Prefrontal default-mode network interactions with posterior hippocampus during exploration

Andrew E. Papale[1], Vanessa M. Brown[2], Angela M. Ianni[1], Michael N. Hallquist[3], Beatriz Luna[1], Alexandre Y. Dombrovski[1]

1. Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, USA
2. Department of Psychology, Emory University, Atlanta, GA, USA
3. Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Correspondence: Alexandre Y. Dombrovski, 728 Bellefield Towers, 3811 O'Hara St., Pittsburgh, PA, 15213; ayd1@pitt.edu

Number of pages: 68
Number of figures: 7
Number of tables: 1
Words for abstract (147), significance (101), introduction (649), and discussion (1419)

Conflict of interest statement: The authors declare no competing financial interests.

Author Contributions
Conceptualization: AYD and MNH. Investigation: MNH and BL. Resources: MNH, BL and AYD. Formal analysis: AYD, MNH, AMI, BEL, and AEP. Writing – original draft: AEP, VMB, AYD. Writing – editing: All authors. Project administration: AYD.

Acknowledgements

Abbreviated title: Hippocampal-prefrontal interactions and exploration

1 **Abstract**

2 Hippocampal maps and ventral prefrontal cortex (vPFC) value and goal representations

3 support foraging in continuous spaces. How might hippocampal-vPFC interactions

4 control the balance between behavioral exploration and exploitation? Using fMRI and

5 reinforcement learning modeling, we investigated vPFC and hippocampal responses as

6 humans explored and exploited a continuous one-dimensional space, with out-of-

7 session and out-of-sample replication. The spatial distribution of rewards, or value

8 landscape, modulated activity in the hippocampus and default network vPFC

9 subregions, but not in ventrolateral prefrontal control subregions or medial orbitofrontal

10 limbic subregions. While prefrontal default network and hippocampus displayed higher

11 activity in less complex, easy-to-exploit value landscapes, vPFC-hippocampal

12 connectivity increased in uncertain landscapes requiring exploration. Further,

13 synchronization between prefrontal default network and posterior hippocampus scaled

14 with behavioral exploration. Considered alongside electrophysiological studies, our

15 findings suggest that locations to be explored are identified through coordinated activity

16 binding prefrontal default network value representations to posterior hippocampal maps.

17

18 Word count (147)

19

20 **Significance Statement**

21 The ventral prefrontal cortex represents goals and values, while the hippocampus

22 contains maps of physical and abstract spaces. In recent years, neuroscientists have

23 sought to understand how hippocampal-prefrontal interactions help us to solve complex

24 problems such as deciding whether to exploit known rewards or explore in search of

25 better alternatives. Using functional magnetic resonance imaging and reinforcement

26 learning modeling, we examine these interactions as humans explore and exploit an

27 environment with a complex distribution of rewards or value landscape. We describe

28 how this value landscape modulates activity and connectivity between prefrontal cortex

29 and hippocampus, controlling the balance between exploration and exploitation.

30

31 Word count (101)

32

33 **Introduction** (649 words)

34 In uncertain environments, animals face the dilemma of whether to *exploit* known good

35 options or *explore* potentially better alternatives. Behavior shifts from exploration to

36 exploitation as valuable options are discovered. In environments with a few discrete

37 options mammalian striatum and amygdala can resolve this dilemma efficiently (Costa

38 et al., 2019), but more complex spaces require hippocampal cognitive maps (Johnson

39 et al., 2012; Mehlhorn et al., 2015; Dombrovski et al., 2020) and flexible reward

40 representations in ventral prefrontal cortex (vPFC; Domenech et al., 2020; Trudel et al.,

41 2021).

42

43   Locations of desirable outcomes can be predicted by a cognitive map integrating

44   available cues with spatial memory. The hippocampus creates value-laden maps based

45   on trajectories through space (Stachenfeld et al., 2017; Dombrovski et al., 2020). These

46   maps contain more fine-grained representations in posterior/dorsal divisions and

47   reward- and context- based representations in anterior/ventral divisions (Royer et al.,

48   2010; Biane et al., 2023). Moreover, the hippocampus replays trajectories (Skaggs and

49   McNaughton, 1996; Liu et al., 2019; Schuck and Niv, 2019) or considers future goals

50   online during deliberation (Johnson and Redish, 2007; Papale et al., 2012; Redish,

51   2016). vPFC, in turn, estimates option values based on current needs and goals

52   (Padoa-Schioppa, 2007; Bartra et al., 2013; Tang et al., 2021). We have found that

53   human anterior hippocampus represents spatially structured values, and these

54   representations scale with successful exploitation (Dombrovski et al., 2020). However,

55   how interactions between various hippocampal and vPFC subregions may support

56   exploration and exploitation remains unclear.

