Supplementary Material:

Additional exclusion criteria for each sample:

1) Younger than 18 years.
2) Impaired hearing.
3) Pregnancy.
4) Head trauma with loss of consciousness in the last year or any evidence of functional impairment due to and persisting after head trauma.
5) Having a known risk from exposure to high magnetic fields (e.g. having pace makers) and having metallic implants (e.g. braces) in the head region (likely to create artifact on the MRI scans).
6) Inability to give informed consent
7) IQ < 80
8) A history of substance abuse within the past 6 months
9) Any significant neurological and medical conditions revealed by clinical and MRI evaluation.
10) Presence of psychotic disorders in first-degree relatives of healthy subjects, assessed by using the Family Interview for Genetic Studies (FIGS; Maxwell, 1992).

DNS BOLD fMRI Data Acquisition

Each participant was scanned using one of two identical research-dedicated GE MR750 3T scanners equipped with high-power high-duty-cycle 50-mT/m gradients at 200 T/m/s slew rate, and an eight-channel head coil for parallel imaging at high bandwidth up to 1MHz at the Duke-UNC Brain Imaging and Analysis Center. A semi-automated high-order shimming program was used to ensure global field homogeneity. A series of 34 interleaved axial functional slices aligned with the anterior commissure-posterior commissure plane were acquired for full-brain coverage using an inverse-spiral pulse sequence to reduce susceptibility artifacts (TR/TE/flip angle=2000 ms/30 ms/60; FOV=240mm; 3.75×3.75×4mm voxels; interslice skip=0). Four initial radiofrequency excitations were performed (and discarded) to achieve steady-state equilibrium. To allow for spatial registration of each participant's data to a standard coordinate system, high-resolution three-dimensional T1-weighted structural images were obtained in 162 axial slices using a 3D Ax FSPGR BRAVO sequence (TR/TE/flip angle = 8.148 ms / 3.22 ms / 12°; voxel size=0.9375x0.9375x1mm; FOV=240mm; interslice skip=0; total scan time = 4 min and 13 s). In addition, high-resolution structural images were acquired in 34 axial slices coplanar with the functional scans and used for spatial registration for participants without Ax FSPGR BRAVO images (TR/TE/flip angle=7.7 s/3.0 ms/12; voxel size=0.9×0.9×4mm; FOV=240mm, interslice skip=0).

DNS BOLD fMRI Data Pre-Processing

Anatomical images for each subject were skull-stripped, intensity-normalized, and nonlinearly warped to a study-specific average template in a standard stereotactic space (Montreal Neurological Institute template) using ANTs (Klein *et al.*, 2009). BOLD time series for each

subject were processed in AFNI (Cox, 1996). Images for each subject were despiked, slice-time-corrected, realigned to the first volume in the time series to correct for head motion, co-registered to the anatomical image using FSL's Boundary Based Registration (Greve and Fischl, 2009), spatially normalized into MNI space using the non-linear warp from the anatomical image, resampled to 2mm isotropic voxels, and smoothed to minimize noise and residual difference in gyral anatomy with a Gaussian filter, set at 6-mm full-width at half-maximum. All transformations were concatenated so that a single interpolation was performed. Voxel-wise signal intensities were scaled to yield a time series mean of 100 for each voxel. Volumes exceeding 0.5mm frame-wise displacement (FD) or 2.5 standardized DVARS (Power *et al.*, 2014; Nichols, 2017) were censored from the GLM.

DNS fMRI Quality Assurance Criteria

Quality control criteria for inclusion of a participant's imaging data were: >5 volumes for each condition of interest retained after censoring for FD and DVARS and sufficient temporal SNR within the bilateral AAL hippocampus ROI, defined as greater than 3 standard deviations below the mean of this value across subjects. Additionally, data were only included in further analyses if the participant demonstrated sufficient engagement with the task, defined as 66% accuracy recalling the names.

d'

During both the encoding and retrieval phases of the SDMT, button box inputs were logged in order to monitor reaction time and response accuracy. In particular, responses during the retrieval portion of the task were coded according to four categories. For images that the participant had previously viewed during encoding, right and wrong responses were coded as "correct

recognitions" and "failed recognitions" respectively. For novel images that were not displayed during encoding, right and wrong responses were coded as "correct rejections" and "false alarms".

