PNAS
www.pnas.org

Supplementary Information for

Anatomical and Microstructural Determinants of Hippocampal Subfield Functional Connectome Embedding

Reinder Vos de Wael, Sara Larivière, Benoît Caldairou, Seok-Jun Hong, Daniel S. Margulies, Elizabeth Jefferies, Andrea Bernasconi, Jonathan Smallwood, Neda Bernasconi, Boris C. Bernhardt

Email: boris.bernhardt@mcgill.ca

This PDF file includes:

Supplementary Materials and Methods
Figs. S1 to S8

www.pnas.org/cgi/doi/10.1073/pnas.1803667115

Supplementary Materials and Methods

Participants Among the initial 898 participants provided in the HCP S900 release, we selected only those 816 that completed all four rs-fMRI scans and in whom structural MRI data (*i.e.,* T2w, T1w, 1.25mm resampled T1w) were available. Remaining participants were members of 371 unique families, among which 92 contained a monozygotic twin pair. We defined 4 groups: two that each contained one twin member (n=92 each), and two groups of unrelated participants, selected randomly (n=139 and n=140). We combined one group of monozygotic twins with one group of unrelated individuals as discovery dataset and used the remaining group of unrelated participants for validation. For a participant selection schema, see (**Fig. S1**) Our cohort selection excluded family relations between and within groups. All hippocampal subfield segmentations (*below*) were visually inspected. Participants with erroneous segmentations (n=20) were discarded. Details of each dataset were as follows: discovery (n=217, 122 women, mean±SD age=28.5±3.7 years) and validation (n=134, 77 women, age=28.7±3.8 years).

MRI acquisition MRI data were acquired on the HCP's custom 3T Siemens Skyra. Two T1w images with identical parameters were acquired using a 3D-MPRAGE sequence (0.7mm isotropic voxels, matrix=320×320, 256 sagittal slices; TR=2400ms, TE=2.14ms, TI=1000ms, flip angle=8°; iPAT=2). Two T2w images were acquired with identical geometry (TR=3200ms, TE=565ms, variable flip angle; iPAT=2). Four rs-fMRI scans were acquired using multi-band accelerated 2D-BOLD echo-planar imaging (2mm isotropic voxels, matrix=104×90, 72 sagittal slices; TR=720ms, TE=33ms, flip angle=52°; mb factor=8; 1200 volumes/scan). Subjects were instructed to keep their eyes open, look at fixation cross, and not fall asleep. While T1w and T2w scans were acquired on the same day, rs-fMRI scans were split over two days (two scans/day).

Image preprocessing Structural and functional MRI data underwent HCP's minimal preprocessing (1-3)**.**
a) Structural MRI. Images underwent gradient nonlinearity correction. When repeated scans were available, these were co-registered and averaged. Following brain extraction and readout distortion correction, T1w and T2w images were co-registered using rigid body transformations. Subsequently, non-uniformity correction using T1w and T2w contrast was applied (4). Segmentations of subcortical structures were extracted using FSL FIRST (5). Preprocessed images were nonlinearly registered to MNI152 space and cortical surfaces were extracted using FreeSurfer 5.3.0-HCP (6-8), with minor modifications to incorporate information from both T1w and T2w (1). Cortical surfaces in individual subjects were aligned using MSMAll (9) to the hemisphere-symmetric Conte69 template (10).
b) rs-fMRI. Timeseries were corrected for gradient nonlinearity, and head motion was corrected using a rigid body transformation. The R-L/L-R blipped scan pairs were used to correct for geometric distortions. Distortion corrected images were warped to T1w space using a combination of rigid body and boundary-based registrations (11). These transformations were concatenated with the transformation from native T1w to MNI152, to warp functional images to MNI152. Further processing removed the bias field (as calculated for the structural image), extracted the brain, and normalized whole brain intensity. A high-pass filter (>2000s FWHM) corrected the time series for scanner drifts, and additional noise was removed using the ICA-FIX procedure (2). Additional tissue-specific signal regression was not performed (12, 13).

Hippocampal subfield surface generation and feature sampling
a) Subfield surface generation. Based on each subjects' minimally processed T1w images in MNI152-space, we automatically segmented the left/right hippocampus into Subiculum (consisting of presubiculum, parasubiculum, subiculum proper), CA1-3, and CA4-DG using a

validated multi-template surface-patch algorithm (14). The algorithm was trained on an open-access database of manual subfield segmentations and corresponding high-resolution 3T MRI data (15). Segmentations were visually quality controlled in all subjects and are available upon request. The algorithm incorporates a spherical harmonic shape parameterization and point distribution model of the surfaces (16). We generated medial sheet representations running through each subfield's core using a Hamilton-Jacobi approach (17), to minimize partial volume effects during feature sampling. Furthermore, we propagated the spherical harmonic parameterization from the outer shell to the medial sheet to improve vertex-correspondence across individuals based on shape-inherent information.

b) Surface-wide sampling of structural and functional features. Using the non-linear registration matrices between native and MNI152 space provided by HCP, we mapped medial sheet meshes and volumetric rs-fMRI data to native T1w. Functional time-series were sampled at each sheet vertex and at each vertex of the MSMAll-registered (9) mid-thickness cortical surfaces. We also sampled T1w/T2w intensity (4). Surface-sampled features were smoothed using a Gaussian diffusion kernel with 5 mesh units as full-width-at-half-maximum (FWHM) in all subfields and cortex. Sampling was carried out in native T1w and rs-fMRI space to minimize interpolation. We also synthesized lower dimensionality representations of the neocortex by averaging unsmoothed timeseries within 360 cortical areas (9) and 7 intrinsic functional networks (18). To reduce blurring between communities at the boundaries of the 7-network parcellation, we eroded each community label by two vertices. Subcortical volumetric timeseries were extracted from all subcortical voxels and averaged within each structure.