57

58   Hippocampal-vPFC interactions are facilitated by long-range synchronization of local

59   field potentials (Buzsáki, G., 2006). Consistent with behavioral evidence, offline delta

60   synchronization (Schultheiss et al., 2020) – linked to sharp wave ripples in hippocampus

61   – promotes memory consolidation and retrieval of contextually relevant information

62   (Buzsáki, 2015; Joo and Frank, 2018; Tang and Jadhav, 2019), potentially including

63   option values (Weilbächer and Gluth, 2016; Schmidt and Redish, 2021). Additionally,

64   theta synchronization during directed attention conveys hippocampal spatial

65 representations to the vPFC (Jones and Wilson, 2005; Gluth et al., 2015; Weilbächer

66 and Gluth, 2016).

67

68 Value-related information differentially modulates vPFC subregions that fall into three

69 large-scale connectivity networks: the ventral limbic network (LIM), the default mode

70 network (DMN) and the frontoparietal control network (CTR). The value of the chosen or

71 best available option positively modulates medial and central orbitofrontal and

72 subcallosal LIM as well as peri-callosal and fronto-polar DMN (Hare et al., 2008; Bartra

73 et al., 2013; Smith et al., 2014; Lopez-Gamundi et al., 2021; Knudsen and Wallis, 2022).

74 By contrast, lateral OFC regions such as 11/47 belonging to CTR are negatively

75 modulated by expected value (Ursu et al., 2008), suggesting competition among

76 similarly-valued options for action selection. Connectivity between the putative rodent

77 analogue of LIM (infralimbic PFC and ventral orbitofrontal cortex, OFC) and ventral

78 hippocampus appears necessary for exploitation of learned values (Chudasama et al.,

79 2012). In addition, interactions between hippocampus and more dorsal rodent prelimbic

80 PFC (putatively homologous to human DMN; Lu et al., 2012) may govern exploitation

81 (Laureiro-Martínez et al., 2015) and exploration (Morici et al., 2022). Given the diverging

82 value responses of the CTR subregion, we reasoned that it would serve as an

83 informative comparison when examining activity and hippocampal connectivity of DMN

84 and LIM vPFC.

85

86   We therefore investigated (1) how spatially structured value functions or value

87   landscapes modulate activity of the human vPFC (DMN, LIM, and CTR), as well as

88   vPFC connectivity with anterior vs. posterior hippocampus, and (2) how such

89   connectivity relates to behavioral exploration. To address these aims, we examined

90   fMRI BOLD activity during exploration and exploitation of a continuous one-dimensional

91   space. We inferred the value distribution over the space using our previously validated

92   information-compressing RL model of adaptive transition from exploration to exploitation

93   (Hallquist and Dombrovski, 2019). Multi-level analyses of event-locked deconvolved

94   BOLD signal (Dombrovski et al., 2020) enabled us to characterize value encoding

95   across within-trial timepoints, during online navigation, outcome receipt and offline

96   processing seconds after an outcome. We replicated key findings out-of-session and

97   out-of-sample.

98

99   **Materials and Methods**

100   *PARTICIPANTS – EXPERIMENT 1*

101   Participants were 71 individuals aged 14-30 (mean +/- SD: 21±5 years; 37 female, 34

102   male) with no history of neurological disorder, brain injury, or developmental or

103   psychiatric disorder (self or first-degree relatives). Participants or their legal guardians

104   provided informed consent prior to participation. Procedures complied with the Code of

105   Ethics of the World Medical Association (1964 Declaration of Helsinki) and were

106   approved by the Institutional Review Board of the University of Pittsburgh (protocol

107   PRO10090478). Participants were compensated $75 for this experiment.

108

PARTICIPANTS – EXPERIMENT 2

Participants were psychiatrically and neurologically healthy individuals (N=43) age 49-

80 (mean +/- SD: 63±8 years; 23 female, 20 male. These data were drawn from the

healthy comparison group of a study of suicidal behavior in the second half of life.

Participants were compensated $150 for completing the experiment and other studies.

Experimental procedures for this study complied with Code of Ethics of the World

Medical Association (1964 Declaration of Helsinki) and the Institutional Review Board at

the University of Pittsburgh (STUDY19030288).

