and mechanisms with how people learn information about other types of statistical relationships. Specifically, we hypothesized that brain regions involved in general memory processes such as the hippocampus (Eichenbaum, 2000; Kernbach et al., 2018) should operate as hubs supporting both social and non-social network learning. The reasoning behind our hypothesis is that the hippocampus is important for encoding relational knowledge in a variety of domains (Garvert et al., 2017; Gheysen et al., 2010; Javadi et al., 2017; Long et al., 2018; Schapiro et al., 2016, 2014). We also hypothesized that individual differences in hippocampal connectivity would be associated with individual differences in learning for both the social task and the non-social control task.

Notably, the network of hubs supporting social network learning may reconfigure relative to non-social network learning, in which case one would expect that social brain areas would operate as hubs for the social condition but not for the non-social control condition. In line with this reasoning, we hypothesized that brain regions involved in social processing such as the amygdala, dorsal medial prefrontal cortex (dmPFC), and temporoparietal junction (TPJ; Adolphs, 2009; Lieberman, 2007; Tso et al., 2018) should operate as hubs preferentially supporting social network learning. Our reasoning is based on the fact that temporoparietal regions including TPJ operate as hubs supporting social perception (Lahnakoski et al., 2012) and integrate sensory information with information about the broader social context to support decision-making (Carter et al., 2012; Carter and Huettel, 2013). We further expected that the strength of the connections between hubs would differ for social network learning and non-social network learning, such that hippocampus is more strongly connected to social hubs during the social network learning task, and more strongly connected to non-social hubs during the non-social network learning task. Finally, we hypothesized that individual differences in network learning would be associated with individual differences in functional brain network architecture, and that this relationship would be moderated by social context. Individuals who recruit social brain regions more should display enhanced performance in the social network learning task but no change in performance in the non-social control task.

## 2. Methods

### 2.1 Ethics statement

Thirty-two participants (12 male and 20 female) were recruited from the general population of Philadelphia, Pennsylvania USA. All participants gave written informed consent and the Institutional Review Board at the University of Pennsylvania approved the procedure.

### 2.2. Description of participants

All participants were between the ages of 18 and 65 years ( $M=25.20$ ,  $SD=9.66$ ), were right-handed, and met standard MRI safety criteria. Two participants opted out of the study during the scanning session due to claustrophobia, two additional participants were excluded due to an average accuracy below 50%, and two participants were excluded due to artifact that resulted in large signal dropout in the parietal lobe.

### 2.3 Procedure

While their brain activation was measured in an MRI scanner, participants completed a social network learning task and a non-social control task. The structure of the experiment was identical to Study 5 in Tompson *et al.* (2018). The order of the two tasks was counterbalanced across participants. In each task, participants viewed a sequence of fractal images that we created using the Qbist filter in the GNU Image Manipulation program (v.2.8.14; [www.gimp.org](http://www.gimp.org)), converted to grayscale, and then matched for average brightness. Images were presented for 1500 ms. To ensure that participants were attending to the stream of images, they were instructed to press the J key with their right index finger if the image was rotated (30% of trials) and to press the F key with their left index finger if the image was not rotated (70% of trials). Each task was broken into 5 runs and participants were given a break between runs to reduce fatigue.

Each image was unique, and for each participant, each image was randomly assigned to a network node in either the social or non-social condition. The sequence of fractal images that each participant saw for each task was generated by a random walk through the network (see Figure 1A). This random walk ensured that the probability of one image being presented after the current trial was equivalent across trials and determined by the network structure. Each node was connected to exactly four other nodes, ensuring that all transition probabilities were equivalent. The structure of transition probabilities is an important cue that signals event structure, which can influence how quickly participants learn information (Fiser and Aslin, 2005; Saffran *et al.*, 1996; Turk-Browne *et al.*, 2005). Therefore, keeping the network structure uniform to remove transition probabilities as a potential source of information about which trials to expect next is important for testing whether participants can learn higher-order network topology.

Participants completed a brief training procedure prior to starting each task. First, they were shown each image in its non-rotated orientation. Then, they were shown the rotated and non-rotated versions side by side and asked to pick the non-rotated image. Next, they completed a practice version of the rotation detection task, where they saw each image once in random order. During each task, participants were also given audio feedback to assist them in learning the rotation of images. Specifically, they heard a high tone when they made an incorrect response and a low tone when they responded too slowly (greater than 1500 ms).

The network structure for each task consisted of two clusters each composed of five nodes, and participants viewed a sequence of 1000 fractal images in each task. For the social task, participants were told that “the images that you will see are taken from an online social media platform where people can choose one of these images as their avatar to represent themselves, much like you might use a photo to represent yourself on Facebook or Twitter. While completing the task (described in more detail on the next page), please make sure you focus on the people these avatars represent.” In the non-social condition, participants were told that the “images were abstract patterns frequently found in rock formations. Some of these patterns are visible to the naked eye, whereas others are only visible with a microscope. These rock patterns are often created by natural forces, including tectonic plate shifts, wind and water erosion, and volcanic activity. While completing the task (described in more detail on the next page), please make sure you focus on the patterns in the rock formations.”

## 2.4 Cross-cluster surprisal

In order to estimate how well participants learned each network, we examined the cross-cluster surprisal, which is measured as the difference in RT between *transition trials* that occurred immediately after a transition from one cluster to another and *non-transition trials* where the image is part of the same cluster as the previous trial (Karuza et al., 2017). If participants learn the cluster membership, then they should anticipate seeing a within-cluster image rather than an image from another cluster (Kahn et al., 2018; Karuza et al., 2019, 2017; Tompson et al., 2018). This surprisal effect should slow participants’ response to the rotation judgment on the next trial (Karuza et al., 2017; Schapiro et al., 2013).

To examine group-level effects of the community structure on participants’ responses, we used linear mixed effects models implemented with the *lmer()* function (library lme4, v. 1.1-10) in R (v. 3.2.2; R Development Core Team, 2015). The primary mixed effects model included node type (transition versus non-transition), network type (social versus non-social), order (social network first versus non-social network first), trial number (standardized), and the two-way and three-way interactions between these variables, as predictors of log-transformed RT; node type, network type, and trial number were included as within-subjects variables, and order was included as a between-subjects variable). For all models, all predictors were mean-centered and we included the fullest set of random effects that allowed the model to converge, which included a random intercept for participant and a by-participant random slope for trial number, network type, and node type. We then conducted simple effects analyses to examine whether the effect of node type was significant in both the social and non-social task conditions. We excluded incorrect trials (15.1% data loss), rotated trials (22.5% data loss), trials with RT less than 100 ms (less than 1% of trials), and outliers greater than 3 SD from the mean RT for each condition (less than 1% of trials). We also excluded a small number of trials (1.6% of trials) where the random walk transitioned from one cluster to another and then immediately transitioned back to the first cluster, which resulted in the middle trial counting as both a pre-transition and post-transition trial.

## 2.5 fMRI acquisition

Data were acquired on a 3T Siemens Prisma scanner with a 64-channel head/neck array. Functional images were acquired using a T2\*-weighted image sequence with a repetition time (TR) of 1000ms, an echo time (TE) of 32ms, a flip angle of 60°, and a 20cm FOV consisting of 56

with 2.5mm thickness acquired at a negative 30° tilt to the AC-PC axis, with a 2.5mm isotropic voxel size, and a multiband factor of four. We also acquired a high-resolution structural image using a T1-weighted axial MPRAGE sequence yielding 160 slices with a 0.9 by 0.9 by 1.0mm voxel size.

## 2.6 fMRI data preprocessing

Neuroimaging data was preprocessed using nipyne (Gorgolewski et al., 2017) and Nilearn (Abraham et al., 2014) implemented in python 2.7 using a combination of AFNI(Cox, 1996), FSL (Jenkinson et al., 2012; Smith et al., 2004), and ANTs (Avants, B. B., Tustison, N., & Song, 2009). The functional data underwent de-spiking to smooth outliers in each voxel using AFNI's 3dDespike, rigid transformation to correct for head motion using FSL's MCFLIRT (Jenkinson et al., 2002), and slice-time correction using FSL's Slicetimer to control for temporal differences in the order of the acquisition of the slices in each brain volume. Skull stripping was performed on the structural data using FSL's BET. The mean functional image was computed and bias-corrected, and then the skull-stripped structural image was bias-corrected. Advanced Normalization Tools (ANTs) was used to compute the transformation parameters for the mean functional image to the high resolution structural image. ANTs segmentation was performed to obtain a warped structural image, a skull-stripped brain mask, and masks for white matter and cerebrospinal fluid. Confound regression was then conducted. The time series was detrended by regressing the time series on the mean and polynomial trends up to quadratic terms. AFNI's 3dbandpass was used to filter out very high or very low fluctuations in the signal (with a high pass of 0.01 and a low pass of 0.12). We included 36 regressors in the confound regression, with six head-motion regressors, three physiological signal regressors (global signal, white matter, and cerebrospinal fluid), as well as their derivatives, quadratics, and squared derivatives. ANTs was used to warp the high-resolution structural image to the MNI template. The transformation parameters from the ANTs functional to structural co-registration and the transformation parameters from the ANTs structural to MNI co-registration were used to warp the 4D functional image to the MNI template. Finally, high variance compounds were removed (Behzadi et al., 2007) using Nilearn. The current preprocessing stream was chosen based on studies that evaluated the performance of a wide variety of preprocessing pipelines in mitigating motion artifact in studies of BOLD functional connectivity (Ciric et al., 2017; Lydon-Staley et al., 2018).

## 2.7 Functional connectivity

In order to examine the functional brain network architecture of network learning, we used the local-global Schaefer cortical parcellation(Schaefer et al., 2018) that divides the human cerebral cortex into 400 functionally homogenous regions (Figure 1B). Given our interest in memory, learning, and social processes, we also added 10 subcortical regions in the left and right hippocampus, left and right amygdala, left and right ventral striatum, left and right caudate, and left and right thalamus using the Harvard-Oxford subcortical atlas (Frazier et al., 2005).

To assess functional connectivity between the regions identified above, we first extracted the average timeseries of activation in each region and standardized the timecourse in each region using the Nilearn (Abraham et al., 2014) package in Python 2.7. We then computed the task-dependent connectivity between each pair of regions using a whole-brain psychophysiological interaction (WB-PPI) approach (Gerchen et al., 2014). For each pair of regions  $i$  and  $j$ , we

computed a multiple regression with the timeseries of activation in region  $i$  as the dependent variable and the timeseries of activation in region  $j$  as an independent variable; we also included a boxcar function to represent the timeseries of each task (coded as 1 during transition trials and 0 during non-transition trials), and we used a separate interaction term to represent the interaction between activation in region  $j$  and the task timeseries. In order to compute the interaction term, the timeseries in region  $j$  was first deconvolved from the canonical HRF function and then multiplied by the boxcar function. The boxcar function and interaction term were then re-convolved with the canonical HRF before computing the multiple regression model. Six head motion parameters and a constant term were also included in the model. These processes were implemented in Python 2.7 using a combination of functions from the `nilearn`, `nipy`, and `nistats` packages, and were designed to follow as closely as possible the implementation of generalized PPI in SPM (McLaren et al., 2012). We constructed a  $410 \times 410$  functional connectivity matrix where the  $ij^{\text{th}}$  element of the matrix represented the task-dependent connectivity (beta weight for the interaction term) between region  $i$  and region  $j$ . We then symmetrized the matrices for each run by averaging the upper and lower triangles, and then we averaged the functional connectivity matrices for the five runs for each task for each subject to yield two functional connectivity matrices for each subject (one for the social task and one for the non-social task).

