

# Transfer of Learning Relates to Intrinsic Connectivity between Hippocampus, Ventromedial Prefrontal Cortex, and Large-Scale Networks

Raphael T. Gerraty,<sup>1\*</sup> Juliet Y. Davidow,<sup>1\*</sup> G. Elliott Wimmer,<sup>2</sup> Itamar Kahn,<sup>3</sup> and Daphna Shohamy<sup>1</sup>

<sup>1</sup>Department of Psychology, Kavli Institute for Brain Science, Columbia University New York, New York 10027, <sup>2</sup>Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany, and <sup>3</sup>Department of Physiology and Biophysics, The Ruth and Bruce Rappaport Faculty of Medicine Technion, Israel Institute of Technology, Haifa 31096, Israel

An important aspect of adaptive learning is the ability to flexibly use past experiences to guide new decisions. When facing a new decision, some people automatically leverage previously learned associations, while others do not. This variability in transfer of learning across individuals has been demonstrated repeatedly and has important implications for understanding adaptive behavior, yet the source of these individual differences remains poorly understood. In particular, it is unknown why such variability in transfer emerges even among homogeneous groups of young healthy participants who do not vary on other learning-related measures. Here we hypothesized that individual differences in the transfer of learning could be related to relatively stable differences in intrinsic brain connectivity, which could constrain how individuals learn. To test this, we obtained a behavioral measure of memory-based transfer outside of the scanner and on a separate day acquired resting-state functional MRI images in 42 participants. We then analyzed connectivity across independent component analysis-derived brain networks during rest, and tested whether intrinsic connectivity in learning-related networks was associated with transfer. We found that individual differences in transfer were related to intrinsic connectivity between the hippocampus and the ventromedial prefrontal cortex, and between these regions and large-scale functional brain networks. Together, the findings demonstrate a novel role for intrinsic brain dynamics in flexible learning-guided behavior, both within a set of functionally specific regions known to be important for learning, as well as between these regions and the default and frontoparietal networks, which are thought to serve more general cognitive functions.

**Key words:** generalization; intrinsic connectivity; learning; networks; rest

## Introduction

Different people learn in different ways; some individuals readily integrate old and new information to support novel inferences, while others fail to transfer what they learn to new situations (Shohamy and Wagner, 2008; van Kesteren et al., 2010; Daw et al., 2011; Wimmer and Shohamy, 2012; Zeithamova et al., 2012). Transfer of learning is highly variable even among healthy populations (Shohamy and Wagner, 2008; Wimmer and Shohamy, 2012). Moreover, transfer is dissociable from other complementary forms of learning, such as learning based on trial-by-trial reinforcement (Shohamy and Wagner, 2008; Daw et al., 2011;

Wimmer and Shohamy, 2012; Doll et al., 2014). Prior studies have shown that variability in transfer is related to differences in task-evoked brain activation during learning (Shohamy and Wagner, 2008; Wimmer and Shohamy, 2012). However, a critical open question is why this variability initially emerges. It is particularly unclear whether differences in activation merely represent transient changes related to task demands, or whether they also reflect intrinsic differences in brain organization.

Flexible learning can be measured using paradigms such as “sensory preconditioning,” where participants first learn associations and then make choices about novel options (Brogden, 1939; Port et al., 1987; Wimmer and Shohamy, 2011, 2012; Jones et al., 2012). Responses to novel options provide an opportunity for spontaneous transfer of initial associations without requiring or rewarding it, capturing individual variability in this tendency. Previous work suggests that transfer may emerge from associative encoding in the hippocampus (Port et al., 1987; Wimmer and Shohamy, 2012) or from inference-based processes in the ventromedial prefrontal cortex (vmPFC; Jones et al., 2012). However, beyond a functional role for each region during learning and decision making, these studies do not address the central question of whether variability in transfer behavior relates to stable variation in functional brain organization.

Received Jan. 14, 2014; revised June 11, 2014; accepted July 11, 2014.

Author contributions: J.Y.D., G.E.W., I.K., and D.S. designed research; J.Y.D. performed research; R.T.G. contributed unpublished reagents/analytic tools; R.T.G., J.Y.D., and G.E.W. analyzed data; R.T.G., J.Y.D., G.E.W., I.K., and D.S. wrote the paper.

This work was funded by National Institute of Neurological Disorders and Stroke Grant R01 078784 to D.S., National Science Foundation (NSF) Early Career Award 0955454 to D.S., and NSF Grant GRFP DGE-11-44155 to J.Y.D. We thank Nina Rouhani for helpful comments on a previous version of this manuscript.

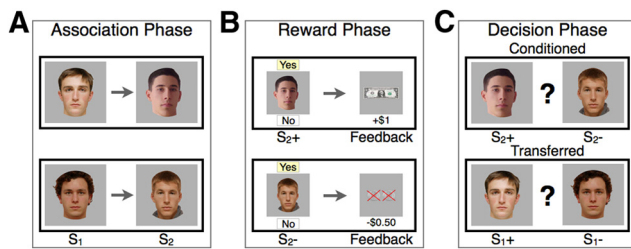
\*R.T.G. and J.Y.D. contributed equally to this work.

The authors declare no competing financial interests.

Correspondence should be addressed to Daphna Shohamy, Columbia University, Psychology Department, 368 Schermerhorn Extension, 1190 Amsterdam Avenue, MC 5501, New York, NY 10027. E-mail: ds2619@columbia.edu.

DOI:10.1523/JNEUROSCI.0185-14.2014

Copyright © 2014 the authors 0270-6474/14/3411297-07\$15.00/0



**Figure 1.** Task design. The learning and transfer task consisted of the following three phases: association phase, reward phase, and decision phase. **A**, In the association phase, participants viewed a series of four pairs of  $S_1$  and  $S_2$  face stimuli while performing a cover task. **B**, In the reward phase, participants learned that  $S_2$  stimuli led to monetary gain ( $S_2+$ ), monetary loss ( $S_2-$ ), or no outcome ( $S_2=$ ) through instrumental conditioning. **C**, In the decision phase, participants chose between pairs of  $S_1$  or pairs of  $S_2$  stimuli. The learning of directly reinforced associations (“conditioning”) was measured as participants’ tendency to choose  $S_2+$  and to avoid  $S_2-$ . Transfer was measured as participants’ tendency to choose  $S_1+$  and to avoid  $S_1-$ , reflecting the transfer of learned value from  $S_2$  to  $S_1$  stimuli.