Subfield-to-cortex connectivity analysis

a) Per-subfield connectivity mapping. To summarize subfield-specific connectivity to the neocortex in each subject, we computed Pearson correlation coefficient maps between the average timeseries across all medial sheet vertices for a given subfield and each cortical vertex. Analysis was carried out separately for left and right subfields. Correlations underwent Fisher-z transformations and maps were averaged across scans.

b) Connectome gradient mapping. After parcellating the cortex into 360 regions (9), we systematically computed z-scored connectivity matrices *Z* between each of the 4k medial sheet vertices and each cortical area. We utilized diffusion embedding (19), an unsupervised learning algorithm, to identify principal modes of spatial variations in connectivity when going from one subfield vertex to another. This technique has recently described spatial gradients in neocortical rs-fMRI connectivity (20). To adapt the approach to hippocampus-to-cortex connectivity, we converted Z into a cosine similarity matrix C, which scales with the angle between the 360-dimensional connectivity profiles of each vertex pair. To allow for negative numbers, we applied an arccosine to C, divided the result by π (to scale to [0 1]), and subtracted it from 1 (such that 1 denotes identical orientation). The other parameters were identical to those previously described (20), i.e., $\boldsymbol{\alpha}$= 0.5, automated diffusion time estimation, thresholding of every column in Z at its 90^{th} percentile. Left/right hippocampal components underwent Procrustes alignment (100 iterations) (21). In a separate analysis, we performed the gradient mapping with 13 subcortical regions (brainstem, amygdala, caudate nucleus, nucleus accumbens, putamen, pallidum and thalamus) as additional regions of interest, with separate left/right regions for all areas except the brainstem. A MATLAB implementation of the diffusion embedding algorithm is available at https://github.com/MICA-MNI/micaopen/, the Python code on which our implementation is based on is available at https://github.com/satra/mapalign.

Data analysis Analysis was carried out using SurfStat for MATLAB (22), available at http://www.math.mcgill.ca/~keith/surfstat.

a) Subfield-specific connectivity analysis. We mapped the connectivity of each subfield in each hemisphere separately using one-sample t-tests on Fisher-to-z-transformed connectivity maps.

We also assessed differences in connectivity between subfields in each hemisphere, using mixed-effects models with *subject* as random and *subfield* as fixed effect. Findings were corrected using random field theory for non-isotropic images, with cluster-defining thresholds (CDT) of CDT=0.001 and a family-wise error rate of 0.025 (23).

b) Assessing the relation between subfield gradients and hippocampal anatomy. Within each subfield, we determined the correspondence between functional gradients and long-axis anatomy, by computing overlaps with manual segmentations of hippocampal head, body, and tail obtained from a previously published T1w MRI protocol and dataset (24). Specifically, we clustered gradients using k-means (k=3, no other constraint) and computed maximal Dice indices between clusters and head-body-tail segmentations.

We also correlated gradient values across all subfield surface with T1w/T2w intensity, modeling subfield as a fixed effect. Results were verified after also controlling for per-vertex estimates of cerebrospinal fluid (CSF) partial volume, columnar volume, and temporal signal to noise ratio. CSF partial volume as estimated using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/), columnar volume was computed as the average volume of the columns formed by a medial sheet's vertex's surrounding triangles with their corresponding outer shell triangles, and temporal signal-to-noise ratio was computed as the mean of the signal in each vertex divided by its standard deviation. All analyses were systematically carried out for the first three gradients. Notably, G1, G2, and G3 collectively explained 52% of variance (**Fig. S3**).

All correlations between gradients and other markers were compared across gradients using Steiger's test (25). As the sign of the gradient is arbitrary, we flipped the sign of gradients with negative correlations resulting in only positive correlations. This provides more conservative significance testing as the difference between correlation values decreases or remains the same.

c) Relation to task co-activation patterns and reverse inference. Support for long-axis specialization in humans has so far been provided mainly by task-based fMRI. We confirmed this using Neurosynth (download date: November 2, 2017) (26), which performed an automated meta-analysis of 11406 neuroimaging studies. We produced task co-activation networks for each voxel by placing 6mm spheres in each hippocampal voxel and assessing which voxels in neocortical gray matter were likely to be reported as active when the seed region is active. We used nearest neighbor interpolation to assign co-activation patterns to hippocampal subfield vertices and performed diffusion embedding on the resulting co-activation matrix.

To assess cognitive differences across the principal gradient, we calculated the centroids of its top and bottom 33%. Using the task co-activation networks of these centroids, we performed an automated reverse inference (26). In brief, for each term appearing in the Neurosynth database, we computed a reverse inference map which is the probability of a voxel being reported as active given that the term appeared in the manuscript. To determine which terms are associated with the top and bottom of the principal gradients, we correlated the centroids' task co-activation networks with all reverse inference maps.

