Supplementary Materials

Supplementary Method

Brain Activation Analyses

In addition to examining connectivity between brain regions, we also examined task-related activation for transition versus non-transition trials for the social versus non-social network learning tasks. Neuroimaging data was preprocessed using nipype[70] implemented in python 2.7 using a combination of AFNI[72], FSL[73,74], and ANTs[75]. The functional data first underwent de-spiking to smooth outliers in each voxel using AFNI's 3dDespike. We next used FSL's MCFLIRT[76] to apply a rigid transformation to correct for head motion, and then we used 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 using N4 Bias Field Correction implemented in ANTs, and then the skull-stripped structural image was bias-corrected. ANTs was used to compute the transformation parameters for the bias-corrected mean functional image to the bias-corrected and skull-stripped high resolution structural image. ANTs was used to warp the skull-stripped high-resolution structural image to the MNI152 2mm brain template provided by FSL. 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 MNI152 2mm brain template. Finally, FSL's fslmaths was used to apply a 6mm Gaussian kernel to smooth the warped 4D functional image.

We then constructed a first-level general linear model (GLM) for each participant and each task with a boxcar function for transition versus non-transition trials convolved with a canonical hemodynamic response function. Additional regressors of no interest included six motion parameters derived from the motion correction preprocessing step. All first-level analyses steps were conducted with nistats in python 2.7. We next constructed a second-level GLM with an intercept modeling the trial type (transition versus non-transition), a regressor for network type (social versus non-social), and a regressor for subject (to account for the within-subjects design). We tested the effect of trial type using a one-sample t-test and the effect of network type using a paired-samples t-test. Results were thresholded using a whole-brain FDR-corrected $p<0.05$ and cluster-threshold of $k>5$. All second-level analyses steps were conducted with nistats in python 2.7.

Ruling out confounds

In the main manuscript, we also tested whether individual differences in hub connectivity might be associated with individual differences in learning. Here, we add mean brain activation from the supplementary GLM analysis as a covariate to demonstrate that hub connectivity predicts learning above and beyond mean brain activation. We extracted mean activation in each node in our atlas from the transition versus non-transition first-level contrast for each participant and each task, and then averaged activation across hubs identified in each set of hubs that had been identified in the main manuscript as moderating the cross-cluster surprisal effect (left hippocampus node strength, social hub node strength, left TPJ node strength, left dmPFC node strength, and connectivity between social hubs in left and right TPJ).

We then constructed a separate linear mixed effects model for each connectivity metric. Each model included connectivity, 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 plus mean activation in the same set of hubs as was included in that connectivity metric as predictors of log-transformed RT; node type, network type, and trial number were included as within-subjects variables, and order, hub connectivity, and mean hub activation were included as between-subjects variables. We then repeated these analyses with head motion (mean framewise displacement across all runs) as a covariate in place of mean activation to rule out individual differences in head motion as a confound.

Supplementary Results

Behavioral evidence for network learning

In the main manuscript, we performed an initial assessment of the behavioral data that sought to extract evidence for network learning across both the social network learning task and the non-social control task. Here, we also ask whether there were important differences in behavior across the two tasks. We found that when the non-social control task was presented first, reaction times were significantly smaller than when the social network task was presented first (B=0.026, SE=0.012, $t(23.590)$=2.111, p=0.046). Moreover, we observed an interaction between node type and order, such that the cross-cluster surprisal effect was larger for participants who completed the non-social control task first (B=0.035, SE=0.014, $t(23.644)$=2.579, p=0.017). We did not hypothesize these effects *a priori*, and we suggest that further work should be done to ensure their reliability and replicability across other similar experiments.

Activation for transition versus non-transition trials

There was significantly greater activation for transition versus non-transition trials in visual cortex, parahippocampal gyrus, and dlPFC (see Supplementary Figure S1). There was significantly greater activation for non-transition trials versus transition trials in mPFC, vlPFC, striatum, PCC, inferior parietal lobe (IPL), and inferior temporal gyrus (ITG; see Supplementary Figure S1).