To answer this question, we used resting-state fMRI (rs-fMRI), which provides a stable measure of distributed brain connectivity. We measured transfer with sensory preconditioning and collected rs-fMRI on a separate day to test whether differences in behavior related to intrinsic connectivity between regions implicated in transfer, including the hippocampus and the vmPFC. rs-fMRI has also been used for characterizing large-scale networks, including the default-mode network (DMN), fronto-parietal network (FPN), and cingulo-opercular network (CON; Power et al., 2011; Yeo et al., 2011). Interestingly, there is overlap between learning-activated regions and a subset of DMN areas—including the hippocampus and the vmPFC, which may compose a distinct subnetwork (Andrews-Hanna et al., 2010; Roy et al., 2012).

Given these findings, we hypothesized that variability in transfer would relate to differences in intrinsic connectivity between the hippocampus and the vmPFC. We additionally hypothesized that these regions’ participation in large-scale functional networks, particularly the DMN, may be related to transfer. Crucially, our task was designed to test whether connectivity effects are selective to transfer versus incremental reinforcement-driven learning measured in the same task.

## Materials and Methods

**Experimental design.** Forty-nine healthy right-handed participants (mean age,  $21.4 \pm 3.3$  years; 30 females) were recruited from Columbia University and the surrounding community. Each participant underwent the following two testing sessions: a behavioral session; and a separate rs-fMRI scan. Informed consent was obtained according to procedures approved by the Columbia University Institutional Review Board. Seven participants were excluded from imaging analysis for excessive motion (see fMRI preprocessing and nuisance regression in Materials and Methods below), leaving a sample of 42 participants. Participants were paid \$20 for behavioral testing, a percentage of their earnings from the task, and \$20 for the MRI scan.

**Task and behavioral analysis.** Participants underwent a behavioral testing session outside the scanner, several days before the brain-imaging session (mean time between sessions,  $55.0 \pm 6.7$  h). Testing was conducted first to ensure that all scanned participants had behavioral data and to screen for MRI contraindications. The transfer paradigm (Fig. 1) was an adapted version of a previously described sensory preconditioning task (Wimmer and Shohamy, 2011, 2012).

The task consists of three phases. In the association phase, participants made button responses as they were exposed to sequential presentations of neutral stimuli pairs (faces, denoted  $S_1$  and  $S_2$ ). A particular  $S_1$  stimulus always preceded its paired  $S_2$  stimulus. Participants performed a

cover task responding to “target” upside-down images, which were aimed at making the encoding incidental. Each of the four pairs was presented 10 times in pseudo-random order, intermixed with 10 target trials for 2 s, with a 2 s interstimulus interval and a 4 s intertrial interval. Participants were not informed of the trial structure.

Next, during the reward learning phase, participants learned to associate  $S_2$  stimuli with a monetary gain ( $S_2+$ ), a monetary loss ( $S_2-$ ), or a neutral outcome [ $S_2=$  (i.e., no win or loss)]. On each trial, an  $S_2$  face was presented, and participants pressed a button to choose to bet or not. The  $S_2+$  stimulus led to a reward of \$1 [80% of trials (16 of 20 trials)] or nothing, while the  $S_2-$  stimulus led to a loss of \$0.50 [80% of trials (16 of 20 trials)] or nothing. The  $S_2=$  stimulus led to neither gain nor loss for all 20 trials. The  $S_2$  stimulus from the fourth pair was not presented during this phase ( $S_2\sim$ ), providing a control stimulus with no value association for the subsequent decision phase. Participants were instructed that they would receive a percentage of their total earnings.

In the critical decision phase, participants were presented with two face stimuli and instructed to choose the face they thought would be more likely to win. Participants were instructed that they would not receive trial-by-trial feedback but that they would be rewarded for a percentage of their choices. Each choice could lead to a win of \$1 or no reward. Stimuli for choice trials were either two  $S_1$  or two  $S_2$  faces. Choices between  $S_1$  stimuli consisted of the following five trial types: positive versus neutral ( $S_1+ - S_1=$ ); positive versus held out ( $S_1+ - S_1\sim$ ); positive versus negative ( $S_1+ - S_1-$ ); negative versus neutral ( $S_1- - S_1=$ ); and negative versus held out ( $S_1- - S_1\sim$ ). Analogous choices were presented for  $S_2$  stimuli. Each choice was presented three times, intermixed in pseudo-random order for 30 total trials.

We computed a transfer score for each participant by averaging the number of  $S_1+$  stimuli chosen or  $S_1-$  stimuli avoided across the five trial types described above, providing a single estimate of individual differences in transfer, while limiting the number of comparisons with the fMRI data. Similar results were obtained using the first principal component across  $S_1$  pairs, which explained 48.9% of the total variance as measured by the ratio of eigenvalues and was highly correlated with the mean. Population-level transfer was compared with chance performance. We used the same analysis strategy for the  $S_2$  choices during the decision phase, allowing a comparison of “conditioned” versus “transfer” responses to test the selectivity of any connectivity associations.

To ensure that participants did not have pre-existing preferences for the stimuli, images were selected based on pretask ratings. Finally, a recognition memory test for  $S_1$ – $S_2$  pairs was administered after the experiment to test whether subjects’ transfer responses were related to explicit awareness of task structure.