Reliability and reproducibility assessment Several analyses assessed consistency of our main findings. To estimate subject-specific *test-retest stability*, we split each subject's functional scan into two sets (each set contained 1 scan from each day and each phase encoding direction). Hippocampal-cortical connectivity profiles of both sets, as well as gradient maps were correlated for each subject. As components with diffusion embedding are not necessarily in the same order, we computed gradients for each set within a subject, aligned these to the group-level components of all other subjects with Procrustes alignment, and correlated aligned components. *Reproducibility* was determined by re-computing results in the validation dataset and correlating hippocampal-cortical connectivity profiles and the group-level components between discovery and validation datasets.

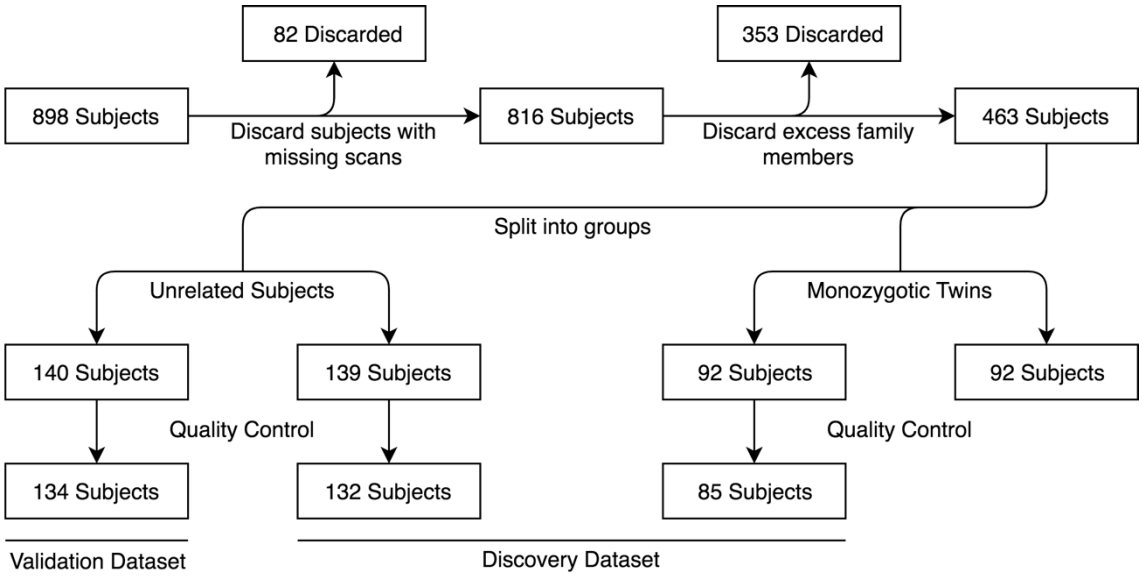

Fig. S1. Schematic of the subject selection. From the initial 898 subjects, 82 were discarded for missing scans. Another 353 to remove familial relationships across subjects. Where possible, a monozygotic twin pair was retained. The monozygotic twins were split into two datasets of 92 subjects each, with each group containing one member of each pair. The remaining subjects were split into two groups. We then build the discovery and verification cohorts based on these subgroups. All groups underwent a visual quality control to exclude inaccurate hippocampal segmentations.

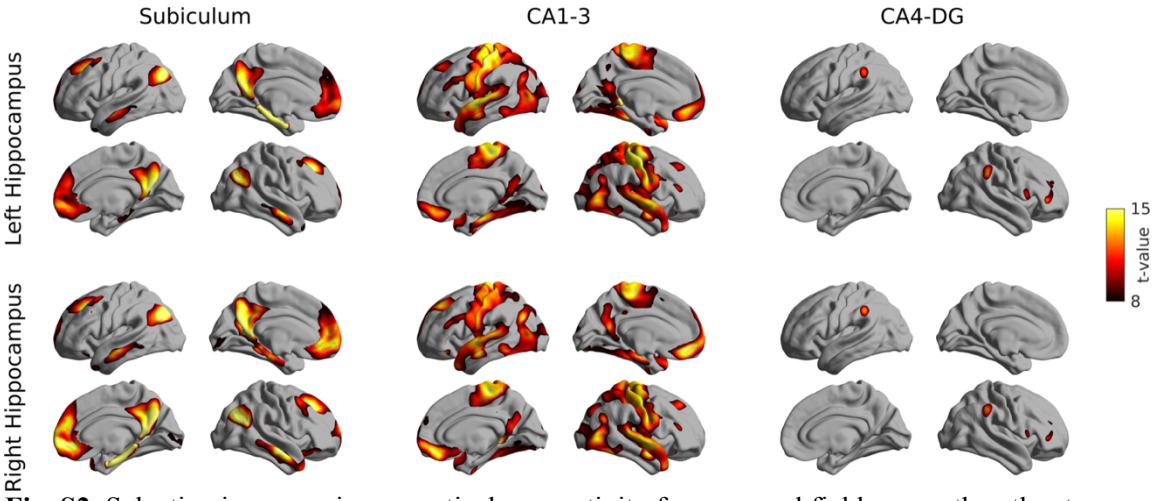

Fig. S2. Selective increases in neocortical connectivity from one subfield versus the other two. Subiculum connected more strongly to the default mode network and CA1-3 connected more strongly to somatomotor and limbic regions, while CA4-DG did not show a selective increase in neocortical integration compared to the other regions. Findings were corrected for multiple comparisons.