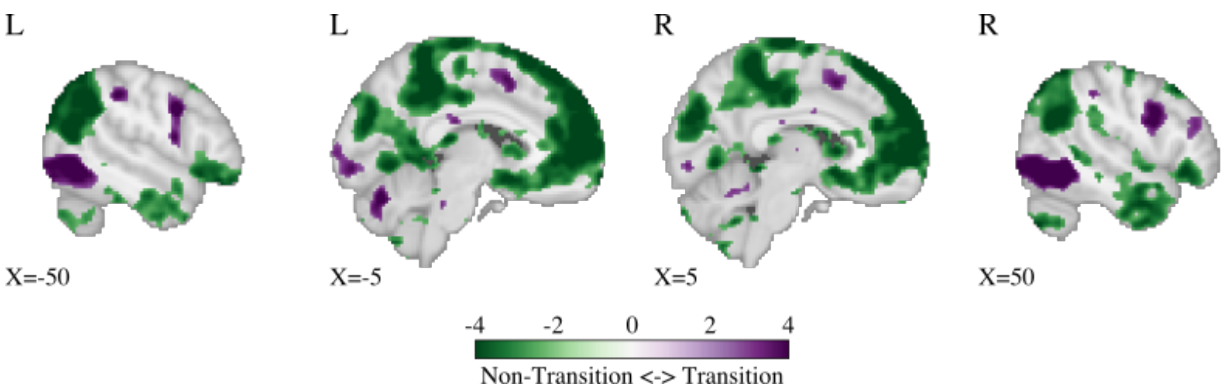

Figure S1. Differences in brain activation for transition versus non-transition trials. Whole-brain map showing regions with significantly different BOLD activation for transition (purple) versus non-transition (green) trials (whole-brain FDR corrected p<0.05).

Activation for social versus non-social networks

There were no brain regions with significantly different activation for social versus non-social tasks (FDR-corrected $p<0.05$).

Correlation between functional connectivity and social traits

There were no significant correlations between subject's node strength for social versus non-social networks in domain-general, social, or non-social hubs and social traits, including self-construal, holistic cognitive style, or perspective-taking (see Table S6). There were also no significant correlations between connectivity within each set of hubs for social versus non-social networks and social traits. We also did not observe any significant correlations between these connectivity metrics and IQ.

Ruling out mean activation as a confound

The significant interactions between node type, network type, and connectivity reported in the main manuscript were also still significant after including mean activation for the transition versus non-transition trials as a covariate in our mixed effects models. First, we observed a significant 3-way interaction between node type, network type, and left hippocampus node strength (B=0.007, SE=0.003, $t(231.577)=2.552$, FDR-corrected $p=0.019$). Second, 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, $t(209.359)=2.921$, FDR-corrected $p=0.011$). Third, we observed a significant 3-way interaction between node type, network type, and left dmPFC node strength (B=-0.008, SE=0.003, $t(192.413)=-2.876$, FDR-corrected $p=0.011$). Fourth, we observed a significant 2-way interaction between node type and social hubs node strength (B=0.008, SE=0.003, $t(190.633)=2.773$, FDR-corrected $p=0.015$). Finally, we observed a significant 2-way interaction between node type and left TPJ node strength (B=0.011, SE=0.003, $t(254.461)=4.049$, FDR-corrected $p<0.001$).

Ruling out head motion as a confound

The significant interactions between node type, network type, and connectivity reported in the main manuscript were also still significant after including head motion as a covariate in our mixed effects models. First, we observed a significant 3-way interaction between node type, network type, and left hippocampus node strength (B=0.007, SE=0.003, $t(230.148)=2.560$, FDR-corrected $p=0.019$). Second, 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, $t(210.401)=2.878$, FDR-corrected $p=0.011$). Third, we observed a significant 3-way interaction between node type, network type, and left dmPFC node strength (B=-0.008, SE=0.003, $t(191.665)=-2.887$, FDR-corrected $p=0.011$). Fourth, we observed a significant 2-way interaction between node type and social hubs node strength (B=0.008, SE=0.003, $t(191.014)=2.719$, FDR-corrected $p=0.018$). Finally, we observed a significant 2-way interaction between node type and left TPJ node strength (B=0.011, SE=0.003, $t(254.044)=3.997$, FDR-corrected $p<0.001$).