	Response: yes	**Response: no**
Viewed: yes	correct recognitions	failed recognitions
Viewed: no	false alarms	correct rejections

Using these four categories, d' was then calculated as a measure of retrieval accuracy as follows:

$$d' = Z(p \text{ correct recognition}) - Z(p \text{ false alarm})$$

where function $Z(p)$, $p \in [0,1]$, is the inverse of the cumulative distribution function of the Gaussian distribution.

Adjusted PRS:

Adjusted PRS was calculated as residual of the following regression model in *R* 3.02 (*R*, *http://www.r-project.org/*). The purpose of this procedure was to remove confounding effects of genotyping batch and population stratification from the original PRS (Schizophrenia Working Group of the Psychiatric Genomics, 2014). These covariates are confounding factors in calculating PRS, whereas they are not confounding factors in demographical data and neuroimaging data. Instead of using these factors as covariates, we adjusted PRS by regressing out (projecting out) these covariates before doing further correlation analyses for demographic variables, and regression analyses for neuroimaging data.

$$fit0\ (reduced\ model): ROI = Batch + PC1 + PC2 + \cdots$$

$$adjusted\_PRS = residual(fit0)$$

Percent variance in BOLD-fMRI activation explained by PRS

For each fMRI analysis described in the main manuscript, averaged ROI signals were extracted from SPM and evaluated in *R* 3.02 (*R*, *http://www.r-project.org/*) to calculate the percent variance of BOLD-fMRI activation explained by PRS. ROI signal was defined as unadjusted contrast beta values extracted from a cluster including peak within right posterior hippocampus under P<0.05 in statistical map of each PRS regression analysis in SPM. In R, association of ROI signal and each adjusted RPS was tested using a regression model covaried with sex and age. Difference of R-square values between full model and reduced model were reported as the percent variance explained.

$fit1\ (full\ model)$: $ROI = adjusted\_PRS + Sex + Age$

$fit2\ (reduced\ model)$: $ROI = Sex + Age$

$R^2 = R^2(fit1) - R^2(fit2)$

Association between other PRS's and hippocampal activation for Lieber sample

We analyzed the relationship between hippocampal activation during neutral encoding and the different PRS, calculated using sub-sets of the SNPs under different thresholds based on the PGC2 GWAS p-values of association with schizophrenia: 5e-08 (PRS1), 1e-06 (PRS2), 1e-04 (PRS3), 0.001 (PRS4), 0.01 (PRS5), 0.05 (PRS6; NB: results presented in the main text)), 0.1 (PRS7), 0.2 (PRS8), 0.5 (PRS9), and 1 (PRS10). We found that the correlation between PRS and hippocampal activation increased from PRS1 to PRS6, indicating that the PRS calculated using the larger number of SNPs with less significant association with schizophrenia in PGC2 GWAS but with putatively more true positive loci show more significant association with hippocampal activation during memory encoding, compared with PRS calculated using a smaller number of

SNPs and putatively fewer true positive associated loci, even those individually more significantly associated with schizophrenia (Figure S2 and Table S4). Interestingly, this relationship is analogous to the relationship between PRS and schizophrenia diagnosis liability identified in the clinical GWAS (Schizophrenia Working Group of the Psychiatric Genomics, 2014).

Percent variance of BOLD-fMRI activation explained by PRS, sex, age, IQ, and d'

Percent variance of BOLD-fMRI activation at hippocampus explained by PRS, sex, age, IQ and d' was evaluated in *R* 3.02 (*R*, *http://www.r-project.org/*). In R, association of ROI signal and each variable was tested using stepwise regression model. Differences of R-square values between full model and reduced model were reported as the percent variance explained and calculated as the following formulas.