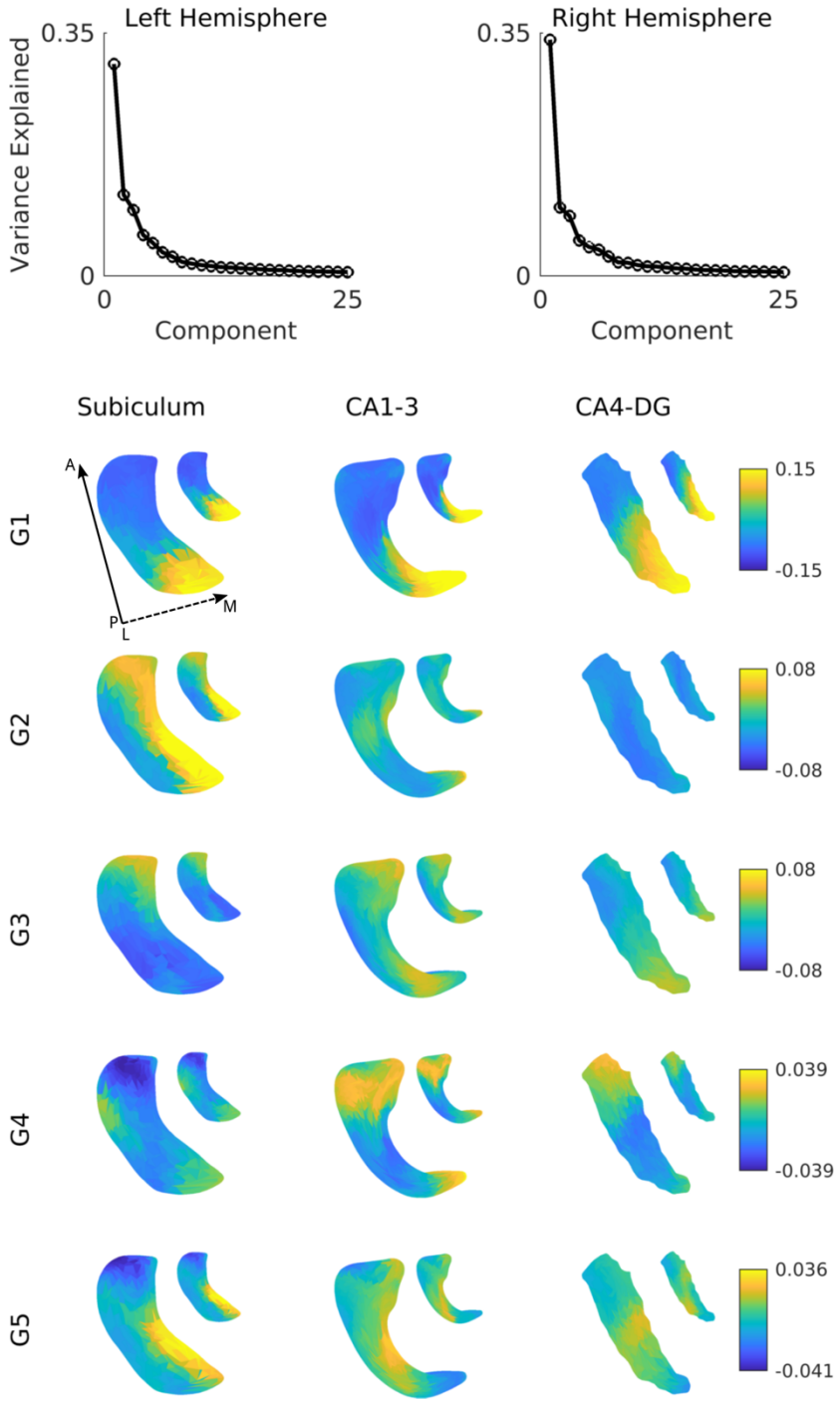

Fig. S3. Overview of the hippocampal gradients. **(A)** Variance explained by the first 25 components for the left and right hippocampi. **(B)** First 5 gradients of the subiculum (*left*) CA1-3 (*middle*) and CA4-DG (*right*). Large and small surfaces represent the left and right hippocampus, respectively. Solid and dashed arrows denote posterior (P) to anterior (A) and lateral (L) to medial (M) direction, respectively.