117

REGIONS OF INTEREST: PREFRONTAL CORTEX AND HIPPOCAMPUS

We analyzed 17 subdivisions of vPFC, encompassing medial and lateral OFC, frontal

pole, and vmPFC/pericallosal ACC. These were selected from the 7-network 200-region

parcellation of (Schaefer et al., 2018; Fig. 1A). Following this parcellation, we assigned

regions to one of three networks: the Default Mode (DMN, orange), the Frontoparietal

(Cognitive) Control (CTR, yellow), or the Limbic (LIM, blue). Sixteen of these regions

were divided into symmetric left-right pairs and labeled using the Eickhoff-

Zilles Cytoarchitectonic MPM atlas 2.2 in AFNI (Fig. 1C; (Eickhoff et al., 2005, 2007;

Henssen et al., 2016). To preserve the left-right symmetry, the 17th region, right

frontopolar BA 10 (fp10 region 161, MNI x=13, y=64, z=-14) was moved from the limbic

network to the DMN. Our sensitivity analyses showed that this change had no

qualitative impact on the results in Experiment 1 for Fig. 4B, 5B, and Fig. 6C,E,G (Fig.

130   S4, Fig. S5). The human frontal DMN expands dorsally, thus we included region dorsal

131   BA 10 (d10 region 89 left MNI x=-7, y=52, z=11, d10 region 194 right MNI x=7, y=52,

132   z=10) in our analyses.

133

134   The bilateral hippocampal mask was derived from the Harvard-Oxford subcortical atlas

135   in MNI152 space (https://identifiers.org/neurovault.collection:262) and resampled to

136   2.3mm (Experiment 1) voxels or 3.1mm (Experiment 2) voxels to match the native

137   resolution of the BOLD data. To delineate anterior (AH) vs. posterior hippocampus

138   (PH), the postero-anterior long axis was defined as described previously (Dombrovski et

139   al., 2020) with the posterior 1/3 comprising PH and anterior 2/3 comprising AH (Fig. 1B).

140

141   *BEHAVIORAL TASK – EXPERIMENT 1*

142   On the clock task (Moustafa et al., 2008), participants were shown a clock face and

143   asked to make a single choice per trial, stopping a dot which revolves clockwise around

144   a central stimulus in 4s. The choice led to immediate feedback indicating points earned

145   on that trial, displayed for 0.9s (Fig. 2A). Unknown to participants, earnings were

146   controlled by one of four difficult reinforcement contingencies with probability-magnitude

147   tradeoffs, in 50-trial runs. Thus, one learned that certain locations on the clock yielded

148   higher average rewards. Two reinforcement contingencies were learnable: *increasing*

149   *expected value* (IEV), where later response times yielded higher average payoffs, vs.

150   *decreasing expected value* (DEV). In IEV and DEV, as well as one unlearnable

151   condition, *constant expected value* (CEV), reward probabilities decreased while

152   magnitudes decreased throughout the interval. In a final, unlearnable *constant expected*

153   *value–reversed* (CEVR) condition, the probability-magnitude tradeoff was reversed. If

154   the participant failed to make a response within 4s (360 degrees) they received zero

155   points (Fig. 2B). Subjects consistently chose longer response times in the IEV condition

156   (green), and shorter response times in the DEV condition (orange) in Experiment 1 (Fig.

157   2C) both in the fMRI (left panel) and out-of-session MEG replication (center panel)

158   sessions.

159

160   Participants completed eight 50-trial runs of the clock task during an fMRI scan. The

161   clock task was implemented in PsychToolbox 3.0.10 (Brainard, D.H., 1997; Kleiner, M

162   et al., 2007) and run in MATLAB 2012a (The Mathworks, Inc. Natick, Massachusetts,

163   www.mathworks.com). The inter-trial interval varied in length according to an

164   exponential distribution, with the sequence and length optimized for detection power

165   using a Monte Carlo approach implemented by the *optseq2* command in Freesurfer 5.3.

166   Specifically, 5 million possible inter-trial interval sequences were simulated, with the top

167   320 sequences retained based on their estimation efficiency. For each subject, 8 of

168   these efficient inter-trial interval sequences were selected, one for each scanner run.

169

170   Participants also completed 8 runs of the clock task during a separate MEG session

171   (Experiment 1 - non-fMRI Replication Session, or Experiment 1 - Out-of-Session MEG

172   Replication), which was counterbalanced in order with the fMRI session and separated

173   by an average of 3.7 weeks. For the purposes of this paper, this second session

174   allowed examination of stable tendencies to explore or exploit and replication of effects

175   across time. Due to lower signal-to-noise ratio of MEG relative to fMRI, runs consisted

176   of 63 trials.