$$fit1\ (full\ model): ROI = adjusted\_PRS + Sex + Age + IQ + d'$$

$$fit2\ (reduced\ model): ROI = Sex + Age + IQ + d'$$

$$Variance\ explained\ by\ PRS: R^2 = R^2(fit1) - R^2(fit2)$$

$$fit3\ (reduced\ model): ROI = adjusted\_PRS + Age + IQ + d'$$

$$Variance\ explained\ by\ Sex: R^2 = R^2(fit1) - R^2(fit3)$$

$$fit4\ (reduced\ model): ROI = adjusted\_PRS + Sex + IQ + d'$$

$$Variance\ explained\ by\ Age: R^2 = R^2(fit1) - R^2(fit4)$$

$fit5\ (reduced\ model)$: $ROI = adjusted\_PRS + Sex + Age + d'$

$Variance\ explained\ by\ IQ$: $R^2 = R^2(fit1) - R^2(fit5)$

$fit6\ (reduced\ model)$: $ROI = adjusted\_PRS + Sex + Age + IQ$

$Variance\ explained\ by\ d'$: $R^2 = R^2(fit1) - R^2(fit6)$

Table S1. Demographic and performance data of the Italian (Bari) sample with statistics of the relationship between each variable and PRS.

	Sex (M/F)	**Age year (sd)**	**WAIS score (sd)**	**d' (sd)**
Mean (SE)	33/43	27.8(6.1)	110.6 (12.7)	2.6 (0.5)
	Association with PRS			
Test	two-sample T test	correlation	correlation	correlation
PRS1	T=0.7593, P=0.4503	r=0.0792, P=0.4965	r=-0.0455, P=0.7189	r=-0.0727, P=0.5441
PRS2	T=0.9480, P=0.3464	r=0.0228, P=0.8453	r=-0.1252, P=0.3204	r=-0.1120, P=0.3489
PRS3	T=-1.1015, P=0.2743	r=0.0897, P=0.4410	r=-0.0333, P=0.7923	r=-0.0752, P=0.5301
PRS4	T=-0.4480, P=0.6555	r=0.0992, P=0.3937	r=0.0262, P=0.8361	r=-0.1463, P=0.2202
PRS5	T=-0.3204, P=0.7495	r=0.0630, P=0.5888	r=0.0303, P=0.8104	r=-0.1429, P=0.2313
PRS6	T=-0.1122, P=0.9109	r=0.0526, P=0.6520	r=-0.0232, P=0.8547	r=-0.2727, **P=0.0205**
PRS7	T=-0.2844, P=0.7769	r=-0.0049, P=0.9662	r=0.0275, P=0.8279	r=-0.1738, P=0.1442
PRS8	T=-0.2689, P=0.7888	r=-0.0086, P=0.9415	r=0.0437, P=0.7296	r=-0.2021, P=0.0887
PRS9	T=-0.4136, P=0.6804	r=-0.0728, P=0.5322	r=0.0126, P=0.9208	r=-0.1417, P=0.2352
PRS10	T=-0.4477, P=0.6557	r=-0.0602, P=0.6057	r=0.0283, P=0.8229	r=-0.1345, P=0.2599

Table S2. Demographic data of the Duke Neurogenetics Study (DNS) sample with statistics of the relationship between each variable and PRS.

	Sex (M/F)	**Age year (sd)**	**WAIS score (sd)**	**d' (sd)**
Mean (SE)	63/74	19.8 (1.2)	123.2 (12.8)	2.7 (0.5)
	Association with PRS			
Test	two-sample T test	correlation	correlation	correlation
PRS1	T=1.7390, P=0.0843	r=-0.0138, P=0.8732	r=-0.1433, P=0.0948	r=-0.0170, P=0.8441
PRS2	T=1.9108, P=0.0582	r=-0.0203, P=0.8138	r=-0.0784, P=0.3623	r=0.0610, P=0.4807
PRS3	T=1.2204, P=0.2247	r=-0.0863, P=0.3162	r=-0.1267, P=0.1402	r=-0.0201, P=0.8164
PRS4	T=0.7898, P=0.4311	r=-0.0288, P=0.7387	r=-0.0793, P=0.3571	r=0.0362, P=0.6754
PRS5	T=-0.0122, P=0.9903	r=-0.0844, P=0.3268	r=-0.1344, P=0.1173	r=0.0344, P=0.6908
PRS6	T=0.1683, P=0.8666	r=-0.0359, P=0.6769	r=-0.0836, P=0.3316	r=0.0793, P=0.3589
PRS7	T=0.0740, P=0.9411	r=-0.0140, P=0.8706	r=-0.0761, P=0.3766	r=0.0388, P=0.6542
PRS8	T=-0.3103, P=0.7568	r=0.0144, P=0.8674	r=-0.0552, P=0.5220	r=0.0497, P=0.5657
PRS9	T=-0.4914, P=0.6239	r=-0.0143, P=0.8682	r=-0.0461, P=0.5925	r=0.0505, P=0.5596
PRS10	T=-0.6268, P=0.5319	r=-0.0121, P=0.8887	r=-0.0494, P=0.5666	r=0.0550, P=0.5248