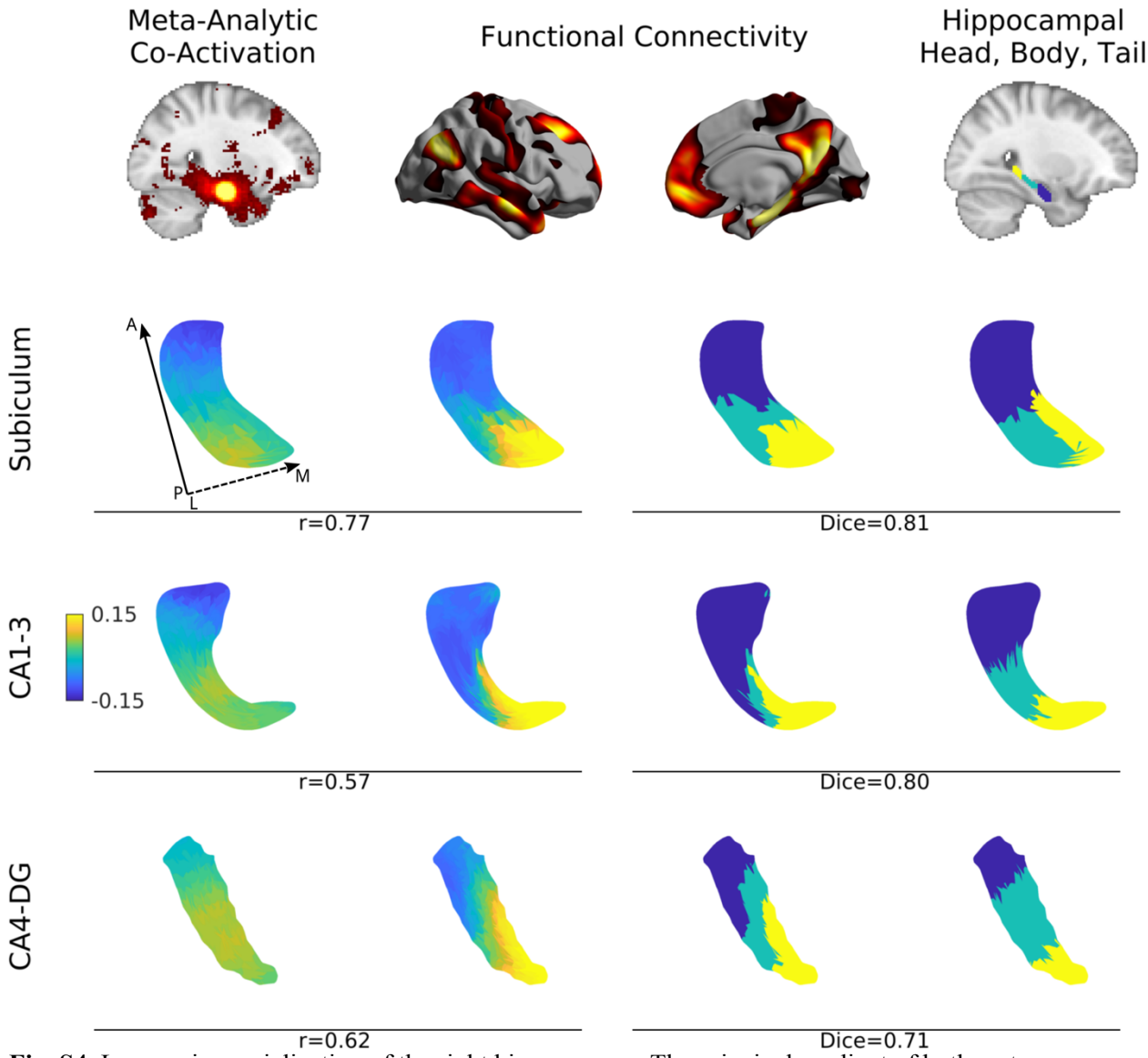

Fig. S4. Long-axis specialization of the right hippocampus. The principal gradient of both meta-analytic task-fMRI co-activation (*first column*) and rs-fMRI (*second column*) ran in anterior-posterior direction. K-means clusters (k=3) of the rs-fMRI connectivity derived gradient (*third column*) overlapped strongly with manual segmentations of hippocampal head, body, and tail based on a previous atlas (24) (*fourth column*). Solid and dashed arrows denote posterior (P) to anterior (A) and lateral (L) to medial (M) direction, respectively.