**fMRI acquisition.** Participants returned for an rs-fMRI scan that enabled us to assess intrinsic connectivity. We acquired two series of 104 interleaved T2\*-weighted, single-shot, gradient echo, echoplanar functional images on a 3T Philips MRI scanner with an eight-channel head coil. Sequence parameters were as follows: TR, 3000 ms; TE, 30 ms; flip angle, 84°; array, 80 × 80; 41 slices; effective voxel resolution, 3 mm<sup>3</sup>. Participants were told they would be in the scanner for <30 min and to remain still with their eyes open looking at a fixation cross, which was back-projected and visible through a mirror mounted on the head coil. A high-resolution T1-weighted MPRAGE image was acquired for registration.

**fMRI preprocessing and nuisance regression.** Preprocessing of fMRI data was performed using the Functional MRI of the Brain Software Library (www.fmrib.ox.ac.uk/fsl). Brain Extraction Tool (BET) was used to skull-strip anatomical and functional images (Smith, 2002). The first three volumes were removed to account for saturation effects. Functional images were slice-time corrected, motion corrected to the median image using trilinear interpolation with 6 df (Jenkinson et al., 2002), spatially smoothed with a 5 mm full-width at half-maximum Gaussian kernel, grand-mean scaled, and high-pass filtered with frequency  $f > 0.01$  Hz. Functional scans were coregistered with anatomical images and linearly transformed to a standard template (T1 Montreal Neurological Institute template, voxel dimensions of 4 mm<sup>3</sup>) using FLIRT (Jenkinson et al., 2002).

Subject-level functional data were further preprocessed with a nuisance regression using FILM with local autocorrelation implemented in

FEAT (Woolrich et al., 2001). Given the well established effects of movement on functional connectivity (Power et al., 2012; Satterthwaite et al., 2012), we used a recently proposed extended regression technique shown to successfully control for motion artifact in connectivity measures (Satterthwaite et al., 2013). Thirty-six nuisance variables (time series for CSF, white matter, whole-brain, and six motion parameters, the first temporal derivative of each time series, and the square of each time series and its derivative) were regressed against the spatially smoothed and high-pass-filtered data from initial pre-preprocessing. Each model also included single spike regressors for any time point where relative motion was  $>2$  SDs from the global mean across subjects. Any subject with  $>20\%$  of time points in any scanning session “scrubbed” in this manner was excluded from further analysis, resulting in a final sample of 42 of 49 participants. The residual time series from each regression were then low-pass filtered at  $f < 0.08$  Hz. For each subject, the filtered time series from both sessions were concatenated, resulting in 202 time points for each region of interest (ROI) and network of interest.

**Connectivity analyses.** To define regions and networks of interest, we performed a spatial independent component analysis (sICA), implemented in MELODIC. The implementation of probabilistic sICA with MELODIC seeks to maximize the independence of latent spatial components (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MELODIC>). For this analysis, we temporally concatenated all sessions and subjects into a 2D matrix, as this does not assume consistency of temporal patterns across subjects or sessions.

Next, regions from within one of the resulting networks, a putative subcomponent of the default mode network, labeled here as the ventral medial network (VMN), were chosen for preliminary within-network analysis based on their association with transfer in previous reports (Jones et al., 2012; Wimmer and Shohamy, 2012; Zeithamova et al., 2012). The DMN, FPN, and CON (Table 1) were included for between-network analyses examining the relationship between transfer and their connectivity with the hippocampus and vmPFC, based on previous association with the performance of cognitive tasks (Dosenbach et al., 2008; Fornito et al., 2012).

All ROIs and networks were functionally defined based on group-level ICA effects. For connectivity analysis within the ventral medial component, we isolated discrete functional ROIs using anatomical boundaries from the Harvard-Oxford atlas where available. Thus, the hippocampus and striatum ROIs were defined based on overlap between group-level VMN effects and anatomical masks of these regions derived from the Harvard-Oxford atlas, with a threshold probability of 25%. For the vmPFC, given its lack of clear anatomical demarcation, and taking advantage of its robust position in the VMN, the ROI was defined by thresholding the VMN component at  $Z > 7$ , providing a region consistent with those in prior reports (Kumaran et al., 2009; van Kesteren et al., 2010). DMN, FPN, and CON were defined by thresholding their respective component  $Z$  statistics until there was no overlap between the networks. Mean time series were then extracted from all regions and networks for each subject.

Within the VMN, we computed pairwise correlations among five regions: left and right hippocampus, left and right ventral striatum, and vmPFC. The correlation coefficients for each pair were Fisher  $Z$  transformed. Average pairwise connectivity was used to characterize the network structure. We tested the correlation between pairwise connectivity and transfer score for each set of regions (yielding 10 region–region pairs). The resulting correlations were corrected for multiple comparisons using a parametric bootstrap approach, in which we selected randomly from a 10-dimensional normal distribution with the same means and covariances as the VMN connectivity matrices across subjects. These  $42 \times 10$  matrices were generated 10,000 times, with each column then correlated with composite transfer score, to calculate the familywise error (FWE) rate for a given correlation. In addition, to probe the selectivity of any effects to transfer, we correlated pairwise connectivity with first-order conditioning performance, using both a parametric test (Meng et al., 1992) and a nonparametric bootstrap to examine the difference between these overlapping correlations.

Because of their significance for learning-based transfer (see Results), we computed connectivity between both the left hippocampus and