Given that the transition trials were determined by a random walk through the graph, there were a varying number of transition trials across runs and subjects. In order to account for this variation, we adopted a nonparametric permutation test of significance. We constructed 500 null model networks for each subject by shuffling the trial labels and rerunning the PPI analyses described above, keeping the number of transition trials in each run consistent. The true PPI beta weight for each edge (or other summary statistic) can be compared to the distribution of null model values to derive a  $p$ -value or  $z$ -score. Higher positive  $z$ -scores represent stronger functional connectivity in transition trials versus non-transition trials than would be expected given a random trial order.

## 2.8 Hub analysis

To identify which brain regions support network learning, we used node strength to identify hubs. Node strength is defined as the sum of all the connection weights to an ROI. Node strength was measured by averaging the subject functional connectivity matrices and then computing the sum of the connectivity of each node with all other nodes. We computed the node strength for the combined social and non-social brain network (averaging across the social and non-social tasks) and the social versus non-social brain network (subtracting the average non-social connectivity matrix from the average social connectivity matrix). We repeated this procedure for our 500 null model networks, and  $z$ -scored the node strength statistics. We identified hubs that had significantly greater node strength (FDR-corrected  $p < 0.05$ ) for transition trials than for non-transition trials across both the social network learning condition and the non-social control condition (domain-general hubs), hubs that had significantly greater node strength for the social network condition than for the non-social control condition (social hubs), and hubs that had significantly greater node strength for the non-social control condition than for the social network condition (non-social hubs).

After identifying hubs that were recruited during network learning, we next sought to determine which other brain regions were most strongly connected to those hubs. We began by computing the average strength of connectivity within and between sets of hubs in *a priori* regions

of interest. In particular, we were interested in hippocampus, mPFC, and IPFC based on previous work (Schapiro et al., 2013), as well as TPJ based on its role in social processing (Tso et al., 2018). We further broke this down into left and right portions of the following regions: hippocampus, TPJ, vmPFC, dmPFC, and IPFC. Left and right hippocampus were defined based on the Harvard-Oxford atlas labels. Unfortunately, the Schaefer parcellation does not provide labels for mPFC, IPFC, or TPJ, and we therefore used a data-driven approach to assign parcels to these anatomical regions. Specifically, parcels were considered part of dorsal medial prefrontal cortex (dmPFC) if their center of gravity had x-coordinates between -20 and +20, y-coordinates greater than +30, and z-coordinates greater than 0. Parcels were considered part of ventral medial prefrontal cortex (vmPFC) if their center of gravity had x-coordinates between -20 and +20, y-coordinates greater than +20, and z-coordinates less than 0. Parcels were considered part of lateral prefrontal cortex (IPFC) if their center of gravity had x-coordinates more negative than -20 or more positive than +20, y-coordinates greater than 0, and z-coordinates between -40 and +40. Parcels were considered part of TPJ if they had a center of gravity with x-coordinates more negative than -40 or more positive than +40, y-coordinates between -10 and -60, and z-coordinates between 0 and 30.

We computed the average strength of connectivity within and between each set of ROIs. We first averaged the connectivity strength across all hubs within each of the *a priori* ROIs across both the social and non-social networks to obtain each ROI's within-region domain-general connectivity. Next, we averaged the connectivity between each pair of ROIs to compute the between-region domain-general connectivity. We also computed the within-region and between-region average connectivity for the social versus non-social connectivity matrix. We then repeated this procedure for our 500 null model networks, and *z*-scored the average ROI connectivity statistics for the domain-general analysis and the social versus non-social analysis.

## 2.9 Cognitive systems involved in network learning

We next examined whether the hubs identified based on node strength were clustered in particular cognitive systems. Using the Yeo 17-system parcellation which subdivides cognitive control, default mode, dorsal attention, ventral attention, somatomotor, limbic, and visual systems into 2-3 subsystems each (Schaefer et al., 2018; Yeo et al., 2011), we counted the number of hubs that were located in each system. Because the hubs were defined based on the null models, we randomly permuted the system labels for each hub 500 times; this procedure allowed us to maintain the total number of hubs and the total number of nodes in each system. We then compared the actual number of hubs in each system to the distribution of hubs in each system from the 500 permutations to calculate *z*-scores and *p*-values for each system. These statistics thus represented the probability that at least that many hubs were located in each system compared to randomly assigned system labels.

We also computed the average connectivity within and between each set of systems. We first averaged the connectivity across all regions within each of the Yeo cognitive systems across both the social network condition and the non-social control condition to compute each system's within-system domain-general connectivity. Next, we averaged the connectivity between each pair of systems to compute the between-system domain-general connectivity. We also computed the within-system and between-system average connectivity for the social versus non-social connectivity matrix. We then repeated this procedure for our 500 null model networks, and *z*-scored the average system connectivity statistics for the domain-general analysis and the social versus non-social analysis.

## 2.10 Association between network learning and brain hubs

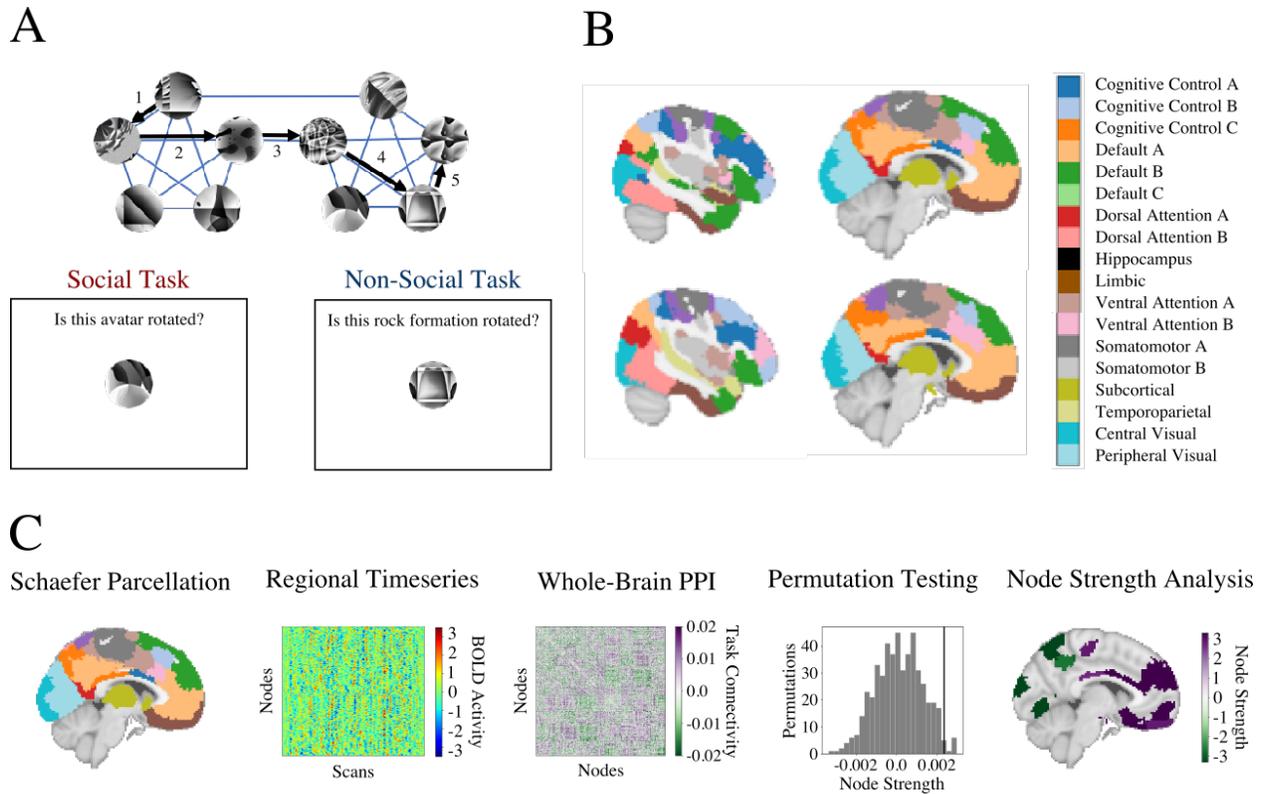
In addition to examining which brain regions operate as hubs supporting network learning, we were also interested in whether individual differences in hub connectivity might be associated with individual differences in learning. That is, are individuals with stronger connectivity in the hubs identified above also better at learning network structure? To address this question, we took the subject-level connectivity matrices and computed the average node strength for each set of hubs (domain-general hubs, social hubs, and non-social hubs) for each subject, as well as the average node strength of hubs in *a priori* ROIs (hippocampus, dmPFC, IPFC, and TPJ). This process yielded 11 node strength metrics (social hub node strength, left TPJ hub node strength, right TPJ hub node strength, non-social hub node strength, left dmPFC node strength, right dmPFC node strength, left IPFC node strength, right IPFC node strength, domain-general cortical hubs node strength, left hippocampus node strength, and right hippocampus node strength).

For each subject, we converted the node strength values to *z*-scores by comparing the average node strength to a distribution of average node strengths obtained from 500 null model networks. We then constructed a separate linear mixed effects model for each node strength metric. Each model included node strength, node type (transition versus non-transition), network type (social versus non-social), order (social network first versus non-social network first), trial number (standardized), and the two-way, three-way, and four-way interactions between these variables as predictors of log-transformed RT; node type, network type, and trial number were included as within-subjects variables, and both order and hub connectivity were included as between-subjects variables. We then examined simple effects within network type for each hub with a significant node type x node strength interaction or a significant node type x network type x node strength interaction. Adding head motion as a covariate or adding mean activation for transition versus non-transition trials as a covariate did not alter the results that we report (see Supplementary Results).

In addition to examining the relationship between node strength and network learning, we also examined whether average connectivity within different types of hubs was associated with network learning. To address this question, we computed the average connectivity within each set of hubs (domain-general hubs, social hubs, and non-social hubs) for each subject, converted all values to *z*-scores based on the non-parametric statistical approach described above, and constructed linear mixed effects models also using the approach described above. Adding head motion as a covariate or adding mean activation for transition versus non-transition trials as a covariate did not alter the results that we report (see Supplementary Results).

## 2.11 Data and code availability

The data and code to reproduce all analyses and figures in this paper are available in Github repository [[https://github.com/stompson/Tompson\\_Network\\_Learning\\_fMRI](https://github.com/stompson/Tompson_Network_Learning_fMRI)].



**Figure 1. Schematic of methods.** (A) Illustration of a random walk on a graph in which nodes are fractal images, and edges indicate allowable transitions between nodes. In the social network learning task, participants were told that the images were avatars representing people; in the control condition, participants were told that the images were rock formations. During the experiment, participants only saw the continuous stream of images, without any explicit indication of the graph from which the stream was drawn. Here we depict the graph to convey to the reader the statistical relationships among the images critical for the analysis. (B) Functional MRI data was collected continuously as participants performed the task. A network model of functional connectivity was constructed by first parcellating the brain into 400 cortical regions of interest using the Schaefer atlas (Schaefer et al., 2018) and then complementing that cortical parcellation with a parcellation of the subcortex composed of 10 additional regions of interest using the Harvard-Oxford atlas (Frazier et al., 2005). Each parcel was represented as a node in a brain network. (C) We used a whole-brain psychophysiological interaction model (Gerchen et al., 2014) to simultaneously estimate connectivity and its relation to the task. We first extracted the average timeseries from 410 brain regions defined by the Schaefer and Harvard-Oxford atlases. For each pair of regions  $i$  and  $j$ , we computed a multiple regression with the timeseries of activation in region  $i$  as the dependent variable and the timeseries of activation in region  $j$  as the independent variable; we also included a boxcar function to represent the timeseries of each task, and we used a separate interaction term to represent the interaction between activation in region  $j$  and the task timeseries. We constructed a  $410 \times 410$  functional connectivity matrix where the  $ij^{\text{th}}$  element of the matrix represented the task-dependent connectivity (beta weight for the interaction term) between region  $i$  and region  $j$ . This process was then repeated 500 times by shuffling the condition labels for the trials to generate null models of connectivity for each pair of brain regions. We computed node strength as the sum of the connectivity of each brain region with all other brain regions. We then compared the true node strength metric to the distribution of null model values to derive a  $p$ -value or  $z$ -score, with higher  $z$  values indicating a score that is stronger than would be expected given a random trial order.