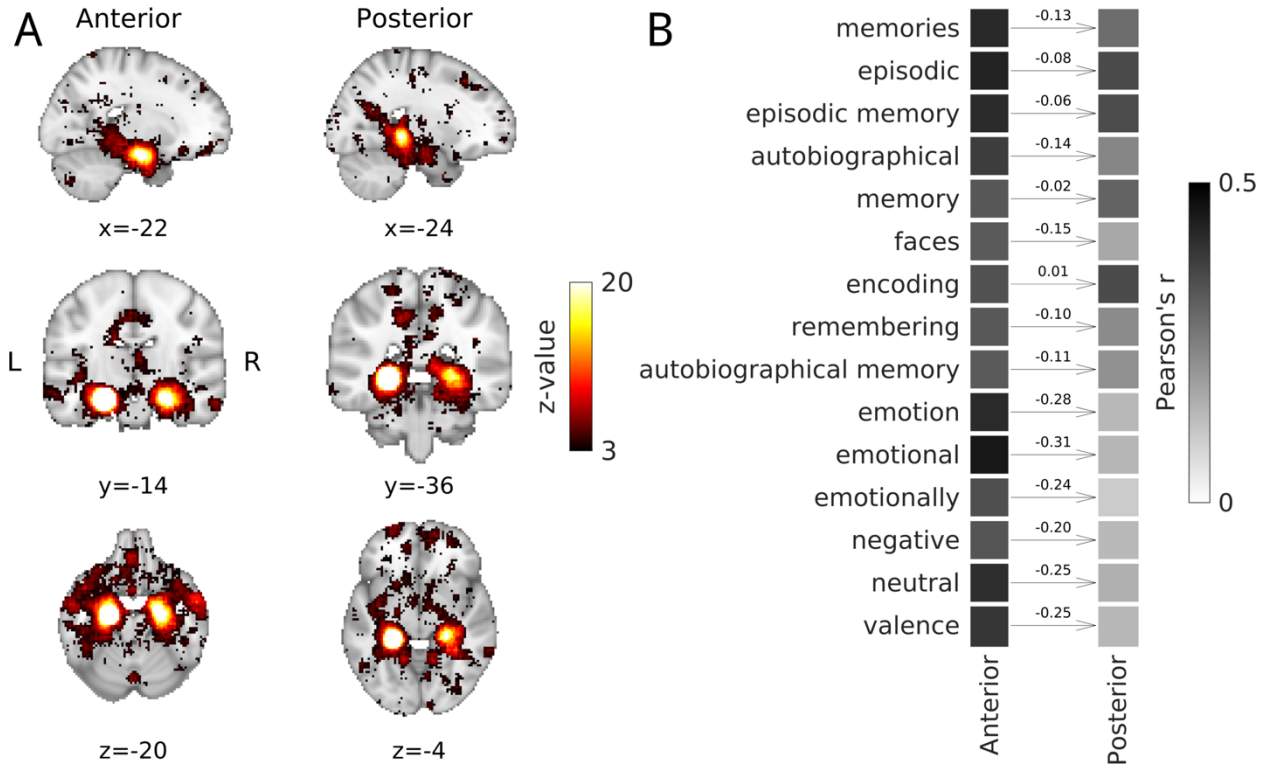

Fig. S5. Automated reverse inference based on Neurosynth. **(A)** Task-based co-activation networks of the centroids of the top (*posterior*) and bottom (*anterior*) 33% of the principal gradient. **(B)** Top 15 terms most strongly associated with the task co-activation derived networks of the anterior and posterior seeds. Terms associated with anatomy and task design were discarded.

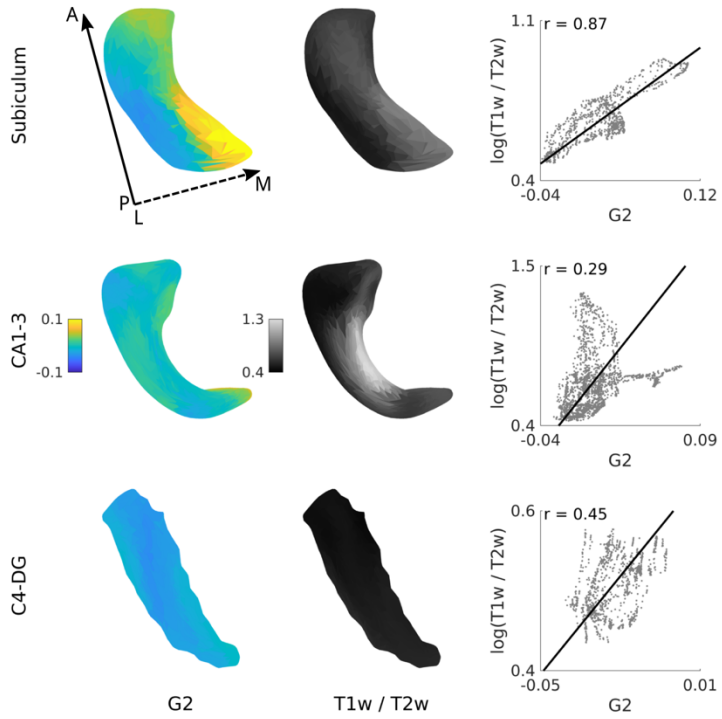

Fig. S6. Right hippocampal second gradient and T1w/T2w intensities. Systematic correlation analyses indicated highest correlations between surface-sampled T1w/T2w and the second gradient, which runs in along the hippocampal infolding. Solid and dashed arrows denote posterior (P) to anterior (A) and lateral (L) to medial (M) direction, respectively.