177

178   During each trial, the central stimulus was an abstract phase-scrambled image, or a

179   happy or fearful face. The phase-scrambled images were intended to provide equal

180   luminance and coloration. Faces were selected from the NimStim database (Tottenham

181   et al., 2009). All 4 contingencies were collected with scrambled images, while only IEV

182   and DEV had emotion manipulations. We do not report on central stimulus effects here,

183   presenting average estimates across stimuli.

184

185   *BEHAVIORAL TASK – EXPERIMENT 2*

186   Participants completed the clock task described above, but with six modifications. 1) To

187   dissociate value entropy from novelty, the contingency reversed every 40 trials unknown

188   to the participants. 2) Participants completed 240 trials of the clock task in two runs. 3)

189   Only IEV and DEV contingencies were employed. 4) Trials were extended to 5s to

190   accommodate slower psychomotor speed in this older sample. 5) No MEG data were

191   collected in a separate session. 6) The central stimulus was always a generic clock

192   face. Subjects consistently chose longer response times in the IEV condition (green),

193   and shorter response times in the DEV condition (orange) in Experiment 2 (Fig. 2C).

194

195   *COMPUTATIONAL MODELING OF BEHAVIOR*

196   To capture learning and value representation on the clock task, we used a previously

197   validated reinforcement learning (RL) model (StrategiC Exploration/Exploitation of

198   Temporal Instrumental Contingencies, or SCEPTIC; Hallquist and Dombrovski, 2019).

199   The SCEPTIC model represents the one-dimensional value landscape of the clock task

200   using a set of unnormalized Gaussian radial basis functions spaced evenly over the

201   interval. Each function has a temporal receptive field with a mean and variance defining

202   its point of maximal sensitivity and the range of times to which it is sensitive. The

203   quantity of interest tracked by the bases is the expected value of a given choice

204   indicated by the response time. To represent time-varying value, the heights of the

205   basis functions are scaled according to a set of *B* weights, $w = [w_1, w_2, ..., w_b]$. The

206   contribution of each basis function to the integrated value representation depends on its

207   temporal receptive field:

$$\varphi_b(x) = \exp\left[-\frac{(x - \mu_b)^2}{2s_b^2}\right] \tag{1}$$

208   where *x* is an arbitrary point within the time interval *T*, μ_b is the center (mean) of the

209   RBF and s_b^2 is its variance. And more generally, the temporally varying expected value

210   function on a trial *i* is obtained by the multiplication of the weights with the basis:

$$V(i) = \sum_{b=1}^{B} w_b(i)\varphi_b \tag{2}$$

211

212   For the clock task, where the probability and magnitude of rewards varied over the

213   course of four-second trials, we spaced the centers of 24 Gaussian RBFs evenly across

214   the discrete interval and chose a fixed width s_b^2 for the temporal variance (width) of each

215   basis function. More specifically, s_b^2 was chosen such that the distribution of adjacent

216   radial basis functions overlapped by approximately 50%.

217   The basic model learns the expected values of different response times by updating

218   each basis function *b* according to the equation:

$$w_b(i+1) = w_b(i) + e_b(i|t)\alpha[\text{reward}(t) - w_b(i)] \tag{3}$$

219   where *i* is the current trial in the task, *t* is the observed response time, and $\text{reward}(t)$ is

220   the reinforcement obtained on trial *i* given the choice *t*. Prediction error updates are

221   weighted by the learning rate α and the eligibility e_b. To avoid tracking separate value

222   estimates for each possible moment, feedback obtained at a given response time *t* is

223   propagated to adjacent times. This is achieved using a Gaussian RBF centered on the

224   response time *t,* having width s_g^2. The eligibility e_b of a basis function φ_b to be updated

225   by prediction error is defined as its overlap with the temporal generalization function, *g*:

$$g(x) = \exp\left[-\frac{(x-t)^2}{2s_g^2}\right] \tag{4}$$

$$e_b(i|t) = \int_0^T \min[\,g(\tau), \varphi_b(\tau)]d\tau \tag{5}$$

226   where τ represents an arbitrary timepoint along the interval *T.* Thus, for each radial

227   basis function b, a scalar eligibility e_b between zero and one represents the overlap with

228   the temporal generalization function s_b^2 .