Supplementary Tables

Table S1. Node strength for both social and non-social networks

Region	x,y,z[a]	Z[b]	k[c]
Transition > Non-Transition Trials			
R Dorsal Medial Prefrontal Cortex (dmPFC)	(12, 58, 36)	3.291	3076
L Superior Temporal Lobe	(-36, -26, 2)	3.291	881
R Inferior Temporal Lobe	(64, -26, -28)	3.291	3201
R Rectus Gyrus	(10, 12, -22)	3.291	1703
L Precentral Gyrus	(-12, -30, 62)	3.291	135
R Postcentral Gyrus	(68, -14, 16)	3.291	1031
R Postcentral Gyrus	(50, -16, 60)	3.291	1238
L Hippocampus	(-6, -14, -16)	3.291	2636
R Hippocampus	(26, -20, -15)	3.291	***
L Paracentral Lobule	(0, -36, 54)	2.878	214
Non-Transition > Transition Trials			
R Middle Temporal Lobe	(68, -52, 0)	3.291	9242
L Middle Temporal Lobe	(-36, -66, 18)	3.291	915
L Precentral Gyrus	(-24, -16, 54)	3.291	495
L Inferior Parietal Lobe	(-30, -40, 44)	3.291	284
L Inferior Parietal Lobe	(-40, -66, 56)	3.291	304
L Occipital Lobe	(-20, -70, 36)	3.291	374
L Lingual Gyrus	(0, -82, -10)	3.291	636
L Insula	(-26, 22, -4)	3.291	831
L Superior Frontal Gyrus	(-12, -2, 64)	3.291	460
L Superior Frontal Gyrus	(-18, 56, -4)	3.291	1745
R Middle Frontal Gyrus	(50, 24, 38)	3.291	383
L Calcarine Gyrus	(-4, -60, 8)	3.291	242
R Temporoparietal Junction	(62, -36, 4)	2.878	294
L Inferior Parietal Lobe	(-52, -44, 40)	2.652	199

*** Included in Left Hippocampus cluster

[a] Stereotactic coordinates from MNI atlas, in mm, left/right (*x*), anterior/posterior (*y*), superior/inferior (*z*), respectively, R = right, L = left

[b] Z-score, significant at FDR-correct $p<0.05$.

[c] Spatial extent in cluster size, threshold ≥ 5 voxels.

Table S2. Hubs with significantly greater node strength for social than non-social condition

Region	x,y,z[a]	Z[b]	k[c]
Social > Non-Social Node Strength			
R Temporoparietal Junction	(68, -46, 12)	3.291	2443
L Precentral Gyrus	(-30, 0, 50)	3.291	2424
R Postcentral Gyrus	(68, -10, 22)	3.291	4657
L Paracentral Lobule	(-6, -34, 68)	3.291	783
R Occipital Lobe	(30, -88, 34)	3.291	6316
L Middle Temporal Lobe	(-46, -10, -14)	2.878	507
L Inferior Temporal Lobe	(-22, -4, -44)	2.878	604
R Parahippocampal Gyrus	(38, -36, -18)	2.878	348
L Middle Cingulate	(-2, -32, 46)	2.878	275

[a] Stereotactic coordinates from MNI atlas, in mm, left/right (*x*), anterior/posterior (*y*), superior/inferior (*z*), respectively, R = right, L = left

[b] Z-score, significant at FDR-correct $p<0.05$.

[c] Spatial extent in cluster size, threshold ≥ 5 voxels.