**Table 1. Regions comprising each component included in connectivity analyses**

Region	Size (voxels)	Maximum $Z$	MNI coordinates		
			$x$	$y$	$z$
<b>VMN</b>					
vmPFC/hippocampus/VStr	2003	17.5	−6	42	−12
L frontal pole	17	3.86	−34	10	−36
Posterior cingulate	10	3.46	−6	−54	12
R cerebellum	10	3.63	42	−42	−28
<b>DMN</b>					
L angular gyrus	1214	9.04	−54	−62	20
R angular gyrus	1103	8.04	58	−58	20
Dorsal medial prefrontal cortex	769	9.84	2	54	24
Posterior cingulate cortex	150	6.66	2	−54	32
L cerebellum	71	6.1	−26	−78	−36
R cerebellum	63	5.28	26	−78	−36
L middle frontal gyrus	39	5.04	−42	10	48
R middle frontal gyrus	17	3.48	13	35	26
Frontal pole	15	4.17	2	62	−12
<b>L FPN</b>					
L middle/inferior frontal gyrus	1963	14.8	−50	14	32
L superior lateral occipital cortex	899	13	−34	−70	44
L middle temporal gyrus	434	9.74	−58	−46	−12
R cerebellum	359	8.21	34	−70	−44
R superior lateral occipital cortex	81	5.37	34	−66	44
R middle/inferior frontal gyrus	60	4.63	46	30	16
R middle temporal gyrus	38	5.05	8	21	14
<b>R FPN</b>					
R middle/inferior frontal gyrus	2303	11.5	46	22	36
R superior lateral occipital cortex	1096	13.3	46	−54	48
L cerebellum	461	10	−38	−70	−40
R middle temporal gyrus	285	7.52	62	−46	−12
L superior lateral occipital cortex	140	4.71	33	18	31
R thalamus	19	4.07	19	31	21
<b>CON</b>					
dACC/frontal pole/L anterior insula	3003	13.1	−26	50	20
R anterior insula	126	5.58	50	14	−4
R cerebellum	121	6.41	38	−58	−32
L cerebellum	80	6.17	−38	−58	−32
L angular gyrus	35	4.26	−58	−50	32
R caudate	35	4.37	14	22	0
R angular gyrus	26	4.47	62	−42	28

Components were thresholded at  $Z > 3.1$ , extent  $> 10$  for inclusion in this table. Coordinates are reported in MNI space. L, left hemisphere; R, right hemisphere; dACC, dorsal anterior cingulate cortex; VStr, ventral striatum.

vmPFC and the following three intrinsic connectivity networks: DMN, FPN, and CON. These correlations were Fisher  $Z$  transformed and were included in separate linear models for each intrinsic network, predicting the transfer score. To isolate the individual effects of the network participation of each region, regressors for each of these three models included connectivity between the hippocampus and the network of interest, and connectivity between the vmPFC and the network of interest, as well as a nuisance regressor of hippocampus–vmPFC connectivity.

## Results

The sICA revealed a number of components reflecting well characterized intrinsic connectivity networks, consistent with previous reports (Beckmann et al., 2005; Zuo et al., 2010; Fornito et al., 2012). Supporting the notion that the hippocampus, vmPFC, and ventral striatum compose a distinct subcomponent of the default mode network, our sICA analysis estimated this VMN as a separate component (Fig. 2A). Correlation analyses revealed significant connectivity between the vmPFC and all other regions extracted from the VMN. We also found significant interhemispheric connectivity for both the hippocampus and ventral striatum. These results allowed us to test whether connectivity within this network related to transfer tested on a separate day.

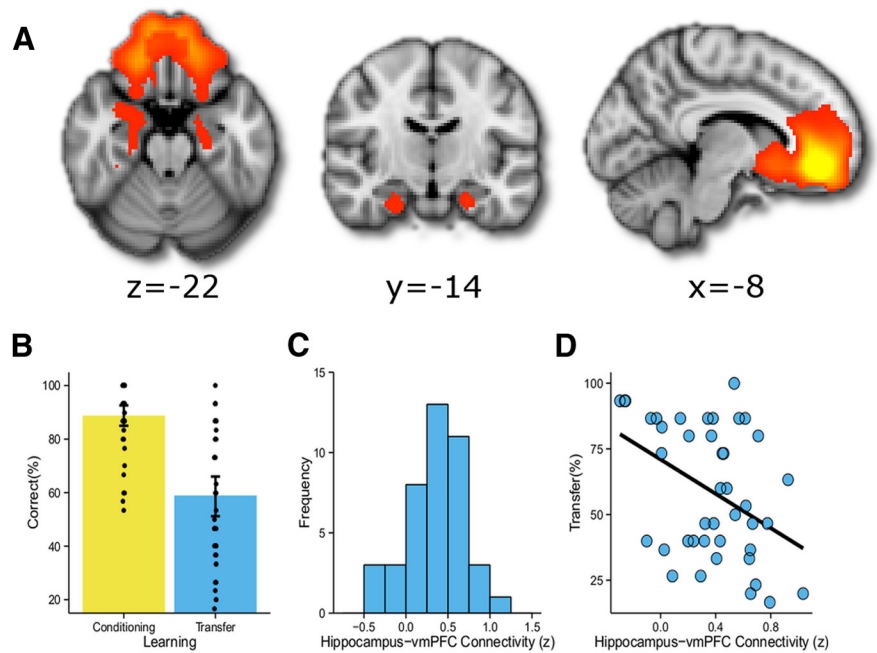
On the separate behavioral session, participants showed robust learning of direct  $S_2$ –outcome associations trained during

the reward phase of sensory preconditioning, evidenced by their average tendency to choose the  $S_2+$  and avoid the  $S_2-$  stimuli during the decision phase [ $88.8 \pm 2.0\%$  (mean  $\pm$  SE),  $t_{(41)} = 19.04$ ,  $p < 0.0001$ ; Fig. 2B]. Average transfer scores were significantly greater than chance ( $58.9 \pm 3.9\%$ ,  $t_{(41)} = 2.3$ ,  $p = 0.01$ , one-tailed; Fig. 2B). A recognition memory test revealed no significant explicit memory for the  $S_1$ – $S_2$  associations (mean accuracy,  $22.6 \pm 3.5\%$ ; not different than chance, 25%), and the correlation between explicit memory for these pairs and transfer did not differ from chance ( $r_{(40)} = 0.11$ ,  $p = 0.49$ ). These findings replicated prior results (Wimmer and Shohamy, 2011, 2012), indicating that any transfer in the current study was implicit and not the result of strategic reasoning. Additionally, transfer was orthogonal to first-order conditioning ( $r_{(40)} = 0.02$ ,  $p = 0.93$ ). As anticipated, transfer was highly variable across participants, ranging from 16.7% to 100% (Fig. 2B); thus, some participants generalized all learned associations, while others showed little to no transfer.