Table S3: Number of SNPs at each p value in the PGC2 GWAS used for calculating PRS

PRS	PGC2 GWAS P	Number of SNPs
1	5e-08	106
2	1e-06	238
3	1e-04	1,253
4	0.001	3,369
5	0.01	10,568
6	0.05	24,670
7	0.1	35,779
8	0.2	51,807
9	0.5	81,224
10	1	102,217

Table S4: Relationship between different PRS and hippocampal activation during memory in the discovery sample. Reported statistics refer to the peak of brain activation in the hippocampal region corresponding to each PRS. FDR correction was conducted within AAL right posterior hippo-parahippocampal mask in SPM12.

PRS	PGC2 GWAS P	Peak	Z Stats	P (uncorrected)	P (FDR corrected)	Number of voxels within peak cluster
1	5e-08	NA*	NA	NA	NA	NA
2	1e-06	NA	NA	NA	NA	NA
3	1e-04	[21, -27, -24]	2.16	0.015	0.760	15
4	0.001	[21, -27, -24]	2.77	0.003	0.747	27
5	0.01	[33, -27, -21]	4.14	<0.001	0.009	100
6	0.05	[30, -30, -18]	4.10	<0.001	0.016	121
7	0.1	[36, -33, -15]	3.34	<0.001	0.095	100
8	0.2	[33, -33, -15]	3.50	<0.001	0.055	134
9	0.5	[30, -33, -15]	3.74	<0.001	0.026	132
10	1	[30, -33, -15]	3.73	<0.001	0.030	137

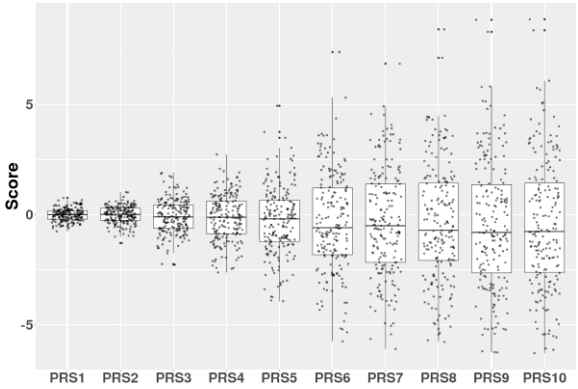
***: NA means no activation under P<0.05.**

Table S5: Variance of hippocampal activation during memory explained by PRS, sex, age, IQ, and d' in the discovery sample. PRS (genetic component) accounted for more percentage of variance of hippocampal activation during memory task compared with sex, age, IQ and d', and PRS6 accounted more variance among different levels of PRSs.

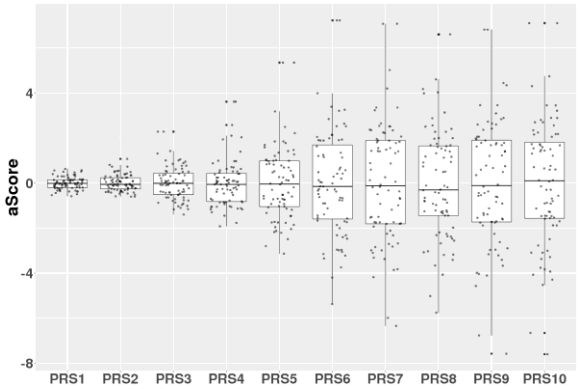
	PRS	**Sex**	**Age**	**IQ**	**d'**
PRS3	3.17%	0.64%	0.01%	0.05%	0.32%
PRS4	3.65%	0.60%	0.05%	0.04%	0.04%
PRS5	8.10%	0.51%	0.71%	0.94%	1.56%
PRS6	**9.11%**	**1.05%**	**1.22%**	**1.41%**	**1.70%**
PRS7	6.98%	0.48%	1.76%	3.04%	3.17%
PRS8	7.14%	0.10%	1.64%	2.95%	3.03%
PRS9	7.23%	0.42%	1.88%	3.13%	3.18%
PRS10	7.17%	0.53%	1.84%	3.09%	3.28%