Table S3. Hubs with significantly greater node strength for non-social than social condition

Region	x,y,z[a]	Z[b]	k[c]
Non-Social > Social Node Strength			
R Dorsal Medial Prefrontal Cortex (dmPFC)	(30, 62, 16)	3.291	7036
R Inferior Temporal Lobe	(70, -46, -4)	3.291	452
L Inferior Temporal Lobe	(-52, -46, -26)	3.291	542
R Inferior Parietal Lobe (IPL)	(64, -52, 32)	3.291	3673
R Occipital Lobe	(48, -84, -4)	3.291	381
L Occipital Lobe	(-14, -106, 8)	3.291	2002
L Occipital Lobe	(-20, -70, 36)	3.291	408
L Superior Frontal Gyrus	(-18, 56, -4)	3.291	709
R Inferior Frontal Gyrus (IFG)	(56, 36, 2)	3.291	282
R Inferior Frontal Gyrus (IFG)	(44, 14, -10)	3.291	996
L Inferior Frontal Gyrus (IFG)	(-22, 16, -16)	3.291	1452
R Inferior Frontal Gyrus (IFG)	(64, 10, 16)	3.291	903
R Middle Cingulate	(16, -44, 32)	3.291	343
R Rectus Gyrus	(10, 12, -22)	2.652	309

[a] Stereotactic coordinates from MNI atlas, in mm, left/right (*x*), anterior/posterior (*y*), superior/inferior (*z*), respectively, R = right, L = left

[b] Z-score, significant at FDR-correct $p<0.05$.

[c] Spatial extent in cluster size, threshold ≥ 5 voxels.

Table S4. Brain regions with significantly greater hippocampal connectivity during the social network learning task

Region	x,y,z[a]	Z[b]	k[c]
Connectivity with Left Hippocampus			
L Middle Cingulate	(0, -28, 30)	3.291	271
R Inferior Temporal Lobe	(64, -54, -16)	3.291	1189
R Temporoparietal Junction	(68, -44, 16)	3.291	3672
R Rolandic Operculum	(66, -4, 6)	3.291	833
L Precuneus	(0, -74, 30)	3.291	550
L Postcentral Gyrus	(-32, -20, 42)	3.291	3313
L Occipital Lobe	(-32, -94, -4)	3.291	390
L Lingual Gyrus	(-4, -66, -4)	3.291	261
R Insula	(48, -14, -6)	3.291	261
L Insula	(-30, 8, 10)	3.291	1471
R Middle Frontal Gyrus	(52, 40, 18)	3.291	1544
L Middle Cingulate	(-2, -32, 46)	3.291	275
R Ventral Striatum	(14, 16, -8)	2.878	83
L Temporoparietal Junction	(-46, -32, 0)	2.878	563
L Fusiform Gyrus	(-24, -62, -16)	2.878	309
R Cuneus	(26, -96, 20)	2.878	469
R Middle Frontal Gyrus	(54, 10, 40)	2.652	521
R Middle Frontal Gyrus	(34, -2, 52)	2.652	304
R Precentral Gyrus	(32, -30, 66)	2.512	138
L Postcentral Gyrus	(-20, -38, 60)	2.512	269
Connectivity with Right Hippocampus			
R Ventral Striatum	(14, 16, -8)	3.291	83
L Middle Temporal Pole	(-16, 0, -28)	3.291	1003
R Inferior Temporal Lobe	(70, -46, -4)	3.291	452
L Supplementary Motor Area	(0, -4, 68)	3.291	206
L Rolandic Operculum	(-42, 0, 6)	3.291	259
L Precentral Gyrus	(-36, -16, 38)	3.291	567
R Postcentral Gyrus	(64, -12, 40)	3.291	636
L Inferior Parietal Lobe (IPL)	(-40, -66, 56)	3.291	304
L Inferior Parietal Lobe (IPL)	(-48, -24, 38)	3.291	1283
L Occipital Lobe	(-10, -100, -18)	3.291	2604
R Middle Frontal Gyrus	(52, 40, 18)	3.291	1553
L Temporoparietal Junction	(-46, -32, 0)	2.878	320

[a] Stereotactic coordinates from MNI atlas, in mm, left/right (*x*),

anterior/posterior (*y*), superior/inferior (*z*), respectively, R = right, L = left

[b] Z-score, significant at FDR-correct $p<0.05$.