### 3. Results

#### 3.1 Behavioral evidence for network learning

We began our investigation by assessing participant reaction time, and its dependence on the graph architecture from which the stream of stimuli were drawn. Initially, the data were collapsed across the social and non-social learning tasks, and we observed significant cross-cluster surprisal, as indicated by an increase in reaction time on trials transitioning from one cluster to another than on trials moving within a cluster ( $B=0.015$ ,  $SE=0.003$ ,  $t(24.461)=4.382$ ,  $p<0.001$ ). We also observed a significant decrease in reaction time in later trials compared to earlier trials ( $B=-0.021$ ,  $SE=0.005$ ,  $t(32.989)=-3.992$ ,  $p<0.001$ ). We did not observe a significant difference in the magnitude of the cross-cluster surprisal effect between social and non-social networks ( $B=-0.002$ ,  $SE=0.002$ ,  $t(31,212.004)=-0.802$ ,  $p=0.423$ ), nor a significant difference in the speed at which the cross-cluster surprisal effect emerged ( $B=-0.004$ ,  $SE=0.002$ ,  $t(31,712.291)=1.778$ ,  $p=0.075$ ). When we examined the two tasks separately, we found a significant cross-cluster surprisal effect for both the social network task ( $B=0.013$ ,  $SE=0.004$ ,  $t(25.614)=3.537$ ,  $p=0.002$ ) and the non-social control task ( $B=0.016$ ,  $SE=0.006$ ,  $t(23.866)=2.901$ ,  $p=0.008$ ), consistent with the idea that participants were able to learn both graph structures. Participants also did not differ in their mean accuracy for the social network task ( $M=84.8\%$ ,  $SD=10.1\%$ ) and for the non-social control task ( $M=84.9\%$ ,  $SD=9.3\%$ ,  $t(25)=0.091$ ,  $p=0.928$ ).

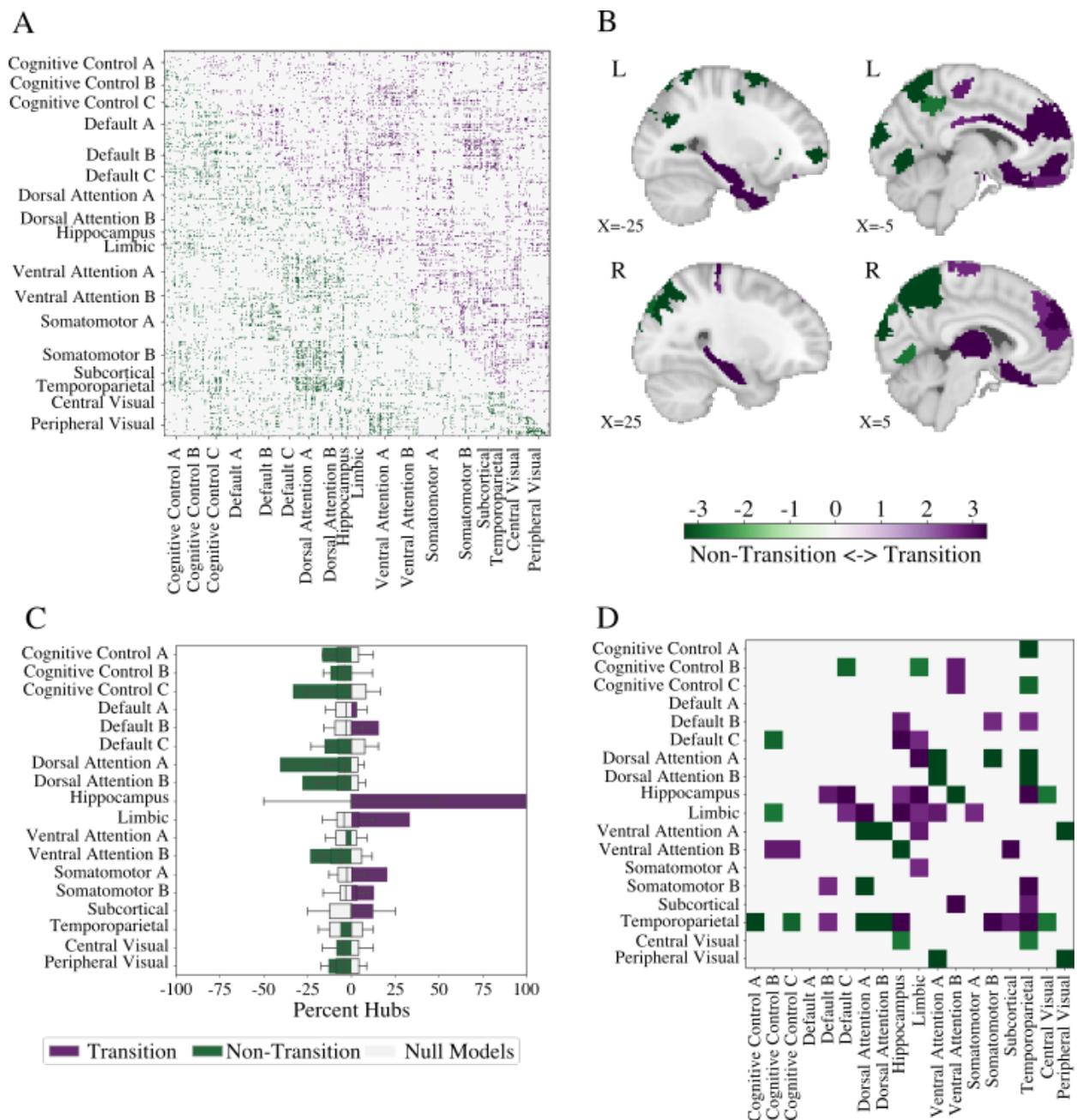
#### 3.2 Functional hubs that support network learning

We next sought to identify the brain network features that support network learning generally. We examined the relationship between brain networks and network learning at two levels of resolution: the region-level and the system-level. This approach enables us to examine both which specific brain regions are involved in network learning as well as which groups of brain regions are involved in network learning.

We began by identifying the functional hubs that were shared across the social network task and the non-social control task. Specifically, we computed the average strength for each node across the brain networks extracted by PPI during both conditions. We normalized each strength value to obtain a z-score, by comparing the node strength to a null distribution (see Methods and Figure 2A). We found that hippocampus, thalamus, dmPFC, STG, temporal pole, and OFC displayed significantly greater node strength for transition trials than for non-transition trials across both tasks (FDR corrected  $p<0.05$  over brain regions; Figure 2B). In contrast, visual cortex, IPL, TPJ, insula, and dlPFC exhibited significantly greater node strength for non-transition trials than for transition trials across both tasks (FDR corrected  $p<0.05$  over brain regions; the full list of regions is included in Table S1 in Supplementary Results).

To better understand this pattern of results, we assessed to what degree functional hubs were found within putative functional systems defined by an *a priori* community assignment (Frazier et al., 2005; Yeo et al., 2011). We found that 86% of the regions displaying significantly greater node strength for transition than non-transition trials were concentrated in hippocampus, limbic system, default mode systems, and sensory systems (see Figure 2C). In particular, the number of such regions in the hippocampus ( $Z=2.748$ , FDR-corrected  $p=0.027$ ) and limbic system ( $Z=3.291$ , FDR-corrected  $p<0.001$ ) were significantly greater than expected in a non-parametric permutation-based null model in which regions were randomly assigned to

systems. In contrast, we found that 78% of the regions displaying significantly greater node strength for non-transition than transition trials were concentrated in cognitive control and attention systems. In particular, the number of hubs in the cognitive control subsystem C ( $Z=2.878$ , FDR-corrected  $p=0.018$ ), the dorsal attention subsystem A ( $Z=3.291$ , FDR-corrected  $p<0.001$ ), and the dorsal attention subsystem B ( $Z=2.226$ , FDR-corrected  $p=0.039$ ) were significantly greater than expected in the non-parametric permutation-based null model. These results indicate that different brain systems may underlie the ability to learn different features of network structures.



**Figure 2. Patterns of functional connectivity in the brain are distinct in transition trials compared to non-transition trials.** (A) Functional connectivity matrix showing edges with significantly different weights (FDR corrected  $p<0.05$  over edges) for transition (upper triangle in purple) versus non-transition (lower triangle in green) trials. (B) Whole-brain map showing regions with significantly different strength z-scores for transition (purple) versus non-transition (green) trials. (C) Box plot showing the percent of hubs for transition (purple), non-transition (green), and null models (grey) across brain systems. (D) Heatmap showing the distribution of significantly different z-scores for transition (purple) versus non-transition (green) trials across brain systems.

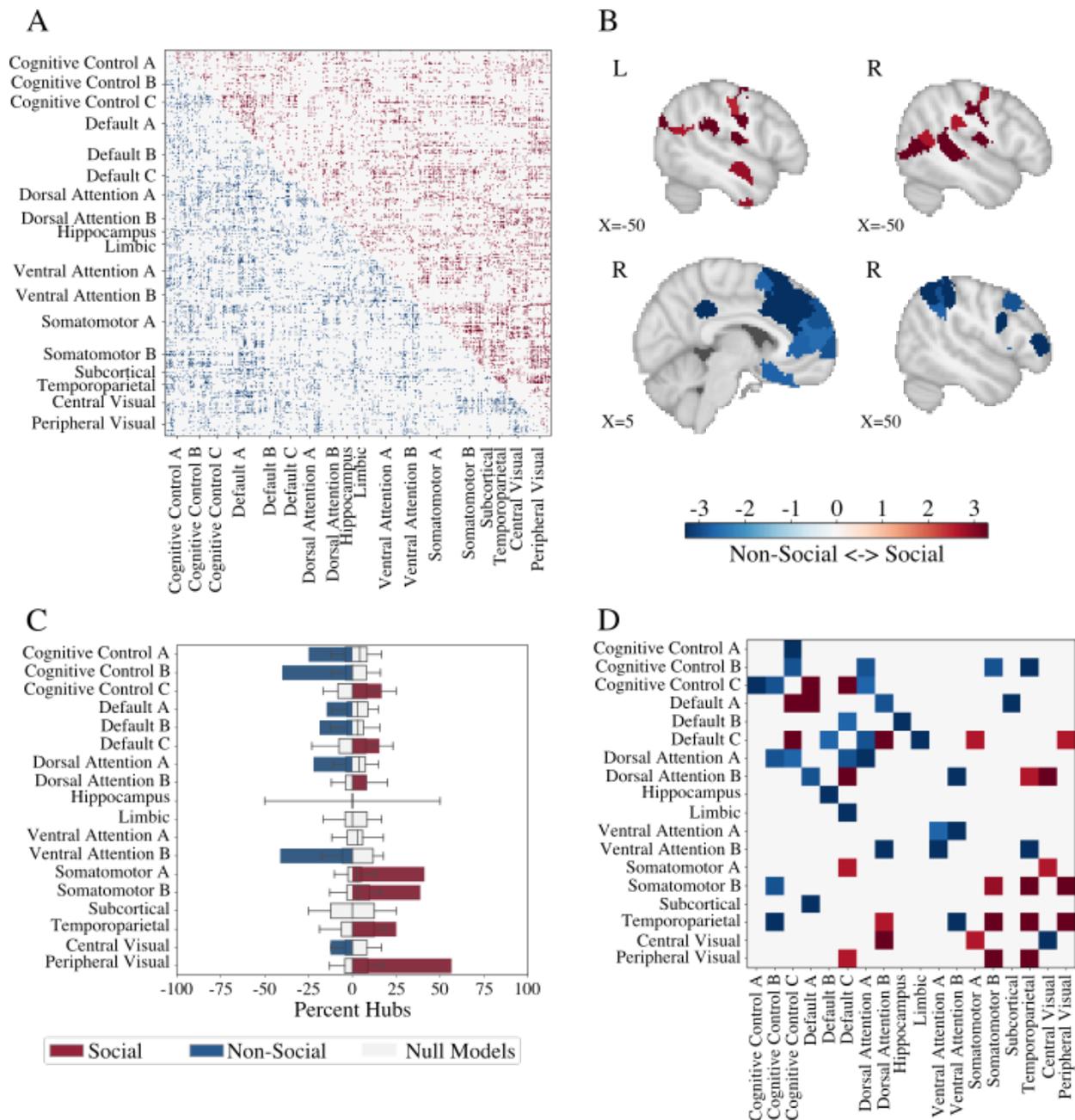
non-transition (green) trials (FDR corrected  $p < 0.05$  over brain regions). (C) Bar graph showing the percentage of network hubs identified in each cognitive system relative to the percentage identified in a non-parametric null model (gray box plots). (D) By averaging edge weights both within and between putative functional systems, we constructed a system-level connectivity matrix. Here we display that matrix showing systems with significantly different connectivity for transition (purple) versus non-transition (green) trials (FDR corrected  $p < 0.05$  over edges in the system-level matrix).