We next tested the central question of whether transfer behavior relates to variability in intrinsic network connectivity (Fig. 2C). Of the 10 correlations among VMN regions (Table 2), only vmPFC connectivity with the left hippocampus showed a significant transfer correlation ( $r_{(40)} = -0.42$ ,  $p = 0.02$  FWE corrected). This correlation was negative, with lower connectivity related to higher transfer (Fig. 2D). A trend in the same direction was found for vmPFC connectivity with the right hippocampus ( $r_{(40)} = -0.29$ ,  $p = 0.06$ , uncorrected), indicating that this effect is not lateralized.

We then tested whether connectivity among the vmPFC, hippocampus, and large-scale networks of interest (Table 1; Figs. 3, 4) was related to transfer. Connectivity between the core DMN and both the hippocampus and the vmPFC was significantly related to transfer (Fig. 3; DMN connectivity with the hippocampus:  $\beta = -0.48$ ,  $p = 0.0002$ ; DMN connectivity with vmPFC:  $\beta = 0.26$ ,  $p = 0.02$ ; adjusted  $r^2 = 0.32$ ,  $p = 0.0002$ ). Connectivity between the hippocampus and the DMN was negatively associated with transfer, while connectivity between the vmPFC and the DMN was positively associated with transfer. Both effects remained significant when including hippocampus–vmPFC connectivity in the model as a nuisance covariate. In keeping with the position of the vmPFC as a network hub (Buckner et al., 2008; Roy et al., 2012; Sreenivas et al., 2012), its connectivity with the FPN was also significantly negatively associated with transfer (Fig. 4;  $\beta = -0.38$ ,  $p = 0.02$ , adjusted  $r^2 = 0.20$ ). This effect remained significant when including hippocampus–vmPFC connectivity as a nuisance covariate. There were no significant effects of connectivity between the cingulo-opercular network and either the hippocampus or vmPFC.

Finally, we tested the selectivity of the above effects to transfer. To control for the possibility that this pattern of results broadly reflects attentional or learning tendencies, we repeated the analyses using each participant's conditioning score (i.e., the propor-



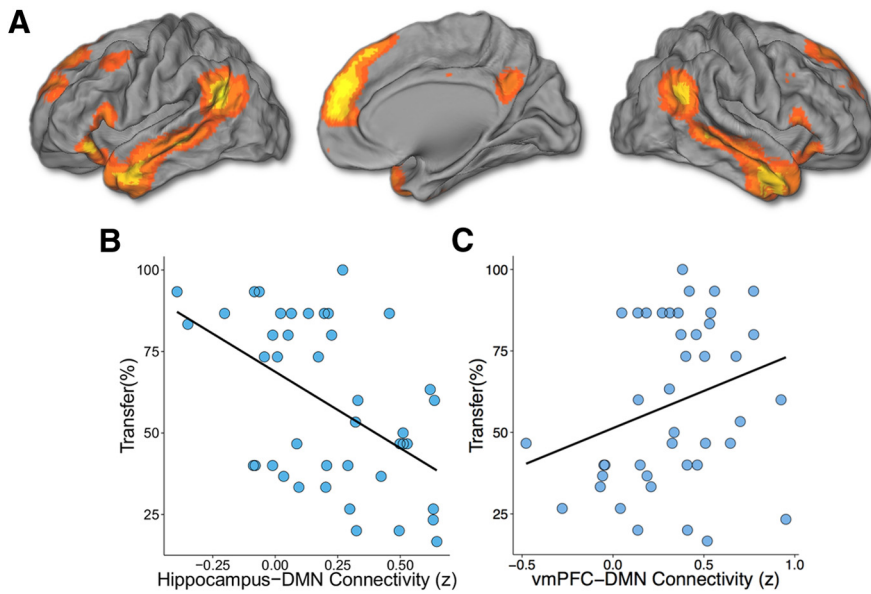
**Figure 2.** Connectivity between the hippocampus and the vmPFC negatively correlates with transfer. **A**, Ventral medial network. Spatial ICA revealed an intrinsic functional connectivity network consisting of the ventromedial prefrontal cortex, bilateral striatum, and bilateral hippocampus. Interindividual variability in pairwise connectivity between regions of this VMN allowed us to test the relationship between within-network connectivity and value transfer. Images thresholded at  $Z < 3.1$  and resampled to  $2 \text{ mm}^3$  for display purposes, and are shown in radiological convention. **B**, Behavioral performance during the decision phase. Average scores for first-order conditioning and transfer were both significantly above chance. Subjects exhibited a large amount of variability in transfer behavior. Bars represent means and 95% CIs, and points show individual subject scores. **C**, The Z-transformed distribution of correlation between the left hippocampus and vmPFC across participants. The y-axis represents the number of participants. **D**, Connectivity between the left hippocampus and vmPFC is significantly negatively correlated with transfer. A trend in the same direction was found for the right hippocampus (see text and Table 2).