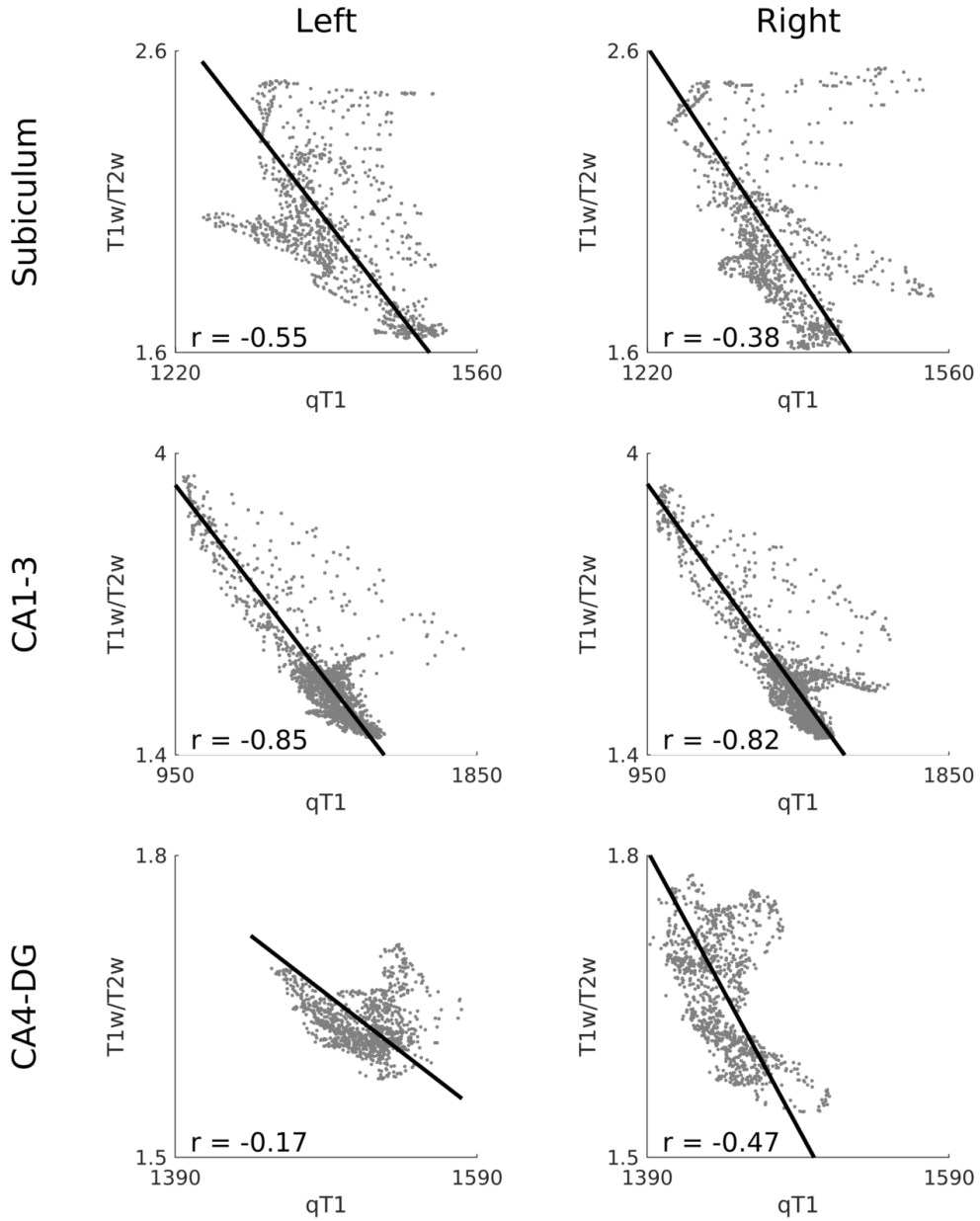

Fig. S7. Correspondence between T1w/T2w and quantitative T1 relaxation time (qT1) measurements from a different dataset of 20 healthy adults (24). Pearson correlations are shown in the bottom left of each scatterplot. Overall, correspondence is good, although differences exist across the subfields.