In a complementary analysis, we examined these relationships at a coarser, system level by averaging edge weights within and between the same putative functional systems (Frazier et al., 2005; Yeo et al., 2011), and then we tested whether these coarse-grained elements differed for transition and non-transition trials. We observed that functional connectivity within the limbic and default mode systems was significantly greater for transition trials than for non-transition trials. Furthermore, functional connectivity between the hippocampus and default mode system, between the hippocampus and somatomotor system, and between the default mode and somatomotor systems, was significantly greater for transition trials than for non-transition trials. In contrast, connectivity within the visual system and connectivity between frontal cognitive control systems and default mode systems was significantly greater for non-transition trials than for transition trials (see Figure 2D). Taken together, these results suggest that memory, default mode, and sensory systems have greater node strength as well as connectivity with each other during transitions between clusters than during trials occurring within a cluster.

### **3.3 Functional hubs that preferentially support social network learning**

After characterizing functional hubs involved in both conditions, we next sought to determine whether some hubs preferentially supported social network learning. We once again examined the relationship between brain networks and network learning at a coarse-grained systems level as well as a more fine-grained region level, which allows us to examine which specific brain regions are preferentially involved in social network learning as well as which groups of brain regions are preferentially involved in social network learning.

To address this question, we computed the difference in average node strength for each potential hub for social versus non-social networks. We normalized each difference value to obtain a  $z$ -score by comparing it to a null distribution (see Methods). We observed significantly greater node strength for the social network condition than the non-social control condition in TPJ, PCC, middle cingulate, precentral gyrus, postcentral gyrus, and visual cortex (FDR-corrected  $p < 0.05$ ; see Figure 3B; full list of regions is included in Table S2 in Supplementary Results). For completeness, we also assessed whether some hubs were stronger in the non-social control condition. We observed significantly greater node strength in the non-social control condition than in the social network condition in bilateral dmPFC, IFG, middle cingulate, inferior temporal lobe, visual cortex, and dlPFC (FDR-corrected  $p < 0.05$ ; see Figure 3B and Table S3). Collectively, these results suggest distinct regional involvement in social network learning in comparison to a non-social control condition.



**Figure 3. Patterns of functional connectivity in the brain are distinct for social network learning compared to a non-social control condition.** (A) Functional connectivity matrix showing edges with significantly different weights for the social network condition (upper triangle in red) versus the non-social control condition (lower triangle in blue). Significance was assessed after an FDR correction at  $p < 0.05$  over edges. (B) Whole-brain map showing regions with significantly different strength z-scores for trials in the social network condition (red) versus in the non-social control condition (blue). Significance was assessed after an FDR correction at  $p < 0.05$  over brain regions. (C) Bar graph showing percentage of network hubs identified in each cognitive system for the social network condition versus the non-social control condition, relative to the percentage identified in a non-parametric null model (gray box plots). (D) By averaging edge weights both within and between putative functional systems, we constructed a system-level connectivity matrix. Here we display that matrix showing systems with significantly different connectivity for the social network condition (red) versus the non-social control condition (blue). Significance was assessed after FDR correction at  $p < 0.05$  over edges in the system-level matrix.

To better understand this pattern of results, we assessed to what degree these functional hubs were found within putative functional systems defined by an *a priori* system assignment (Yeo et al., 2011). We found that 81% of the regions displaying significantly greater node strength in the social network condition than in the non-social control condition were concentrated in the default mode, somatomotor, and visual systems (Figure 3C). In particular, the number of such regions in somatomotor subsystem A ( $Z=3.291$ , FDR-corrected  $p<0.001$ ), somatomotor subsystem B ( $Z=3.291$ , FDR-corrected  $p<0.001$ ), and peripheral visual subsystem ( $Z=3.291$ , FDR-corrected  $p<0.001$ ) were significantly greater than expected in a non-parametric permutation-based null model in which regions were assigned to systems uniformly at random. The number of such regions in the temporoparietal default subsystem trended towards significance after FDR-correction ( $Z=1.995$ , FDR-corrected  $p=0.090$ ). In contrast, we found that 65% of the regions displaying significantly greater node strength for the non-social control condition than the social network condition were concentrated in cognitive control and attention systems. In particular, the number of such regions in cognitive control subsystem A ( $Z=3.291$ , FDR-corrected  $p<0.001$ ), cognitive control subsystem B ( $Z=3.291$ , FDR-corrected  $p<0.001$ ), default subsystem B ( $Z=2.290$ , FDR-corrected  $p=0.050$ ), dorsal attention subsystem A ( $Z=2.290$ , FDR-corrected  $p=0.050$ ), and ventral attention subsystem B ( $Z=3.291$ , FDR-corrected  $p<0.001$ ) were significantly greater than expected in the same non-parametric permutation-based null model. The number of such regions in default subsystem A was marginally significant after FDR-correction ( $Z=1.995$ , FDR-corrected  $p=0.090$ ).

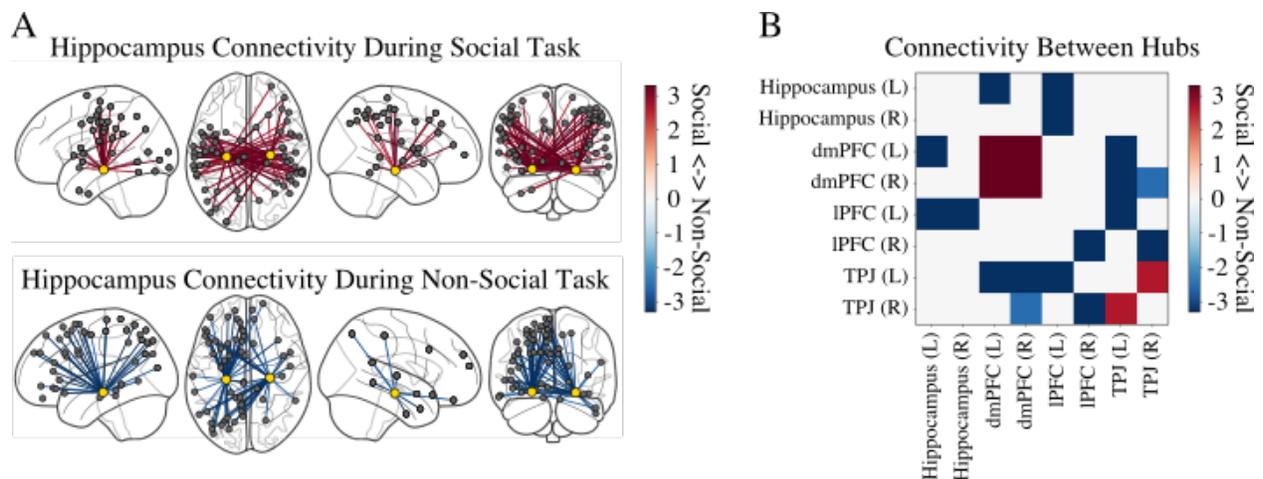
In a complementary analysis, we examined these relationships at a coarser, system level by averaging edge weights within and between the same putative functional systems (Yeo et al., 2011), and then we tested whether these coarse-grained elements differed for the social network condition compared to the non-social control condition. We observed that functional connectivity was significantly greater within the default mode and somatomotor systems for the social network condition than for the non-social control condition (FDR-corrected  $p<0.05$ ; see Figure 3D). In contrast, we observed that functional connectivity was significantly greater within the dorsal attention, ventral attention, and central visual systems for the non-social control condition than for the social network condition (FDR-corrected  $p<0.05$ ). Expanding our assessment to both within- and between-system connectivity, we found that the temporoparietal, default mode, and somatomotor systems were more strongly connected to each other and to the peripheral visual subsystem in the social network condition, suggesting that the hubs that we identified in the temporoparietal cortex may be integrating sensory information with the social context. In contrast, the non-social control condition was primarily characterized by increased connectivity between frontal cognitive control and attention systems, as well as limbic systems and frontal default systems. Thus, cognitive systems involved in social network learning reconfigure to form distinct subnetworks from those involved in non-social network learning.

### **3.4 Specific connectivity patterns of functional hubs support social network learning**

To complement the whole-brain analyses reported in the previous sections, we wished to evaluate specific connectivity patterns emanating from functional hubs that might preferentially support social network learning. We began our investigation by considering the specific role of the hippocampus, largely motivated by important prior work offering evidence in favor of its role in similar tasks (Schapiro et al., 2016, 2014). We found that the hippocampus was connected to different regions during the social network task compared to the non-social control task (Tables

S4 and S5 in Supplementary Results). In particular, the left hippocampus had significantly stronger connectivity with the dmPFC and IFG during the non-social control condition (FDR-corrected  $p < 0.05$ ; Figure 4A), whereas the left hippocampus had significantly stronger connectivity with TPJ during the social network condition than the non-social control condition (FDR-corrected  $p < 0.05$ ). The right hippocampus also had significantly stronger connectivity with IFG during the non-social control condition (FDR-corrected  $p < 0.05$ ) and significantly stronger connectivity with bilateral TPJ during the social network condition (FDR-corrected  $p < 0.05$ ). During both conditions, we found that the hippocampus exhibited significantly greater connectivity with hubs specific to the non-social control task, including the dmPFC and IFG (FDR-corrected  $p < 0.05$ ).