Figure S1: Distributions of adjusted PRS scores of primary discovery sample (A), Bari sample (B), and DNS sample (C).

A.



B.



C.

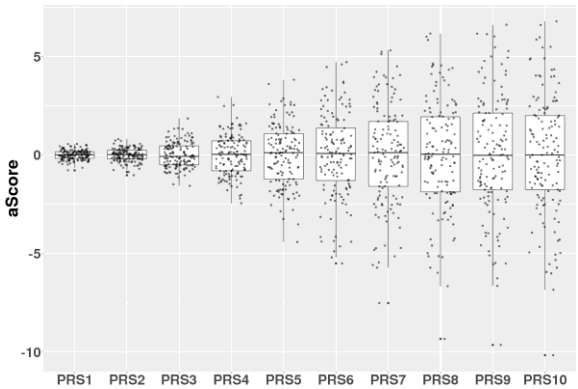
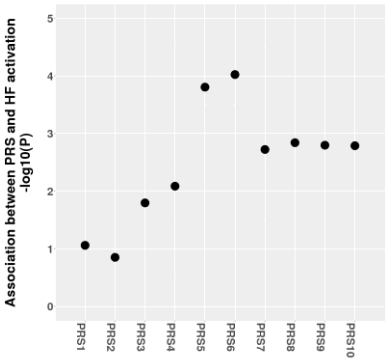


Figure S2: Relationship between different PRS and hippocampal activation during episodic memory in the discovery sample. Analysis was conducted in R using extracted ROI**. A.** Negative logarithm of the p-values of the relationship between activation (BOLD) in right posterior hippocampus and different PRS. PRS6 showed the most significant association (P=9.45e-05). **B.** Variance of brain activation in right hippocampus explained by different PRS. PRS6 explained 8.23% of variance of brain activation (BOLD) in right hippocampus.

A.



B.

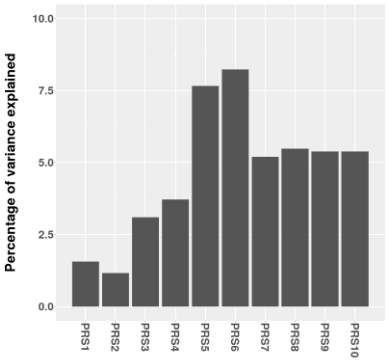
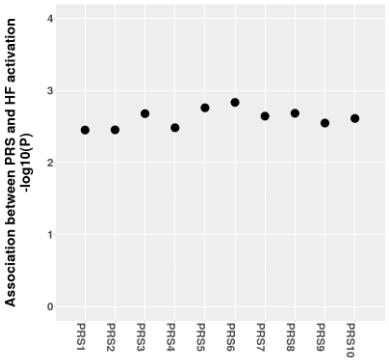


Figure S3: Relationship between different PRS and hippocampal activation during episodic memory in the Bari sample. Analysis was conducted in R using extracted ROI **A.** Negative logarithm of the p-values of the relationship between activation (BOLD) in right posterior hippocampus and different PRS. PRS6 showed the most significant association (0.0015). **B.** Variance of brain activation in right hippocampus explained by different PRS. PRS6 explained 12.82% of variance of brain activation (BOLD) in right hippocampus.

A.



B.