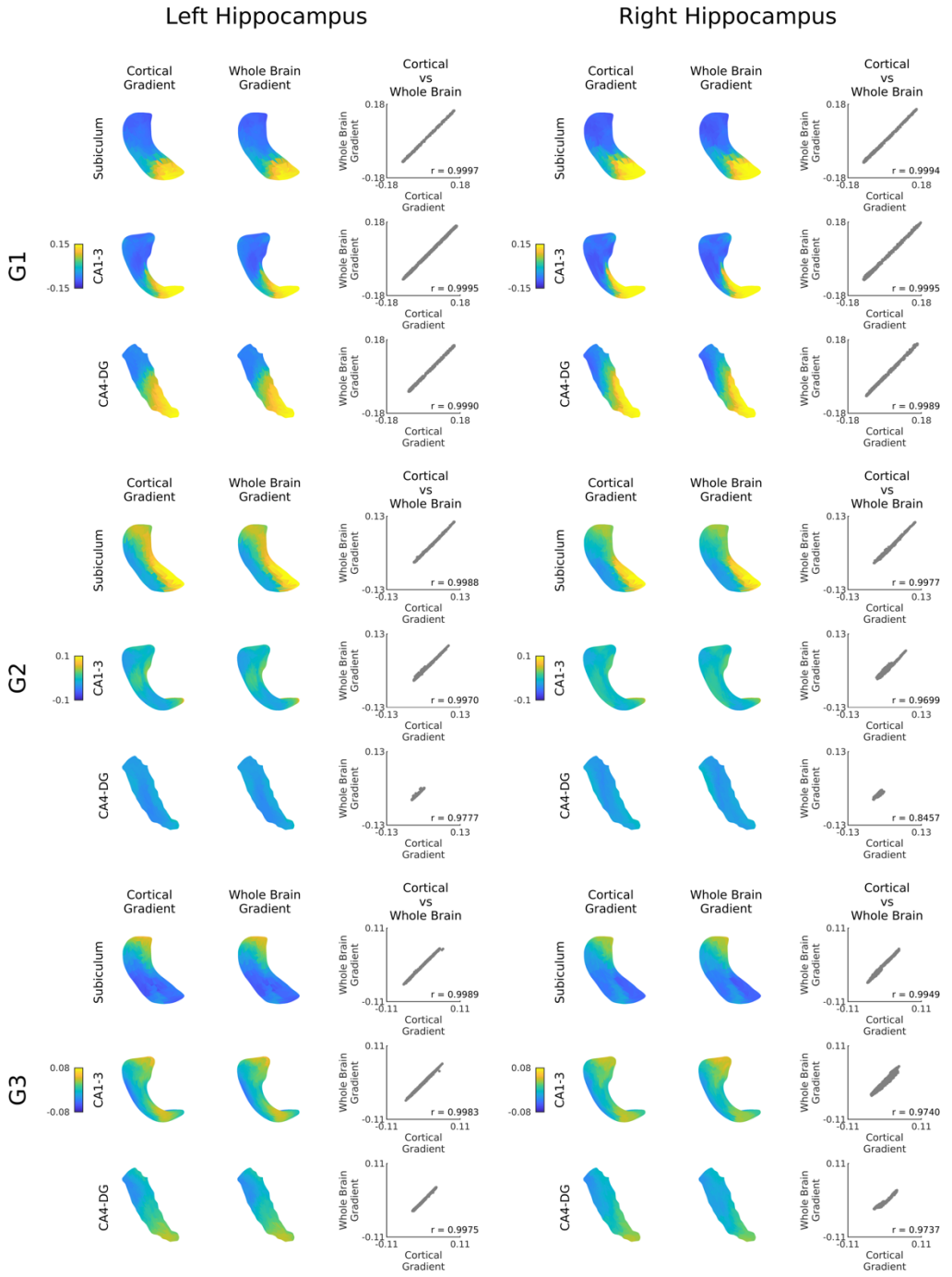

Fig. S8. Comparison of hippocampal-cortical and hippocampal-whole brain G1 (*top*), G2 (*middle*), and G3 (*bottom*) for the left (*left*) and right (*right*) hippocampus. Gradients remain virtually unchanged after including subcortical areas. Cortical gradient surfaces represent the gradients as used in the body of the manuscript. Whole brain gradients are computed after additional including subcortical areas. Scatter plots show the vertex-wise correspondence between these two gradients and their correlation values.

References

1. Glasser MF, *et al.* (2013) The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* 80:105-124.
2. Salimi-Khorshidi G, *et al.* (2014) Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage* 90:449-468.
3. Smith SM, *et al.* (2013) Resting-state fMRI in the human connectome project. *Neuroimage* 80:144-168.
4. Glasser MF & Van Essen DC (2011) Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *J. Neurosci.* 31(32):11597-11616.
5. Patenaude B, Smith SM, Kennedy DN, & Jenkinson M (2011) A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* 56(3):907-922.
6. Dale AM, Fischl B, & Sereno MI (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9(2):179-194.
7. Fischl B, Sereno MI, & Dale AM (1999) Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9(2):195-207.
8. Fischl B, Sereno MI, Tootell RB, & Dale AM (1999) High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp* 8(4):272-284.
9. Glasser MF, *et al.* (2016) A multi-modal parcellation of human cerebral cortex. *Nature* 536(7615):171-178.
10. Van Essen DC, Glasser MF, Dierker DL, Harwell J, & Coalson T (2012) Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. *Cereb Cortex* 22(10):2241-2262.
11. Greve DN & Fischl B (2009) Accurate and robust brain image alignment using boundary-based registration. *Neuroimage* 48(1):63-72.
12. Murphy K & Fox MD (2017) Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage* 154:169-173.
13. Vos de Wael R, Hyder F, & Thompson GJ (2017) Effects of Tissue-Specific Functional Magnetic Resonance Imaging Signal Regression on Resting-State Functional Connectivity. *Brain Connect* 7(8):482-490.
14. Caldairou B, *et al.* (2016) A Surface Patch-Based Segmentation Method for Hippocampal Subfields. *Medical Image Computing and Computer-Assisted Intervention,* Lecture Notes in Computer Science, (Springer), pp 379-387.
15. Kulaga-Yoskovitz J, *et al.* (2015) Multi-contrast submillimetric 3 Tesla hippocampal subfield segmentation protocol and dataset. *Sci Data* 2:150059.
16. Styner M, *et al.* (2006) Framework for the Statistical Shape Analysis of Brain Structures using SPHARM-PDM. *The insight journal* (1071):242-250.
17. Kim H, *et al.* (2014) Multivariate hippocampal subfield analysis of local MRI intensity and volume: application to temporal lobe epilepsy. *Med. Image Comput. Comput. Assist. Interv.* 17(Pt 2):170-178.
18. Yeo BTT, *et al.* (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106(3):1125-1165.
19. Coifman RR & Lafon S (2006) Diffusion maps. *Appl. Comput. Harmon. Anal.* 21(1):5-30.
20. Margulies DS, *et al.* (2016) Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl. Acad. Sci. U. S. A.* 113(44):12574-12579.
21. Langs G, Golland P, & Ghosh SS (2015) Predicting Activation Across Individuals with Resting-State Functional Connectivity Based Multi-Atlas Label Fusion. *Med Image Comput Comput Assist Interv* 9350:313-320.

22. Worsley KJ*, et al.* (2009) SurfStat: A Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. *Neuroimage* 47:S102.
23. Worsley KJ, Andermann M, Koulis T, MacDonald D, & Evans AC (1999) Detecting changes in nonisotropic images. *Hum Brain Mapp* 8(2-3):98-101.
24. Bernasconi N*, et al.* (2003) Mesial temporal damage in temporal lobe epilepsy: a volumetric MRI study of the hippocampus, amygdala and parahippocampal region. *Brain* 126(Pt 2):462-469.
25. Steiger JH (1980) Tests for comparing elements of a correlation matrix. *Psychological bulletin* 87(2):245.
26. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, & Wager TD (2011) Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* 8(8):665-670.