To examine if these relationships existed at the systems-level, we averaged edge weights of hubs within *a priori* regions of interest (hippocampus, dmPFC, IPFC, and TPJ) in order to examine how connectivity between hubs involved in both conditions might interact with hubs preferentially supporting learning in the social network task or the non-social control task. We found that hippocampus exhibited significantly greater connectivity with non-social hubs including dmPFC and IFG for non-social tasks than social tasks (FDR-corrected  $p < 0.05$ ; Figure 4B). This set of findings suggests that brain regions that have the strongest connectivity during non-social (relative to social) tasks are engaging with memory systems in support of learning the network structure of non-social networks.



**Figure 4. Functional connectivity between the hippocampus and other brain areas.** (A, top) Significant edges representing greater connectivity between the hippocampus (indicated in yellow) and other hubs during the social network learning condition than during the non-social control condition (top image). (A, bottom) Significant edges representing greater connectivity between the hippocampus (indicated in yellow) and other hubs during the non-social control condition than during the social network condition (bottom image). Significance is assessed after FDR correction at  $p < 0.05$  over the number of edges. (B) Connectivity matrix representing the average connectivity within and between hubs in *a priori* defined ROIs, thresholded at an FDR-corrected  $p < 0.05$ .

### 3.5 Relation between brain connectivity and behavioral performance

Finally, we sought to test whether individual differences in hub connectivity were associated with how well individuals learned the network architecture from which the sequence of stimuli were drawn. To address this question, we calculated the average node strength for each of the hubs identified above in the TPJ, dmPFC, IPFC, and hippocampus, separately for each participant. Specifically, we considered (i) social hubs as the TPJ regions that had significantly

greater node strength for the social network condition than for the non-social control condition, (ii) non-social hubs as the medial and lateral PFC regions that had significantly greater node strength for the non-social control condition than for the social network condition, (iii) domain-general cortical hubs as TPJ and mPFC regions that had significantly greater node strength than other brain areas across both conditions, and (iv) domain-general subcortical hubs as the hippocampus, which had significantly greater node strength than other brain areas across both conditions. This process yielded node strength values for 11 hubs and three metrics of hub-to-hub connectivity (see Methods and Supplementary Results for a full list and for further details regarding the results).

We constructed a linear mixed effects model with hub connectivity, node type (transition versus non-transition), network type (social versus non-social), trial number (standardized), and the two-way, three-way, and four-way interactions between these variables as predictors of log-transformed RT; node type, network type, and trial number were included as within-subjects variables, and both order and hub connectivity were included as between-subjects variables. Five significant effects survived FDR-correction for multiple comparisons.

First, we observed a significant 2-way interaction between node type and social hub node strength ( $B=0.008$ ,  $SE=0.003$ ,  $t(190.896)=2.730$ , FDR-corrected  $p=0.049$ ), such that the cross-cluster surprisal effect was significantly greater across network types when individuals had greater node strength in the left TPJ (Figure 5A). However, the 2-way interaction between node type and left TPJ was not significant when social networks ( $B=0.008$ ,  $SE=0.004$ ,  $t(25.315)=1.741$ , FDR-corrected  $p=0.117$ ) and non-social networks ( $B=0.006$ ,  $SE=0.005$ ,  $t(20.891)=1.200$ , FDR-corrected  $p=0.528$ ) were examined separately.

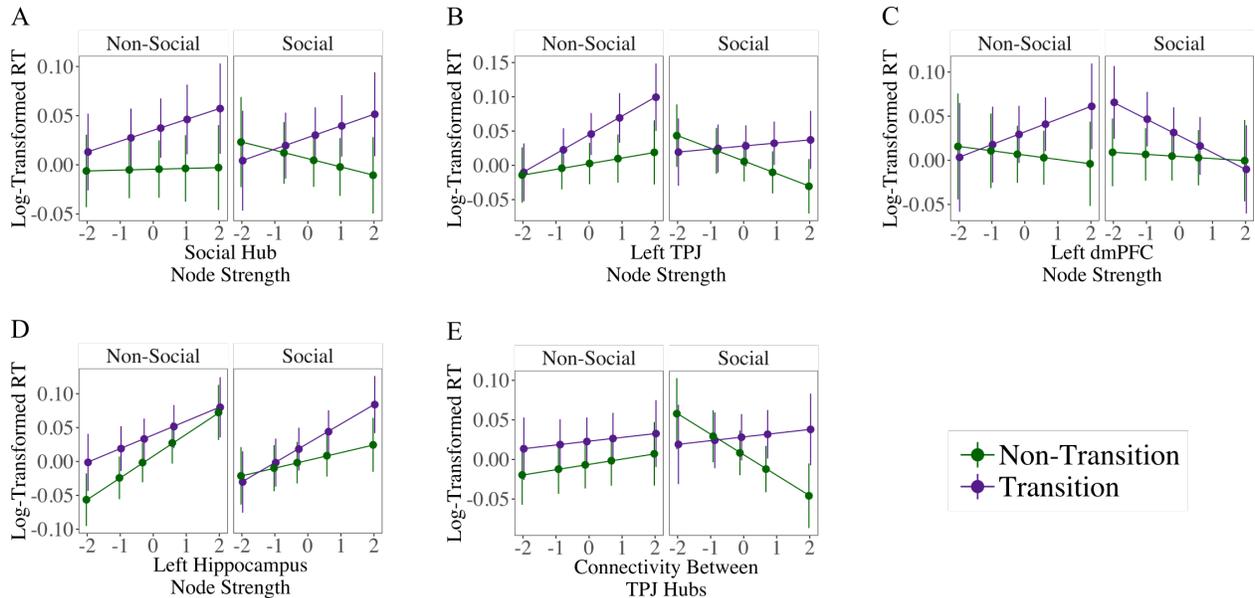
Second, we observed a significant 2-way interaction between node type and left TPJ node strength ( $B=0.011$ ,  $SE=0.003$ ,  $t(253.818)=4.018$ , FDR-corrected  $p=0.001$ ), such that the cross-cluster surprisal effect was significantly greater across network types when individuals had greater node strength in the left TPJ (Figure 5B). The 2-way interaction between node type and left TPJ was significant for social networks ( $B=0.008$ ,  $SE=0.004$ ,  $t(1300.076)=2.360$ , FDR-corrected  $p=0.046$ ) but not for non-social networks ( $B=0.011$ ,  $SE=0.005$ ,  $t(21.105)=2.195$ , FDR-corrected  $p=0.197$ ).

Third, we observed a significant 3-way interaction between node type, network type, and left hippocampus node strength ( $B=0.007$ ,  $SE=0.003$ ,  $t(231.158)=2.571$ , FDR-corrected  $p=0.050$ ; Figure 5C). The cross-cluster surprisal effect was greater for social networks when individuals had greater node strength in the left hippocampus, although this effect was not significant after FDR-correction ( $B=0.008$ ,  $SE=0.004$ ,  $t(24.752)=2.136$ , FDR-corrected  $p=0.071$ ); the interaction was not significant for non-social networks ( $B=-0.005$ ,  $SE=0.006$ ,  $t(20.354)=-0.813$ , FDR-corrected  $p=0.532$ ).

Fourth, we observed a significant 3-way interaction between node type, network type, and left dmPFC node strength ( $B=-0.009$ ,  $SE=0.003$ ,  $t(191.084)=-2.859$ , FDR-corrected  $p=0.033$ ), such that the cross-cluster surprisal effect was greater for non-social networks when individuals had greater node strength in left dmPFC (Figure 5D). However, the 2-way interaction between node type and left dmPFC was not significant when social networks ( $B=-0.002$ ,  $SE=0.003$ ,  $t(19.254)=-0.668$ , FDR-corrected  $p=0.512$ ) and non-social networks ( $B=0.007$ ,  $SE=0.007$ ,  $t(21.650)=1.025$ , FDR-corrected  $p=0.528$ ) were examined separately.

Fifth, we observed a significant 3-way interaction between node type, network type, and connectivity between social hubs in left and right TPJ ( $B=0.008$ ,  $SE=0.003$ ,  $t(208.910)=2.898$ , FDR-corrected  $p=0.033$ ; Figure 5E). The cross-cluster surprisal effect was significantly greater for

social networks when individuals had greater connectivity between social hubs in left and right TPJ ( $B=0.012$ ,  $SE=0.004$ ,  $t(227.291)=2.919$ , FDR-corrected  $p=0.019$ ), whereas the interaction was not significant for non-social networks ( $B=-0.003$ ,  $SE=0.005$ ,  $t(20.066)=-0.600$ , FDR-corrected  $p=0.556$ ). None of the other hub connectivity metrics were significantly associated with behavior (see Supplementary Results for full output of the mixed effects model).



**Figure 5. Association between brain connectivity and network learning.** Greater (A) social hub node strength, (B) left TPJ node strength, (C) left dmPFC node strength, (D) left hippocampus node strength, and (E) connectivity between TPJ hubs was associated with a larger cross-cluster surprisal effect for both social and non-social networks, as evidenced by slower (larger) RT for transition (purple) than non-transition (green) trials.

#### 4. Discussion

In this paper, we investigate the brain networks involved in the implicit learning of community structure for social networks. Navigating interwoven layers of social connections is critical for success in a broad range of social interactions with co-workers, friends, family and strangers (Balkundi and Harrison, 2006; Decostanza et al., 2012; Fitzhugh and DeCostanza, 2016; Liu and Chen, 2003). Although there is an extensive literature examining how people learn relational information in the domains of language learning (Goldstein and Vitevitch, 2014; Lieberman et al., 2004), motor sequence learning (Hung et al., 2018; Kahn et al., 2018; Robertson, 2007), and statistical learning (Karuza et al., 2017; Schapiro et al., 2013; Turk-Browne et al., 2005), little is known about the neural mechanisms underlying how people learn relational information about social networks. Building on this extant literature, we showed both overlap and divergence in the brain networks involved in social and non-social network learning. We found that the hippocampus operated as a hub supporting both types of network learning. Temporoparietal regions of the default mode system operated as hubs preferentially supporting social network learning, where they were also more strongly connected to the hippocampus as well as to somatomotor and visual systems. Medial and lateral prefrontal cortical regions in both the default mode system and frontal cognitive control system operated as hubs preferentially

supporting non-social network learning, where they were also more strongly connected to the hippocampus as well as to cognitive control and attention systems. Furthermore, individuals who had stronger hubs in the temporoparietal default system were better at learning the community structure of the social networks. This work is the first neuroscience study to directly compare relational learning in social and non-social networks, and provides insights into how the brain reconfigures to support this learning process.

#### **4.1 Memory systems support learning both social and non-social networks**

Consistent with prior work examining how people learn relational information in other domains (Hung et al., 2018), the hippocampus and PFC have also been found to play an integral role in the learning of community structure present in networks comprised of visual images (Schapiro et al., 2016, 2013). Evidence suggests that multivariate patterns in the hippocampus and IFG represent information about which communities individual nodes belong to, whereas mPFC represents boundaries and transitions (Schapiro et al., 2016, 2013). Importantly, lesion studies also suggest that the hippocampus is necessary for learning the statistical relationships between non-social objects (Covington et al., 2018; Schapiro et al., 2014). Here, we extend this work by applying graph theory to show that the hippocampus is a primary hub supporting both social and non-social network learning. Both the hippocampus and parahippocampal gyrus have significantly greater node strength for transition trials than for non-transition trials. This pattern of results is consistent with work suggesting that the hippocampus integrates representations of objects and their spatiotemporal context from disparate cortical areas (Opitz, 2010), as well as evidence suggesting that hippocampal-cortical connections support simulation of future navigation in complex spatial maps (Javadi et al., 2017).