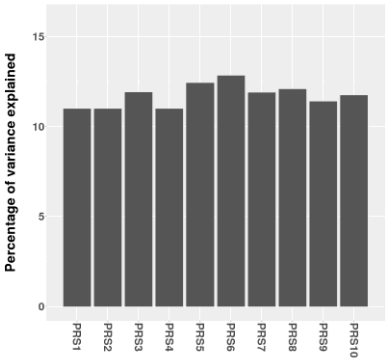
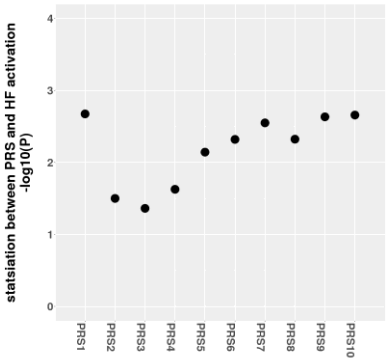


Figure S4: Relationship between different PRS and hippocampal activation during episodic memory in the Duke sample. Analysis was conducted in R using extracted ROI **A.** Negative logarithm of the p-values of the relationship between activation (BOLD) in right posterior hippocampus and different PRS. PRS1 showed the most significant association (P=0.0021). **B.** Variance of brain activation in right hippocampus explained by different PRS. PRS1 explained 6.77% of variance of brain activation (BOLD) in right hippocampus, and PRS6 explained 5.71% of variance.

A.



B.

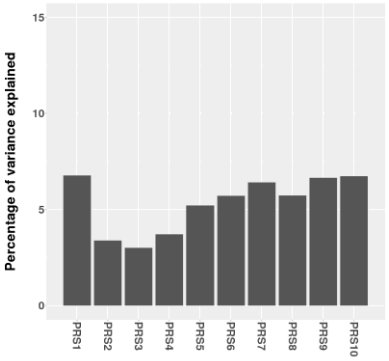


Figure S5: The association between hippocampal activation and PRS6 is not confounded by IQ. Section views (sagittal, coronal and transversal views) of the relationship between PRS6 and fMRI hippocampal activity during neutral encoding in Lieber sample, with IQ as covariate (left side), and scatterplot of the relationship between PRS6 and right posterior hippocampal activity (right side). The inclusion or exclusion of IQ as a covariate does not lead to a change in the association between hippocampal activation and PRS6 (see Figure 1 for comparison).

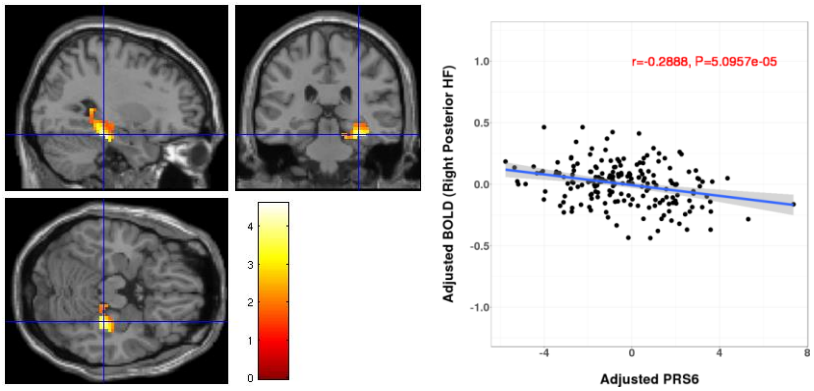
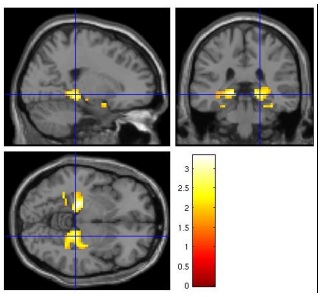
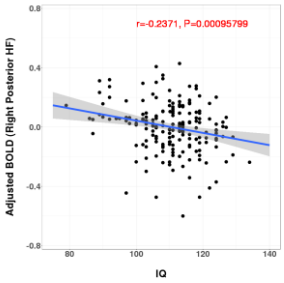