**Table 2. Correlations between behavioral transfer and intrinsic functional connectivity between ROIs**

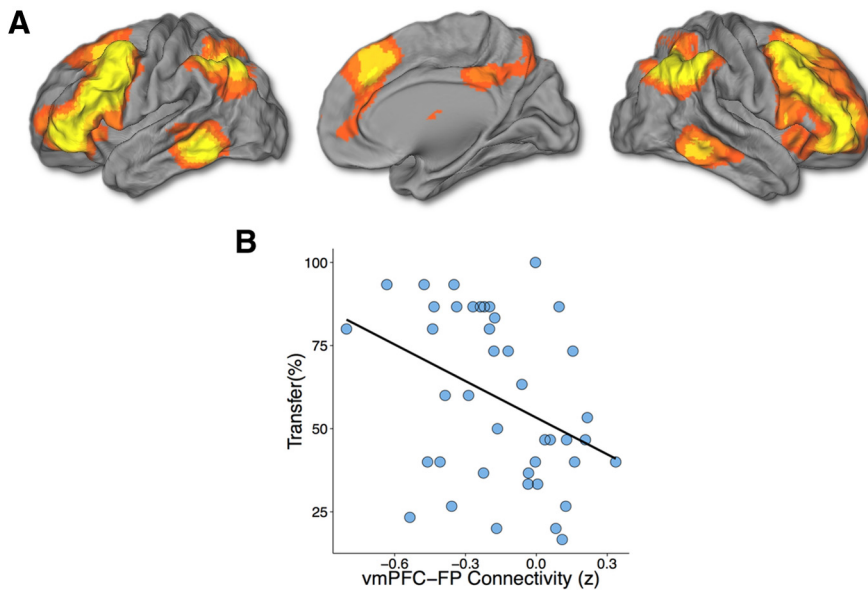
	Hippocampus		Caudate		
	vmPFC	Right	Left	Right	Left
vmPFC		–0.29	–0.42*	0.02	0.09
Hippocampus					
Right	–0.29		–0.18	–0.06	–0.18
Left	–0.42*	–0.18		–0.07	–0.21
Caudate					
Right	0.02	–0.06	–0.07		0.29
Left	0.09	–0.18	–0.21	0.29	

Values in each cell represent the correlation of transfer with intrinsic functional connectivity between the ROIs listed in the corresponding row and column. Only connectivity between the vmPFC and the left hippocampus ROIs is significantly related to transfer behavior,  $r_{(40)} = -0.42$ ,  $p_{\text{FWE}} < 0.05$ .

tion of choosing  $S_2+$  and avoiding  $S_2-$ ), which reflects direct learning of stimulus–outcome associations from the reward learning phase. Importantly, the correlation between hippocampus–vmPFC connectivity and learning was selective to transfer, and was not found for conditioning ( $r_{(40)} = -0.02$ ,  $p = 0.90$ , uncorrected). Moreover, a direct comparison of the two overlapping correlations (Meng et al., 1992) was significant ( $Z = -1.85$ ,  $p = 0.03$ , one-tailed); this difference was also significant when tested using a nonparametric bootstrap to account for non-normality of behavioral performance. These results provide evidence for the specific contributions of this network to transfer. Associations with between-network connectivity were also specific, as follows: conditioning did not relate to DMN connectivity with either the vmPFC or the hippocampus (hippocampus:  $\beta = -0.02$ ,  $p = 0.85$ ; vmPFC:  $\beta = -0.01$ ,  $p = 0.85$ ; adjusted  $r^2 =$



**Figure 3.** Intrinsic DMN connectivity with the vmPFC and the hippocampus is associated with transfer. **A**, DMN as revealed by spatial ICA. **B**, **C**, Intrinsic functional connectivity between the DMN and the hippocampus (**B**) and the vmPFC (**C**) was significantly related to transfer of learned value. Regression lines are from linear models containing DMN–hippocampus and DMN–vmPFC connectivity as predictors. Both effects remained significant when including vmPFC–hippocampus connectivity as a covariate, and both effects are selective to transfer.



**Figure 4.** vmPFC connectivity with the frontoparietal network is negatively associated with transfer. **A**, Frontoparietal network. **B**, Plot showing negative relationship between value transfer and vmPFC–FPN connectivity. The regression line was computed from a linear model containing FPN–hippocampus and FPN–vmPFC connectivity as predictors. This effect remained significant when including vmPFC–hippocampus connectivity as a covariate and is specific to transfer.

$-0.05$ ,  $p = 0.96$ ). The same was true for the effect of vmPFC–FPN connectivity on learning ( $\beta = -0.06$ ,  $p = 0.52$ , adjusted  $r^2 = -0.06$ ).

## Discussion

Individuals vary in the extent to which they leverage past learning when faced with novel choices. The sensory preconditioning paradigm provides a unique opportunity to explore the stable neural basis of variability in the tendency to transfer spontaneously, even in contexts where it provides no explicit benefit. Here we

find that individual variability in the transfer of learning is related to intrinsic connectivity between the hippocampus and the vmPFC, measured during rest and on a separate day. Further, we find that individual differences in this behavior are also related to overall connectivity between these regions and the DMN and FPN. While multiple studies have demonstrated a relationship between task-evoked activation and differences in transfer (Shohamy and Wagner, 2008; Wimmer and Shohamy, 2012; Wimmer et al., 2012; Zeithamova et al., 2012), these studies do not address the potential role of intrinsic functional brain organization. Based on these studies, variability in transfer could have related solely to task-evoked patterns of brain activity, and not to intrinsic connectivity. The finding of a link between these two measures advances our understanding of the role of network connectivity in behavioral variability, and suggests that stable and intrinsic aspects of functional brain organization can explain individual differences in the transfer of learning.

These results thus demonstrate a novel link between intrinsic connectivity and complex measures of learning. Further, their specificity to transfer of value as opposed to first-order conditioning indicates a relationship between resting-connectivity and behavior that is particular to regions involved in a given domain. While we used sICA to functionally isolate intrinsic networks, the selection of VMN regions was based on the a priori hypothesis that connectivity between regions known to be involved in transfer would be related to variability in this behavior. While much rs-fMRI research has focused on commonalities in normal populations, here we exploit the fact that individual connectivity is itself highly variable in healthy adults (Kahn and Shohamy, 2013) and show that this variability has implications for learning behavior.

Previous studies have separately highlighted roles for the hippocampus and the vmPFC in a wide range of behaviors that involve memory-guided learning (Port et al., 1987; Dusek and Eichenbaum, 1997; Myers et al., 2003; Greene et al., 2006; Kumaran et al., 2009; van Kesteren et al., 2010; Schoenbaum et al., 2011; Wimmer and Shohamy, 2012; Zeithamova et al., 2012). Together, these regions are thought to contribute to the formation and use of flexible representations that guide transfer behavior. In the sensory preconditioning paradigm, we previously found in humans that hippocampus activation during learning is related to subsequent transfer (Wimmer and Shohamy, 2012). Jones et al. (2012) separately demonstrated the necessary role of a homolog to human

medial orbitofrontal cortex in transfer behavior using the same paradigm in rodents. The current results suggest that the strength of connectivity between the hippocampus and vmPFC at rest explains a significant portion of the variability in this kind of transfer and links across these pieces of evidence in different species performing the same task.