At the system-level, the hippocampus exhibits stronger connectivity with the default mode system for transition trials than for non-transition trials. This pattern of results is consistent with prior studies providing evidence that the default mode system is involved in developing and maintaining models that predict and simulate future events (Buckner, 2010). In the context of learning community structure in networks, transition trials may be interpreted as prediction errors, given that participants are responding more slowly because they anticipated the next trial to remain within a community (Kahn et al., 2018; Lynn et al., 2018; Schapiro et al., 2013). In the context of spatial navigation, some researchers have argued that prefrontal cortex monitors prediction errors and updates contextual information about the prediction, which is then integrated by the hippocampus into an updated mental model (Spiers and Gilbert, 2015). Similarly, medial default mode network regions process temporal information about motor patterns and support the testing of predictive models of future motor sequences (Carvalho et al., 2016). It is intuitively plausible that the hippocampus and regions of the default mode may be communicating with one another to update predictions about future trials and representations of the community structure.

Although the hippocampus operates as a hub during both social and non-social network learning, the pattern of connections to cortical areas differs for the two conditions. At the region level, the hippocampus is primarily connected to temporoparietal, somatomotor, and visual areas during social network learning, whereas it is primarily connected to medial prefrontal, lateral prefrontal, and inferior temporal areas during non-social network learning. This differentiation of connectivity is consistent with past work showing that temporoparietal regions support updating of prediction errors by validating mental models against sensory inputs (Kluger and Schubotz, 2017). Collectively, the pattern of findings suggests that there exist marked differences in how the

hippocampus is integrating information to update representations of the community structure of social versus non-social networks.

In the current study, we analyzed hippocampus as a single region, but recent work suggests that anterior and posterior subregions of the hippocampus support the learning of different network features. In particular, posterior hippocampus has been linked to learning more local spatial features such as number of nearby edges whereas anterior hippocampus has been linked to learning more global spatial features such as the average shortest path length to get to any other node (Javadi et al., 2017). Moreover, previous work examining the role of hippocampus in learning community structure of non-social networks found that multivariate representations of community structure in the hippocampus are primarily located in the anterior hippocampus (Schapiro et al., 2016). Future work could investigate whether anterior and posterior hippocampus are involved in learning local versus global features of social networks as well.

## **4.2 Nexus hubs versus top-down control of network learning**

Beyond the role of memory systems in network learning, we observed notable connectivity patterns in temporoparietal brain areas including the right TPJ. Temporoparietal, default, somatomotor, and peripheral visual subsystems were more strongly connected to each other during social network learning than during the non-social control condition. This pattern of results is consistent with prior work that suggests that TPJ may integrate sensory information with information about the broader social context to support decision-making (Carter et al., 2012; Carter and Huettel, 2013; Lahnakoski et al., 2012), in part due to its location at the intersection of brain regions involved in many different types of cognitive processes (Carter et al., 2012). In light of these prior studies, it would appear possible that TPJ is integrating information about the visual features of the images with motor responses from the task and with the social/online avatar framing of the task to support the participants' learning. Importantly, the node strength of left and right TPJ hubs as well as the connectivity among TPJ hubs was significantly correlated with cross-cluster surprisal for the social network condition but not for the non-social control condition, supporting the region's role in social network learning specifically.

By contrast, non-social network learning primarily involved connectivity with prefrontal brain areas including the dmPFC and IFG. Frontal cognitive control and attention subsystems were more strongly connected to each other, as well as more strongly connected to medial prefrontal default subsystems, during non-social network learning than during social network learning. One possible interpretation of this result is that cognitive control and attention areas are exerting top-down control to direct processing of sensory information (Lee and D'Esposito, 2012; Miller and Cohen, 2001; Spiers and Gilbert, 2015) in support of developing a mental representation of the non-social network. Here, prefrontal cortical regions may be either modulating sensory processing of stimuli to support the goal of learning the non-social network (Lee and D'Esposito, 2012), or may be modulating attention to information that hippocampus will integrate into an updated mental representation (Spiers and Gilbert, 2015).

It is also interesting to note that we did not find individual differences in any brain hubs or systems that were significantly correlated with individual differences in the magnitude of the cross-cluster surprisal effect during the non-social control condition. Contrasting the group-level findings with these individual-level findings, our data suggest that prefrontal systems are consistently engaged in the cognitive processes that we study here, but that individual differences

in behavior may require explanations that include other regional features, or other brain regions altogether.

### **4.3 Representing features of real-world social networks**

Our study expands upon recent work investigating how people represent the positions of individuals in a social network. The social brain is involved in many overlapping sub-processes, including mentalizing (dmPFC and TPJ), emotion processing (amygdala and insula), social reward (ventral striatum and vmPFC), and autobiographical memory (mPFC and PCC). In the context of representing information about social networks, brain regions associated with mentalizing and reward are involved in representing information about the social value of other college students in their dormitory (Morelli et al., 2018). People also develop neural representations of information about others' social position including their network centrality, brokerage opportunities in the network, and social distance from the perceiver (Parkinson et al., 2018, 2017, 2014). Collectively, these studies focus on the position of individual nodes in a social network, and we expand upon this work by considering meso-scale features of the social network including its community structure.

Here, we find that many of the brain regions implicated in representing the network position of individual nodes are also important for learning community structure (e.g., PCC and right TPJ). Importantly, we also show that functional connectivity between these brain regions, and functional connectivity with memory regions including the hippocampus, is important for learning meso-scale network features. A notable absence from our findings is the set of reward-related regions (e.g., ventral striatum, OFC, and vmPFC), which do not appear to be strongly involved in social network learning.

Although beyond the scope of the current study, it would also be important to examine whether these effects extend to learning real-world social networks. Given the ubiquitous nature of online social interactions in modern social life, we believe that this process closely mirrors how individuals mentally represent their social networks. Nevertheless, it is possible that true social exchange differs from anthropomorphizing images, particularly in the cognitive processes involved. Therefore, future work should investigate whether there are differences in the functional brain networks that support learning real-world social networks versus networks of avatars that represent people.

Furthermore, social networks in turn have been linked to individual differences in neural processes underlying information processing and decision-making. Belonging to the same group influences reward processing (Fareri, Niznikiewicz, Lee, & Delgado, 2012) and recruitment of TPJ has been linked to individual's position in their social network (O'Donnell et al., 2017; Schmäzle et al., 2017). For example, people who occupy more central positions in their social network tend to recruit TPJ more when making sense of social information (O'Donnell et al., 2017). People with less dense social networks also show greater functional connectivity between left and right TPJ when making sense of social contexts such as being excluded (Schmäzle et al., 2017) and those who are more receptive to peer influence show higher node strength of TPJ (Wasylyshyn et al., 2018). Our work expands on this extant literature by showing that people who have high TPJ node strength and greater functional connectivity between left and right TPJ learn social networks better. Learning information about the community structure of social networks might be important for how individuals position themselves in their social network, or the positions

of individuals within their social network might influence which brain regions they use to process social information.

#### **4.4 Methodological considerations and limitations**

Several methodological considerations are pertinent to this work. First, given the contrast between the two conditions in the current paradigm, one might anticipate that brain regions involved in considering the mental states of others might be most strongly involved in the social network condition. While we do find that TPJ is more strongly involved in the social condition, other mentalizing regions including dmPFC were actually more strongly involved in the non-social control condition. Moreover, we did not find any significant relationships between functional connectivity in TPJ and self-report measures of social traits (see Supplementary Results). Thus, it is difficult to conclude that social network learning demands additional mentalizing processes relative to non-social network learning. In the current study we used the same abstract shapes to represent people (social condition) or rock formations (non-social control condition). It is possible that a stronger manipulation, perhaps using real people or real social networks, might yield stronger recruitment of mentalizing brain regions.

Second, one key difference between the social and non-social tasks is that participants are told to focus on the people represented by the images in the social task but to focus on the patterns in the images in the non-social task. Thus, both the type of attentional focus and the type of cognitive process elicited could be different between the social and non-social tasks. However, if focusing on the patterns in the images were driving the connectivity differences, one might anticipate the visual regions would be more strongly implicated in the learning of non-social networks versus social networks. In fact, visual regions are more likely to have significantly greater node strength, connectivity with hippocampus, and connectivity with TPJ for social network task than for non-social network task. Future work could rule out this potential confound by instructing participants to focus on non-social but physically connected objects represented by the abstract images (e.g., the abstract images are symbols for subway stations).

Third, the experimental paradigm that we employed requires dozens of trials per node to detect the cross-cluster surprisal effect, and we were therefore limited in the size and complexity of the social network that we could test. Real-world social networks are often much larger than those we study here, and they also tend to display much more complex patterns of connections (Hill and Dunbar, 2003). There is some evidence that learnability of network structures is robust to the size of the communities within the network (Karuza et al., 2019). But, other work on network learning in non-social domains suggests that whether a network is organized into clusters versus more random networks or networks with a lattice-like structure influences how well it is learned (Kahn et al., 2018). Future work could investigate how people learn other, more naturalistic, configurations of social networks.

Fourth, one limitation of past neuroimaging work on how people learn relational information is that these studies have focused on activation in single brain regions or connectivity between pairs of brain regions (Hung et al., 2018; Lieberman et al., 2004; Schapiro et al., 2016, 2013). Implicitly learning the edges and community structure of a complex network and building a mental representation of that network likely involves processing and coordination from many brain regions and systems (Schapiro et al., 2016). In supplementary analyses we did not find any significant differences in univariate brain activation for social versus non-social network learning, and the relationships between brain connectivity and cross-cluster surprisal were significant even

after controlling for univariate brain activation (see Supplementary Results). These facts underscore the need for an explicitly multivariate approach that takes into account the distributed set of brain regions, and their interconnections, that support network learning. Here we address this need by capitalizing on recent advances in the nascent field of network neuroscience, which applies network science tools to understand how groups of brain regions interact to support both basic and complex cognitive processes across different temporal and spatial scales (Bassett and Sporns, 2017). Specifically, we examine network hubs, or brain areas that have strong connectivity to the rest of the brain, because they are thought to be central to coordinating communication between different brain systems and integrating information from different systems (van den Heuvel and Sporns, 2013).

#### **4.5 Conclusion**

In this study, we apply graph theory and network neuroscience tools to investigate how people learn information about social networks, including which communities each person in a social network belongs to. Navigating novel social contexts is a defining feature of modern social life (Rainie and Wellman, 2012) and learning information about how people in a new social group are connected to each other is an important facet of adapting to and fitting into these new social contexts (Balkundi and Harrison, 2006; Decostanza et al., 2012; Fitzhugh and DeCostanza, 2016; Liu and Chen, 2003). We found that people learned the community structure of both social and non-social networks, but the learning of these two types of networks was differentially associated with brain network architecture. Brain regions implicated in memory processes operate as hubs supporting learning community structure of both social and non-social networks, whereas brain regions implicated in social processing operate as hubs supporting learning of social (but not non-social) community structure. Evidence supports the notion that social hubs integrate information from sensory systems, whereas non-social hubs in cognitive control systems are exerting top-down control of information processing. Broadly, our study provides a promising approach to determine how the brain supports social network learning, extending our knowledge about the impact of functional brain networks and social context on these learning processes.