[c] Spatial extent in cluster size, threshold ≥ 5 voxels.

Table S5. Brain regions with significantly greater hippocampal connectivity during the non-social control task

Region	x,y,z[a]	Z[b]	k[c]
Connectivity with Left Hippocampus			
L Middle Temporal Lobe	(-52, -24, -20)	3.291	438
L Supplementary Motor Area	(0, 2, 62)	3.291	9103
L Precuneus	(0, -62, 58)	3.291	557
L Superior Parietal Lobe	(-12, -84, 46)	3.291	2405
L Paracentral Lobule	(0, -40, 64)	3.291	562
R Dorsal Medial Prefrontal Cortex	(12, 58, 36)	2.878	301
L Middle Temporal Lobe	(-42, -60, 2)	2.878	956
R Inferior Temporal Lobe	(64, -20, -30)	2.878	1159
R Inferior Frontal Gyrus	(56, 30, -6)	2.878	840
Connectivity with Right Hippocampus			
L Middle Temporal Lobe	(-42, -60, 2)	3.291	565
R Inferior Temporal Lobe	(64, -20, -30)	3.291	1460
R Rectus Gyrus	(16, 44, -24)	3.291	1467
L Precuneus	(0, -72, 50)	3.291	2890
R Lingual Gyrus	(26, -54, 0)	3.291	1617
L Inferior Frontal Gyrus	(-22, 16, -16)	3.291	1114
R Temporoparietal Junction	(68, -46, 14)	2.878	195
R Superior Parietal Lobe	(30, -68, 60)	2.878	368
L Posterior Cingulate Cortex	(-10, -48, -6)	2.878	118
R Insula	(50, -16, 2)	2.878	157
L Anterior Cingulate Cortex	(0, -4, 32)	2.878	143

[a] Stereotactic coordinates from MNI atlas, in mm, left/right (*x*), anterior/posterior (*y*), superior/inferior (*z*), respectively, R = right, L = left

[b] Z-score, significant at FDR-correct p<0.05.

[c] Spatial extent in cluster size, threshold ≥ 5 voxels.

Table S6. Correlation between Hub Connectivity and Social Traits

Hub	Self-Construal	Holistic Cognitive Style	Perspective-Taking	Raven's Matrices
Node strength				
Domain-general hubs	-0.234 (0.608)	0.304 (0.564)	-0.004 (0.983)	-0.103 (0.919)
L Hippocampus	0.056 (0.927)	-0.089 (0.919)	-0.399 (0.432)	-0.219 (0.658)
R Hippocampus	-0.069 (0.919)	-0.015 (0.983)	-0.355 (0.468)	-0.070 (0.919)
Social hubs	-0.135 (0.844)	0.328 (0.558)	-0.090 (0.919)	-0.135 (0.844)
L TPJ	-0.049 (0.927)	0.061 (0.927)	-0.083 (0.919)	-0.023 (0.980)
R TPJ	-0.126 (0.865)	0.36 (0.468)	-0.102 (0.919)	-0.151 (0.834)
Non-social hubs	-0.266 (0.608)	0.586 (0.093)	0.357 (0.468)	-0.252 (0.608)
L dmPFC	-0.050 (0.927)	0.184 (0.746)	0.321 (0.558)	-0.247 (0.608)
R dmPFC	-0.237 (0.608)	0.47 (0.288)	0.409 (0.432)	-0.279 (0.608)
L lPFC	0.077 (0.919)	0.078 (0.919)	-0.041 (0.927)	-0.008 (0.983)
R lPFC	-0.256 (0.608)	0.548 (0.105)	0.135 (0.844)	0.006 (0.983)
Connectivity within sets of hubs				
Domain-general hubs	-0.081 (0.919)	0.245 (0.608)	0.182 (0.746)	-0.164 (0.816)
Social hubs	0.394 (0.432)	-0.156 (0.834)	-0.250 (0.608)	0.187 (0.746)
Non-social hubs	-0.265 (0.608)	0.042 (0.927)	0.307 (0.564)	-0.208 (0.691)

Pearson r-values for correlation of hub z-scores with trait measures. FDR-corrected p-values indicated in parentheses.