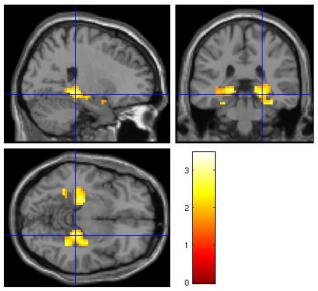
Figure S6. The association between hippocampal activation and IQ is moderated by PRS6. We analyzed the relationship between IQ and hippocampal activation, with and without PRS6 as covariate. Section views (sagittal, coronal and transversal views) of the relationship between IQ and fMRI hippocampal activity during neutral encoding in Lieber sample (left side), and scatterplots of the relationship between IQ and right posterior hippocampal activity (right side). We found a significant relationship between IQ and activity in the posterior hippocampus (**A**). When we included PRS6 as a covariate, the significance of the association between IQ and hippocampal activity was reduced (**B**). We further evaluated the IQ effect in different ranges of PRS6, and we found that the association between IQ and hippocampal activation is much stronger in lower PRS6 (**C**) compared with results in middle (**D**) or higher (**E**) PRS6.

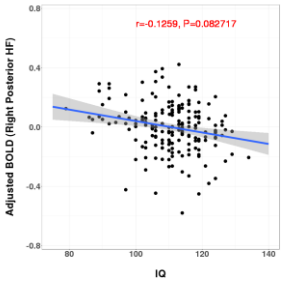
A.



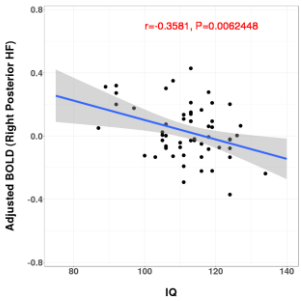


B.

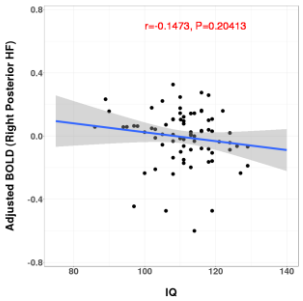




C.



D.



E.

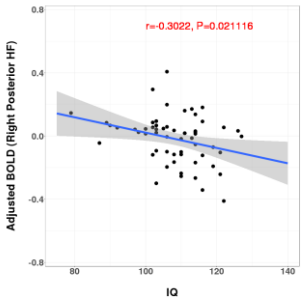
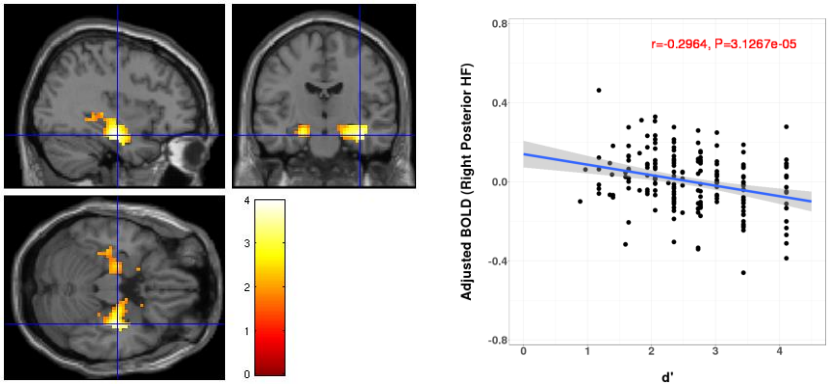


Figure S7: Retrieval accuracy (d') is negatively associated with hippocampal activation. Retrieval accuracy (calculated using d' as detailed in the Supplemental Materials above) was negatively associated with hippocampal activation during memory encoding. The displayed results are thresholded at $p < 0.05$ uncorrected and masked using a bilateral hippocampal ROI. Hippocampal activation remained significant after FWE correction (Peak coordinates [36, -15, -21], $p = 0.035$ FWE corrected within AAL bilateral hippocampal-parahippocampal mask, cluster size = 447).



References:

Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 1996; 29(3): 162-73.
Greve DN, Fischl B. Accurate and robust brain image alignment using boundary-based registration. Neuroimage 2009; 48(1): 63-72.
Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC*, et al.* Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage 2009; 46(3): 786-802.
Nichols T. Notes on Creating a Standardized Version of DVARS, 1–5. arXiv:170401469 2017.
Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 2014; 84: 320-41.
Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. Nature 2014; 511(7510): 421-7.