**Declarations of interest:** none

## **References**

- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., Varoquaux, G., 2014. Machine learning for neuroimaging with scikit-learn. *Front. Neuroinform.* 8, 14. <https://doi.org/10.3389/fninf.2014.00014>
- Adolphs, R., 2009. The social brain: neural basis of social knowledge. *Annu. Rev. Psychol.* 60, 693–716. <https://doi.org/10.1146/annurev.psych.60.110707.163514>
- Avants, B. B., Tustison, N., & Song, G., 2009. Advanced normalization tools (ANTS). *Insight J.* 2, 1–35.
- Balkundi, P., Harrison, D.A., 2006. TIES, LEADERS, AND TIME IN TEAMS: STRONG INFERENCE ABOUT NETWORK STRUCTURE'S EFFECTS ON TEAM VIABILITY AND PERFORMANCE. *Acad. Manag. J.* 49, 49–68. <https://doi.org/10.5465/AMJ.2006.20785500>
- Bassett, D.S., Sporns, O., 2017. Network neuroscience. *Nat Neurosci* 20, 353–364.
- Behzadi, Y., Restom, K., Liau, J., Liu, T.T., 2007. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage.* <https://doi.org/10.1016/j.neuroimage.2007.04.042>
- Buckner, R.L., 2010. The Role of the Hippocampus in Prediction and Imagination. *Annu. Rev. Psychol.* 61, 27–48. <https://doi.org/10.1146/annurev.psych.60.110707.163508>
- Bureau, U.C., 2018. CPS Historical Migration/Geographic Mobility Tables.
- Carter, R.M., Bowling, D.L., Reeck, C., Huettel, Scott A, Lee, D., Danchin, E., Giraldeau, L.-A., Valone, T.J., Wagner, R.H., Rendell, L., Behrens, T.E.J., Hunt, L.T., Rushworth, M.F.S., Olsson, A., Ochsner, K.N., Woolrich, M.W., Asch, S.E., Raafat, R.M., Chater, N., Frith, C.D., Kameda, T., Ohtsubo, Y., Takezawa, M., Bonaccio, S., Dalal, R., Yoshida, W., Seymour, B., Friston, K.J., Dolan, R.J., Festinger, L., Mitchell, J.P., Corbetta, M., Patel, G., Shulman, G.L., Rapoport, A., Erev, I., Abraham, E. V., Olson, D.E., Hampton, A.N., O'Doherty, J.P., Clithero, J.A., Huettel, S. A., Goren, C.C., Sarty, M., Wu, P.Y., Ro, T., Russell, C., Lavie, N., Blakemore, S.-J., Tankersley, D., Stowe, C.J., Fletcher, P.C., Schirmer, J., Ames, D.L., Gilbert, D.T., Heider, F., Simmel, M., Krach, S., Saxe, R., Kanwisher, N., Decety, J., Lamm, C., Bhatt, M., Lohrenz, T., Camerer, C.F., Montague, P.R., Shepherd, S. V., Deaner, R.O., Platt, M.L., Harris, L.T., Fiske, S.T., Brainard, D.H., Jenkinson, M., Bannister, P., Brady, M., Smith, S., Smith, S.M., Desikan, R.S., Carrington, S.J., Bailey, A.J., Macrae, C.N., Banaji, M.R., Gallagher, H.L., Whitfield-Gabrieli, S., Sip, K.E., Abe, N., Suzuki, M., Mori, E., Itoh, M., Fujii, T., Bossaerts, P., King-Casas, B., Krajbich, I., Adolphs, R., Tranel, D., Denburg, N.L., Jackson, P.L., Sommerville, J.A., Chaminade, T., Meltzoff, A.N., Tomlin, D., Sanfey, A.G., Singer, T., Hanke, M., Kim, J.-H., Jones, E., Oliphant, T., Peterson, P., Hunter, J.D., Smith, A.T., Kosillo, P., Williams, A.L., Nichols, T.E., Wexler, A., Perner, J., Aichhorn, M., Kronbichler, M., Staffen, W., Ladurner, G., 2012. A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science* 337, 109–111. <https://doi.org/10.1126/science.1219681>
- Carter, R.M., Huettel, S.A., 2013. A nexus model of the temporal-parietal junction. *Trends Cogn. Sci.* 17, 328–36. <https://doi.org/10.1016/j.tics.2013.05.007>
- Carvalho, F.M., Chaim, K.T., Sanchez, T.A., de Araujo, D.B., 2016. Time-Perception Network and Default Mode Network Are Associated with Temporal Prediction in a Periodic Motion Task. *Front. Hum. Neurosci.* 10, 268. <https://doi.org/10.3389/fnhum.2016.00268>
- Ciric, R., Wolf, D.H., Power, J.D., Roalf, D.R., Baum, G., Ruparel, K., Shinohara, R.T., Elliott, M.A., Eickhoff, S.B., Davatzikos, C., Gur, R.C., Gur, R.E., Bassett, D.S., Satterthwaite,

- T.D., 2017. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *Neuroimage*.  
<https://doi.org/10.1016/j.neuroimage.2017.03.020>
- Covington, N. V., Brown-Schmidt, S., Duff, M.C., 2018. The necessity of the hippocampus for statistical learning. *J. Cogn. Neurosci.* [https://doi.org/10.1162/jocn\\_a\\_01228](https://doi.org/10.1162/jocn_a_01228)
- Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–73.
- Decostanza, A.H., Dirosa, G.A., Rogers, S.E., Slaughter, A.J., Estrada, A.X., 2012. Researching Teams: Nothing’s Going to Change Our World. *Ind. Organ. Psychol.* 5, 36–39.  
<https://doi.org/10.1111/j.1754-9434.2011.01400.x>
- Eichenbaum, H., 2000. A cortical–hippocampal system for declarative memory. *Nat. Rev. Neurosci.* <https://doi.org/10.1038/35036213>
- Fareri, D.S., Niznikiewicz, M.A., Lee, V.K., Delgado, M.R., 2012. Social network modulation of reward-related signals. *J. Neurosci.* 32, 9045–52.  
<https://doi.org/10.1523/JNEUROSCI.0610-12.2012>
- Fine, A.B., Jaeger, T.F., Farmer, T.A., Qian, T., 2013. Rapid expectation adaptation during syntactic comprehension. *PLoS One* 8. <https://doi.org/10.1371/journal.pone.0077661>
- Fiser, J., Aslin, R.N., 2005. Encoding multielement scenes: statistical learning of visual feature hierarchies. *J. Exp. Psychol. Gen.* 134, 521–37. <https://doi.org/10.1037/0096-3445.134.4.521>
- Fiser, J., Aslin, R.N., 2002. Statistical learning of higher-order temporal structure from visual shape sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 458–467.  
<https://doi.org/10.1037/0278-7393.28.3.458>
- Fitzhugh, S.M., DeCostanza, A.H., 2016. Organizational tie preservation and dissolution during crisis, in: *INSNA Sunbelt XXXVI*. Newport Beach, CA.
- Frazier, J.A., Chiu, S., Breeze, J.L., Makris, N., Lange, N., Kennedy, D.N., Herbert, M.R., Bent, E.K., Koneru, V.K., Dieterich, M.E., Hodge, S.M., Rauch, S.L., Grant, P.E., Cohen, B.M., Seidman, L.J., Caviness, V.S., Biederman, J., 2005. Structural Brain Magnetic Resonance Imaging of Limbic and Thalamic Volumes in Pediatric Bipolar Disorder. *Am. J. Psychiatry* 162, 1256–1265. <https://doi.org/10.1176/appi.ajp.162.7.1256>
- Garvert, M.M., Dolan, R.J., Behrens, T.E., 2017. A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *Elife* 6. <https://doi.org/10.7554/eLife.17086>
- Gerchen, M.F., Bernal-Casas, D., Kirsch, P., 2014. Analyzing task-dependent brain network changes by whole-brain psychophysiological interactions: A comparison to conventional analysis. *Hum. Brain Mapp.* 35, 5071–5082. <https://doi.org/10.1002/hbm.22532>
- Gheysen, F., Van Opstal, F., Roggeman, C., Van Waelvelde, H., Fias, W., 2010. Hippocampal contribution to early and later stages of implicit motor sequence learning. *Exp. Brain Res.* 202, 795–807. <https://doi.org/10.1007/s00221-010-2186-6>
- Goldstein, R., Vitevitch, M.S., 2014. The influence of clustering coefficient on word-learning: How groups of similar sounding words facilitate acquisition. *Front. Psychol.* 5, 1307.  
<https://doi.org/10.3389/fpsyg.2014.01307>
- Gorgolewski, K.J., Esteban, O., Ellis, D.G., Notter, M.P., Ziegler, E., Johnson, H., Hamalainen, C., Yvernault, B., Burns, C., Manhães-Savio, A., Jarecka, D., Markiewicz, C.J., Salo, T., Clark, Daniel, Waskom, M., Wong, J., Modat, M., Dewey, B.E., Clark, M.G., Dayan, M., Loney, F., Madison, C., Gramfort, A., Keshavan, A., Berleant, S., Pinsard, B., Goncalves, M., Clark, Dav, Cipollini, B., Varoquaux, G., Wassermann, D., Rokem, A., Halchenko,

- Y.O., Forbes, J., Moloney, B., Malone, I.B., Hanke, M., Mordom, D., Buchanan, C., Pauli, W.M., Huntenburg, J.M., Horea, C., Schwartz, Y., Tungaraza, R., Iqbal, S., Kleesiek, J., Sikka, S., Frohlich, C., Kent, J., Perez-Guevara, M., Watanabe, A., Welch, D., Cumba, C., Ginsburg, D., Eshaghi, A., Kastman, E., Bougacha, S., Blair, R., Acland, B., Gillman, A., Schaefer, A., Nichols, B.N., Giavasis, S., Erickson, D., Correa, C., Ghayoor, A., Küttner, R., Haselgrove, C., Zhou, D., Craddock, R.C., Haehn, D., Lampe, L., Millman, J., Lai, J., Renfro, M., Liu, S., Stadler, J., Glatard, T., Kahn, A.E., Kong, X.-Z., Triplett, W., Park, A., McDermottroe, C., Hallquist, M., Poldrack, R., Perkins, L.N., Noel, M., Gerhard, S., Salvatore, J., Mertz, F., Broderick, W., Inati, S., Hinds, O., Brett, M., Durnez, J., Tambini, A., Rothmei, S., Andberg, S.K., Cooper, G., Marina, A., Mattfeld, A., Urchs, S., Sharp, P., Matsubara, K., Geisler, D., Cheung, B., Floren, A., Nickson, T., Pannetier, N., Weinstein, A., Dubois, M., Arias, J., Tarbert, C., Schlamp, K., Jordan, K., Liem, F., Saase, V., Harms, R., Khanuja, R., Podranski, K., Flandin, G., Papadopoulos Orfanos, D., Schwabacher, I., McNamee, D., Falkiewicz, M., Pellman, J., Linkersdörfer, J., Varada, J., Pérez-García, F., Davison, A., Shachnev, D., Ghosh, S., 2017. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. 0.13.1.  
<https://doi.org/10.5281/ZENODO.581704>
- Halford, G.S., Wilson, W.H., Phillips, S., 2010. Relational knowledge: The foundation of higher cognition. *Trends Cogn. Sci.* <https://doi.org/10.1016/j.tics.2010.08.005>
- Hill, R.A., Dunbar, R.I.M., 2003. Social network size in humans. *Hum. Nat.* 14, 53–72.  
<https://doi.org/10.1007/s12110-003-1016-y>
- Hung, Y.H., Frost, S.J., Molfese, P., Malins, J.G., Landi, N., Mencl, W.E., Rueckl, J.G., Bogaerts, L., Pugh, K.R., 2018. Common Neural Basis of Motor Sequence Learning and Word Recognition and Its Relation With Individual Differences in Reading Skill. *Sci. Stud. Read.* <https://doi.org/10.1080/10888438.2018.1451533>
- Javadi, A.-H., Emo, B., Howard, L.R., Zisch, F.E., Yu, Y., Knight, R., Pinelo Silva, J., Spiers, H.J., 2017. Hippocampal and prefrontal processing of network topology to simulate the future. *Nat. Commun.* 8, 14652. <https://doi.org/10.1038/ncomms14652>
- Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841. [https://doi.org/10.1016/S1053-8119\(02\)91132-8](https://doi.org/10.1016/S1053-8119(02)91132-8)
- Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. Fsl. *Neuroimage* 62, 782–90. <https://doi.org/10.1016/j.neuroimage.2011.09.015>
- Kahn, A.E., Karuza, E.A., Vettel, J.M., Bassett, D.S., 2018. Network constraints on learnability of probabilistic motor sequences. *Nat. Hum. Behav.* 2, 936–947.  
<https://doi.org/10.1038/s41562-018-0463-8>
- Karuza, E.A., Farmer, T.A., Smith, F.X., Fine, A.B., Jaeger, T.F., 2014. On-Line Measures of Prediction in a Self-Paced Statistical Learning Task, in: *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*. pp. 725–730.
- Karuza, E.A., Kahn, A.E., Bassett, D.S., 2019. Human Sensitivity to Community Structure Is Robust to Topological Variation. *Complexity* 2019, 1–8.  
<https://doi.org/10.1155/2019/8379321>
- Karuza, E.A., Kahn, A.E., Thompson-Schill, S.L., Bassett, D.S., 2017. Process reveals structure: How a network is traversed mediates expectations about its architecture. *Sci. Rep.* 7, 12733. <https://doi.org/10.1038/s41598-017-12876-5>
- Karuza, E.A., Newport, E.L., Aslin, R.N., Starling, S.J., Tivarus, M.E., Bavelier, D., 2013. The

- neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain Lang.* <https://doi.org/10.1016/j.bandl.2012.11.007>
- Karuza, E.A., Thompson-Schill, S.L., Bassett, D.S., 2016. Local Patterns to Global Architectures: Influences of Network Topology on Human Learning. *Trends Cogn. Sci.* <https://doi.org/10.1016/j.tics.2016.06.003>
- Kernbach, J.M., Yeo, B.T.T., Smallwood, J., Margulies, D.S., Schotten, M.T. de, Walter, H., Sabuncu, M.R., Holmes, A.J., Gramfort, A., Varoquaux, G., Thirion, B., Bzdok, D., 2018. Subspecialization within default mode nodes characterized in 10,000 UK Biobank participants. *Proc. Natl. Acad. Sci.*
- Kluger, D.S., Schubotz, R.I., 2017. Strategic adaptation to non-reward prediction error qualities and irreducible uncertainty in fMRI. *Cortex* 97, 32–48. <https://doi.org/10.1016/j.cortex.2017.09.017>
- Labor, U.S.D. of, 2017. Number of jobs, labor market experience, and earnings growth among Americans at 50: Results from a longitudinal survey.
- Lahnakoski, J.M., Glerean, E., Salmi, J., Jääskeläinen, I.P., Sams, M., Hari, R., Nummenmaa, L., 2012. Naturalistic fMRI Mapping Reveals Superior Temporal Sulcus as the Hub for the Distributed Brain Network for Social Perception. *Front. Hum. Neurosci.* 6, 233. <https://doi.org/10.3389/fnhum.2012.00233>
- Lee, T.G., D’Esposito, M., 2012. The dynamic nature of top-down signals originating from prefrontal cortex: a combined fMRI-TMS study. *J. Neurosci.* 32, 15458–66. <https://doi.org/10.1523/JNEUROSCI.0627-12.2012>
- Lenhart, A., Smith, A., Anderson, M., Duggan, M., Perrin, A., 2015. Teens, Technology, and Friendships, Pew Research Center.
- Lieberman, M.D., 2007. Social cognitive neuroscience: a review of core processes. *Annu. Rev. Psychol.* 58, 259–289. <https://doi.org/10.1146/annurev.psych.58.110405.085654>
- Lieberman, M.D., Chang, G.Y., Chiao, J., Bookheimer, S.Y., Knowlton, B.J., 2004. An event-related fMRI study of artificial grammar learning in a balanced chunk strength design. *J. Cogn. Neurosci.* <https://doi.org/10.1162/089892904322926764>
- Liu, M., Chen, X., 2003. Friendship networks and social, school and psychological adjustment in Chinese junior high school students. *Psychol. Sch.* 40, 5–17. <https://doi.org/10.1002/pits.10066>
- Long, J., Feng, Y., Liao, H., Zhou, Q., Urbin, M.A., 2018. Motor Sequence Learning Is Associated With Hippocampal Subfield Volume in Humans With Medial Temporal Lobe Epilepsy. *Front. Hum. Neurosci.* 12, 367. <https://doi.org/10.3389/fnhum.2018.00367>
- Lydon-Staley, D.M., Ciric, R., Satterthwaite, T.D., Bassett, D.S., 2018. Evaluation of confound regression strategies for the mitigation of motion artifact in studies of dynamic resting state functional connectivity. <https://doi.org/10.31234/OSF.IO/RW5ZT>
- Lynn, C.W., Kahn, A.E., Bassett, D.S., 2018. Structure from noise: Mental errors yield abstract representations of events.
- McLaren, D.G., Ries, M.L., Xu, G., Johnson, S.C., 2012. A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *Neuroimage* 61, 1277–1286. <https://doi.org/10.1016/j.neuroimage.2012.03.068>
- Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>
- Morelli, S.A., Leong, Y.C., Carlson, R.W., Kullar, M., Zaki, J., 2018. Neural detection of socially valued community members. *Proc. Natl. Acad. Sci. U. S. A.* 115, 8149–8154.

- <https://doi.org/10.1073/pnas.1712811115>
- Niche, 2014. Going Away to College: Data on 350,000 HS Grads - Niche Ink.
- O'Donnell, Matthew Brook, Bayer, J.B., Cascio, C.N., Falk, E.B., O'Donnell, Matthew B., Bayer, J.B., Cascio, C.N., Falk, E.B., 2017. Neural bases of recommendations differ according to social network structure. *Soc. Cogn. Affect. Neurosci.* 12, nsw158. <https://doi.org/10.1093/scan/nsw158>
- Opitz, B., 2010. Neural binding mechanisms in learning and memory. *Neurosci. Biobehav. Rev.* <https://doi.org/10.1016/j.neubiorev.2009.11.001>
- Parkinson, C., Kleinbaum, A.M., Wheatley, T., 2018. Similar neural responses predict friendship. *Nat. Commun.* 9, 332. <https://doi.org/10.1038/s41467-017-02722-7>
- Parkinson, C., Kleinbaum, A.M., Wheatley, T., 2017. Spontaneous neural encoding of social network position. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-017-0072>
- Parkinson, C., Liu, S., Wheatley, T., 2014. A Common Cortical Metric for Spatial, Temporal, and Social Distance. *J. Neurosci.* <https://doi.org/10.1523/jneurosci.2159-13.2014>
- Rainie, L., Wellman, B., 2012. *Networked: The new social operating system.* MIT Press, Cambridge, MA.
- Repeke, L., Benet-Martínez, V., 2018. The (Diverse) Company You Keep: Content and Structure of Immigrants' Social Networks as a Window Into Intercultural Relations in Catalonia. *J. Cross. Cult. Psychol.* 49, 924–944. <https://doi.org/10.1177/0022022117733475>
- Robertson, E.M., 2007. The serial reaction time task: implicit motor skill learning? *J. Neurosci.* 27, 10073–5. <https://doi.org/10.1523/JNEUROSCI.2747-07.2007>
- Saffran, J.R., Aslin, R.N., Newport, E.L., 1996. Statistical learning by 8-month-old infants. *Science* 274, 1926–8.
- Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.-N., Holmes, A.J., Eickhoff, S.B., Yeo, B.T.T., 2018. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb. Cortex* 28, 3095–3114. <https://doi.org/10.1093/cercor/bhx179>
- Schapiro, A.C., Gregory, E., Landau, B., McCloskey, M., Turk-Browne, N.B., 2014. The Necessity of the Medial Temporal Lobe for Statistical Learning. *J. Cogn. Neurosci.* [https://doi.org/10.1162/jocn\\_a\\_00578](https://doi.org/10.1162/jocn_a_00578)
- Schapiro, A.C., Rogers, T.T., Cordova, N.I., Turk-Browne, N.B., Botvinick, M.M., 2013. Neural representations of events arise from temporal community structure. *Nat. Neurosci.* 16, 486–92. <https://doi.org/10.1038/nn.3331>
- Schapiro, A.C., Turk-Browne, N.B., Norman, K.A., Botvinick, M.M., 2016. Statistical learning of temporal community structure in the hippocampus. *Hippocampus* 26, 3–8. <https://doi.org/10.1002/hipo.22523>
- Schmälzle, R., O'Donnell, M.B., Garcia, J.O., Cascio, C.N., Bayer, J., Bassett, D.S., Vettel, J.M., Falk, E.B., 2017. Brain connectivity dynamics during social interaction reflect social network structure. *Proc. Natl. Acad. Sci.* 114, 5153–5158. <https://doi.org/10.1073/pnas.1616130114>
- Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, S208–S219. <https://doi.org/10.1016/j.neuroimage.2004.07.051>
- Spiers, H.J., Gilbert, S.J., 2015. Solving the detour problem in navigation: a model of prefrontal

- and hippocampal interactions. *Front. Hum. Neurosci.* 9, 125.  
<https://doi.org/10.3389/fnhum.2015.00125>
- Stadtfeld, C., Vörös, A., Elmer, T., Boda, Z., Raabe, I.J., 2019. Integration in emerging social networks explains academic failure and success. *Proc. Natl. Acad. Sci. U. S. A.* 116, 792–797. <https://doi.org/10.1073/pnas.1811388115>
- Strayer, N., 2016. The Great Out-of-State Migration: Where Students Go. *New York Times*.
- Tompson, S.H., Kahn, A.E., Falk, E.B., Vettel, J.M., Bassett, D.S., 2018. Individual differences in learning social and nonsocial network structures. *J. Exp. Psychol. Learn. Mem. Cogn.* <https://doi.org/10.1037/xlm0000580>
- Tso, I.F., Rutherford, S., Fang, Y., Angstadt, M., Taylor, S.F., 2018. The “social brain” is highly sensitive to the mere presence of social information: An automated meta-analysis and an independent study. *PLoS One* 13, e0196503. <https://doi.org/10.1371/journal.pone.0196503>
- Turk-Browne, N.B., Jungé, J., Scholl, B.J., 2005. The automaticity of visual statistical learning. *J. Exp. Psychol. Gen.* 134, 552–64. <https://doi.org/10.1037/0096-3445.134.4.552>
- Turk-Browne, N.B., Scholl, B.J., Johnson, M.K., Chun, M.M., 2010. Implicit Perceptual Anticipation Triggered by Statistical Learning. *J. Neurosci.* 30, 11177–11187. <https://doi.org/10.1523/JNEUROSCI.0858-10.2010>
- van den Heuvel, M.P., Sporns, O., 2013. Network hubs in the human brain. *Trends Cogn. Sci.* 17, 683–96. <https://doi.org/10.1016/j.tics.2013.09.012>
- Wasylyshyn, N., Falk, B.H., Garcia, J.O., Cascio, C.N., O’Donnell, M.B., Bingham, C.R., Simons-Morton, B., Vettel, J.M., Falk, E.B., 2018. Global brain dynamics during social exclusion predict subsequent behavioral conformity. *Soc. Cogn. Affect. Neurosci.* 13, 182–191. <https://doi.org/10.1093/scan/nsy007>
- Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., Fischl, B., Liu, H., Buckner, R.L., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106.