The current results indicate that the correlation between hippocampal–vmPFC connectivity and transfer is negative, such that participants transferring the most are those with the weakest connectivity. The direction of this effect is consistent with a recent proposal about how connections between the hippocampus and the vmPFC relate to the incorporation of new memories into existing abstract frameworks, called “schemas.” Schemas also capture an element of flexible memory-guided behavior, and studies show that the extent to which an individual will generalize a schema to new events is negatively correlated with hippocampal–vmPFC connectivity (van Kesteren et al., 2010). Such findings are consistent with neurobiological evidence of dynamic changes in hippocampal–prefrontal circuits during learning (Doyère et al., 1993; Takita et al., 1999) and have led to the proposal that the vmPFC may serve as a gateway to hippocampal learning (van Kesteren et al., 2012). Given the convergence between our results and previous findings, it is possible that the negative relationship we show reflects an overlapping learning mechanism to that proposed by van Kesteren et al. (2012). If prefrontal gating supports multiple forms of transfer, intrinsic connectivity between the hippocampus and vmPFC could be a marker of the tendency to transfer learning across a range of tasks. If so, then it would be expected that groups who show decreased transfer, such as patients with schizophrenia (Ivleva et al., 2012), may actually display increased vmPFC–hippocampus connectivity at rest.

Sensory preconditioning provides a measure of implicit transfer of value across learned associations. The findings we report are broadly relevant for other classes of paradigms that test transfer or generalization of knowledge (for review, see Kumaran and McClelland, 2012), many of which involve generalization of stimulus–stimulus associations outside the domain of reward learning. Yet other related processes have been examined within the context of reward-based decision making, in studies of the neural and cognitive mechanisms underlying learning about structured regularities in the environment, referred to as “model-based” reinforcement learning (Daw et al., 2005, 2011; Hampton et al., 2006). This type of learning relies on prospective models of environmental contingencies that can flexibly guide decision making, as opposed to the rigid and feedback-dependent updating of action or stimulus value, referred to as “model-free” learning (for review, see Doll et al., 2012). Current evidence suggests that there is significant individual variability in the respective weights placed on these types of learning (Daw et al., 2011). Although memory-based generalization and model-based frameworks vary in their origins, there is evidence for shared behavioral mechanisms (Doll et al., 2014). Furthermore, emerging findings across these domains reveal an important role for the hippocampus and vmPFC (McDannald et al., 2011; Zeithamova et al., 2012; Bornstein and Daw, 2013). Together with the current results, this suggests that variability in intrinsic functional connectivity between the hippocampus and the vmPFC may relate to a broader range of stable learning behaviors that involve a common mechanism.

In the present study, we purposefully measure intrinsic connectivity in the absence of a task to address questions regarding variability in transfer behavior regardless of task-evoked activity.

In general, the link between activation evoked by task demands and spontaneous fluctuations remains an open question. Emerging evidence suggests that resting and evoked activity may negatively interact (He, 2013). Such a tradeoff between baseline and event-related connectivity presents a challenge to drawing inferences between these two conditions. These questions are important for interpreting studies of intrinsic connectivity, and should be tested in future studies that directly compare evoked and baseline conditions.

Learning and memory research has historically focused on distinguishing the roles of discrete and functionally specific nodes. The present findings indicate an important but unexplored role for large-scale functional networks in complex learning behavior. Our results highlight the relationship between learning and intrinsic connectivity within a ventral-medial component of the default network centered on the hippocampus and the vmPFC, regions known to be involved in a host of flexible learning behaviors. In addition, we demonstrate that broader connectivity between these regions and the DMN and FPN, whose functions are currently much less clear, also relates to transfer. Together, these findings raise the possibility that communication between localized brain regions involved in specific cognitive functions and distributed domain-general networks may play an important role in learning. These results also suggest that this approach has potential implications for the use of rs-fMRI to further our understanding of individual differences in a range of basic learning behaviors. Because variation in these processes is likely to underlie complex behavioral traits, these results hold additional promise for answering questions about stable differences in complex behavior that are central to research in education as well as studies of clinical populations.

## References

- Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL (2010) Functional-anatomic fractionation of the brain's default network. *Neuron* 65:550–562. [CrossRef Medline](#)
- Beckmann CF, DeLuca M, Devlin JT, Smith SM (2005) Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc Lond B Biol Sci* 360:1001–1013. [CrossRef Medline](#)
- Bornstein AM, Daw ND (2013) Cortical and hippocampal correlates of deliberation during model-based decisions for rewards in humans. *PLoS Comput Biol* 9:e1003387. [CrossRef Medline](#)
- Brogden WJ (1939) Sensory pre-conditioning. *J Experimental Psychology* 25:9.
- Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: anatomy, function, and relevance to disease. *Ann N Y Acad Sci* 1124:1–38. [CrossRef Medline](#)
- Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8:1704–1711. [CrossRef Medline](#)
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69:1204–1215. [CrossRef Medline](#)
- Doll BB, Simon DA, Daw ND (2012) The ubiquity of model-based reinforcement learning. *Curr Opin Neurobiol* 22:1075–1081. [CrossRef Medline](#)
- Doll BB, Shohamy D, Daw N (2014) Multiple memory systems as substrates for multiple decision systems. *Neurobiol Learn Mem*. Advance online publication. Retrieved July 17, 2014. doi:10.1016/j.nlm.2014.04.014. [CrossRef](#)
- Dosenbach NU, Fair DA, Cohen AL, Schlaggar BL, Petersen SE (2008) A dual-networks architecture of top-down control. *Trends in Cognitive Sciences* 12:99–105. [CrossRef Medline](#)
- Doyère V, Burette F, Negro CR, Laroche S (1993) Long-term potentiation of hippocampal afferents and efferents to prefrontal cortex: implications for associative learning. *Neuropsychologia* 31:1031–1053. [CrossRef Medline](#)
- Dusek JA, Eichenbaum H (1997) The hippocampus and memory for orderly