Table S7. Main Effect of Connectivity on Log-Transformed RT

	B	*SE*	*df*	*t-value*	*p-value*
Node strength					
Domain-general hubs	0.003	0.01	23.915	0.293	0.922
L Hippocampus	0.021	0.011	25.168	1.943	0.856
R Hippocampus	0.006	0.009	24.701	0.629	0.922
Social hubs	0.000	0.011	25.655	0.009	0.993
L TPJ	0.004	0.011	24.827	0.333	0.922
R TPJ	0.004	0.011	25.68	0.353	0.922
Non-social hubs	0.021	0.013	25.262	1.548	0.856
L dmPFC	-0.002	0.013	22.996	-0.129	0.967
R dmPFC	0.012	0.012	25.361	0.960	0.922
L lPFC	-0.009	0.013	25.66	-0.665	0.922
R lPFC	0.003	0.012	25.945	0.268	0.922
Connectivity within sets of hubs					
Domain-general hubs	-0.005	0.013	26.687	-0.393	0.922
Non-social hubs	0.016	0.011	24.092	1.370	0.856
Social hubs	-0.006	0.011	26.114	-0.611	0.922

The p-values are FDR-corrected. Significant results are highlighted in bold.

Table S8. Node Type x Connectivity 2-way Interaction Effect on Log-Transformed RT

	B	*SE*	*df*	*t-value*	*p-value*
Node strength					
Domain-general hubs	0.002	0.003	293.818	0.883	0.529
L Hippocampus	0.001	0.003	139.793	0.497	0.698
R Hippocampus	0.002	0.003	334.954	0.728	0.594
Social hubs	**0.008**	**0.003**	**190.896**	**2.730**	**0.049**
L TPJ	**0.011**	**0.003**	**253.818**	**4.018**	**0.001**
R TPJ	0.006	0.003	178.546	2.069	0.137
Non-social hubs	-0.005	0.003	81.713	-1.847	0.137
L dmPFC	0.001	0.003	196.834	0.458	0.698
R dmPFC	-0.005	0.003	81.596	-1.856	0.137
L lPFC	0.006	0.003	162.694	1.914	0.137
R lPFC	-0.004	0.003	114.873	-1.362	0.308
Connectivity within sets of hubs					
Domain-general hubs	0.003	0.003	125.980	0.921	0.529
Non-social hubs	0.001	0.003	220.468	0.305	0.760
Social hubs	0.007	0.003	153.309	2.231	0.126

The p-values are FDR-corrected. Significant results are highlighted in bold.

Table S9. Node Type x Network Type x Connectivity 3-way Interaction Effect on Log-Transformed RT

	B	*SE*	*df*	*t-value*	*p-value*
Node strength					
Domain-general hubs	0.001	0.003	159.367	0.272	0.846
L Hippocampus	**0.007**	**0.003**	**231.158**	**2.571**	**0.050**
R Hippocampus	0.005	0.003	103.790	1.748	0.233
Social hubs	0.002	0.003	188.560	0.697	0.679
L TPJ	0.000	0.003	108.466	0.152	0.880
R TPJ	0.003	0.003	185.538	0.917	0.679
Non-social hubs	0.002	0.003	209.978	0.646	0.679
L dmPFC	**-0.009**	**0.003**	**191.084**	**-2.859**	**0.033**
R dmPFC	0.002	0.003	211.060	0.591	0.679
L lPFC	-0.007	0.003	155.635	-2.205	0.101
R lPFC	0.002	0.003	152.128	0.551	0.679
Connectivity within sets of hubs					
Domain-general hubs	-0.002	0.003	283.351	-0.834	0.679
Non-social hubs	-0.003	0.003	88.465	-0.977	0.679
Social hubs	**0.008**	**0.003**	**208.910**	**2.898**	**0.033**

The p-values are FDR-corrected. Significant results are highlighted in bold.