- stimulus relations. *Proc Natl Acad Sci U S A* 94:7109–7114. [CrossRef Medline](#)
- Fornito A, Harrison BJ, Zalesky A, Simons JS (2012) Competitive and cooperative dynamics of large-scale brain functional networks supporting recollection. *Proc Natl Acad Sci U S A* 109:12788–12793. [CrossRef Medline](#)
- Greene AJ, Gross WL, Elsinger CL, Rao SM (2006) An fMRI analysis of the human hippocampus: inference, context, and task awareness. *J Cogn Neurosci* 18:1156–1173. [CrossRef Medline](#)
- Hampton AN, Bossaerts P, O'Doherty JP (2006) The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci* 26:8360–8367. [CrossRef Medline](#)
- He BJ (2013) Spontaneous and task-evoked brain activity negatively interact. *J Neurosci* 33:4672–4682. [CrossRef Medline](#)
- Ivleva EI, Shohamy D, Mihalakos P, Morris DW, Carmody T, Tamminga CA (2012) Memory generalization is selectively altered in the psychosis dimension. *Schizophr Res* 138:74–80. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez A, Mirenzi A, Schoenbaum G (2012) Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science* 338:953–956. [CrossRef Medline](#)
- Kahn I, Shohamy D (2013) Intrinsic connectivity between the hippocampus, nucleus accumbens, and ventral tegmental area in humans. *Hippocampus* 23:187–192. [CrossRef Medline](#)
- Kumaran D, McClelland JL (2012) Generalization through the recurrent interaction of episodic memories: a model of the hippocampal system. *Psychol Rev* 119:573–616. [CrossRef Medline](#)
- Kumaran D, Summerfield JJ, Hassabis D, Maguire EA (2009) Tracking the emergence of conceptual knowledge during human decision making. *Neuron* 63:889–901. [CrossRef Medline](#)
- McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G (2011) Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *J Neurosci* 31:2700–2705. [CrossRef Medline](#)
- Meng X-I, Rosenthal R, Rubin DB (1992) Comparing correlated correlation coefficients. *Psychol Bull* 111:172–175. [CrossRef](#)
- Myers CE, Shohamy D, Gluck MA, Grossman S, Kluger A, Ferris S, Golomb J, Schnirman G, Schwartz R (2003) Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *J Cogn Neurosci* 15:185–193. [CrossRef Medline](#)
- Port RL, Beggs AL, Patterson MM (1987) Hippocampal substrate of sensory associations. *Physiol Behav* 39:643–647. [CrossRef Medline](#)
- Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE (2011) Functional network organization of the human brain. *Neuron* 72:665–678. [CrossRef Medline](#)
- Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59:2142–2154. [CrossRef Medline](#)
- Roy M, Shohamy D, Wager TD (2012) Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends Cogn Sci* 16:147–156. [CrossRef Medline](#)
- Satterthwaite TD, Wolf DH, Loughhead J, Ruparel K, Elliott MA, Hakonarson H, Gur RC, Gur RE (2012) Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. *Neuroimage* 60:623–632. [CrossRef Medline](#)
- Satterthwaite TD, Elliott MA, Gerraty RT, Ruparel K, Loughhead J, Calkins ME, Eickhoff SB, Hakonarson H, Gur RC, Gur RE, Wolf DH (2013) An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage* 64:240–256. [CrossRef Medline](#)
- Schoenbaum G, Takahashi Y, Liu TL, McDannald MA (2011) Does the orbitofrontal cortex signal value? *Ann N Y Acad Sci* 1239:87–99. [CrossRef Medline](#)
- Shohamy D, Wagner AD (2008) Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron* 60:378–389. [CrossRef Medline](#)
- Smith SM (2002) Fast robust automated brain extraction. *Hum Brain Mapp* 17:143–155. [CrossRef Medline](#)
- Sreenivas S, Boehm SG, Linden DE (2012) Emotional faces and the default mode network. *Neurosci Lett* 506:229–234. [CrossRef Medline](#)
- Takita M, Izaki Y, Jay TM, Kaneko H, Suzuki SS (1999) Induction of stable long-term depression in vivo in the hippocampal-prefrontal cortex pathway. *Eur J Neurosci* 11:4145–4148. [CrossRef Medline](#)
- van Kesteren MT, Fernández G, Norris DG, Hermans EJ (2010) Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proc Natl Acad Sci U S A* 107:7550–7555. [CrossRef Medline](#)
- van Kesteren MT, Ruiters DJ, Fernández G, Henson RN (2012) How schema and novelty augment memory formation. *Trends Neurosci* 35:211–219. [CrossRef Medline](#)
- Wimmer GE, Shohamy D (2011) The striatum and beyond: contributions of the hippocampus to decision making. In: *Decision making, affect, and learning: attention and performance XXIII* (Delgado MR, Phelps EA, Robbins TW, eds), pp 281–310. Oxford, New York: Oxford UP.
- Wimmer GE, Shohamy D (2012) Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science* 338:270–273. [CrossRef Medline](#)
- Wimmer GE, Daw ND, Shohamy D (2012) Generalization of value in reinforcement learning by humans. *Eur J Neurosci* 35:1092–1104. [CrossRef Medline](#)
- Woolrich MW, Ripley BD, Brady M, Smith SM (2001) Temporal autocorrelation in univariate linear modeling of fMRI data. *Neuroimage* 14:1370–1386. [CrossRef Medline](#)
- Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106:1125–1165. [CrossRef Medline](#)
- Zeithamova D, Dominick AL, Preston AR (2012) Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron* 75:168–179. [CrossRef Medline](#)
- Zuo XN, Kelly C, Adelstein JS, Klein DF, Castellanos FX, Milham MP (2010) Reliable intrinsic connectivity networks: test-retest evaluation using ICA and dual regression approach. *Neuroimage* 49:2163–2177. [CrossRef Medline](#)