229   The SCEPTIC model selects an action based on a softmax choice rule, analogous to

230   simpler reinforcement learning problems (e.g., two-armed bandit tasks; Sutton and

231   Barto, 2018). For computational speed, we arbitrarily discretized the interval into 100 ms

232   time bins such that the agent selected among 40 potential responses (i.e., a multinomial

233   representation). At trial i the agent chooses a response time in proportion to its

234   expected value:

$$p[t(i + 1) = j | V(i)] = \frac{\exp\left(V(i)_j / \beta\right)}{\sum_{\tau=0}^{T} \exp\left(V(i)_\tau / \beta\right)} \tag{6}$$

235   where j is a specific response time and the temperature parameter β controls value

236   sensitivity such that choices become more stochastic and less value-sensitive at higher

237   values.

238

239   Importantly, as detailed previously (Hallquist and Dombrovski, 2019), selective

240   maintenance of frequently chosen and preferred actions performs better than other

241   alternative models. Specifically, basis weights corresponding to a non-preferred,

242   spatiotemporally distant action revert toward a prior in inverse proportion to the temporal

243   generalization function:

$$w_b(i + 1) = w_b(i) + e_b(i|t)\alpha[reward(t) - w_b(i)] - \gamma\left(1 - e_b(t)\right)(w_b(i) - h) \tag{7}$$

244

245   where γ is a selective maintenance parameter between zero and one that scales the

246   amount of decay towards a point h, which is taken to be zero here for parsimony, but

247   could be replaced with an alternative prior expectation. Our primary fMRI analyses

248   used signals derived from fitting the information-compressing RL model (eq.7) to

249   participants' behavior.

250   We defined the information content of the learned value distribution as Shannon's

251   entropy of the normalized basis weights (the trial index i is omitted for simplicity of

252   notation):

$$H(\mathbf{w}) = -\sum_{b=1}^{B} w_b \log_{10}(w_b) \tag{8}$$

253

254   The SCEPTIC model was fit to individual choices using the empirical Variational Bayes

255   Approach (Daunizeau et al., 2014). This method relied on a mixed-effects model where

256   individual-level parameters were assumed to be sampled from a normal distribution.

257   The group's summary statistics were then inferred from individual-level posterior

258   parameter estimates using an iterative variational Bayes algorithm that alternates

259   between estimating population parameters and the individual subject parameters. Over

260   algorithm iterations, individual-level priors are minimized toward the inferred parent

261   population distribution, as in standard multilevel regression. The SCEPTIC model was

262   fit separately for Experiment 1 and Experiment 2.

263

264   For analysis of fMRI signals trial-by-trial, we used two variables from the SCEPTIC

265   model, the maximum of $V(i)$, or V_{max} (Eq. 2), and the entropy $H(\mathbf{w})$ (Eq. 8).

266

267   *FMRI ACQUISITION, PREPROCESSING AND GLM ANALYSIS – EXPERIMENT 1*

268   We used a Siemens Tim Trio 3T scanner with the following T2*-optimized sequence:

269   TR=1.0s, TE=30ms, flip angle = 55°, multiband acceleration factor = 5, voxel size =

270   2.3mm^3. T1 anatomical scans for coregistration were acquired with voxel size =

271   1.0mm^3, TR=2.2s, TE=3.58ms, and GRAPPA 2x acceleration. A gradient echo

272   fieldmap TE=4.93ms and 7.39ms was acquired to correct for magnetic field

273   inhomogeneity.

274

275   Anatomical scans were preprocessed as described previously (Dombrovski et al.,

276   2020). Briefly, slice timing and motion coregistration were performed using NiPy in AFNI

277   (v.19.0.26; Roche, 2011) and FMRIB in FSL (v6.0.1). Skull stripping was performed by

278   masking low intensity voxels and via the ROBEX brain extraction algorithm (Iglesias et

279   al., 2011). Functional images were aligned to T1 anatomical images using white matter

280   segmentation and boundary-based registration (Greve and Fischl, 2009). Given that

281   fast-TR sequences have reduced gray-white contrast, we improved the functional-

282   structural coregistration by first aligning functional volumes to a single-band reference

283   image with improved contrast. Functional scans were then transformed into MNI152

284   template space (FSL FNIRT) and smoothed using a 5mm FWHM kernel in FSL SUSAN

285   either within the hippocampal mask (for hippocampus signals) or across the whole

286   brain. Head motion artifacts were reduced via an ICA analysis in FSL MELODIC, with

287   the spatiotemporal components then passed to an ICA-AROMA classification algorithm

288   (Pruim et al., 2015). Components that were classified as noise were regressed out of

289   the signal using FSL regfilt (nonaggressive regression approach). Susceptibility

290   artifacts were reduced using fieldmap correction using FSL FUGUE. An 0.008 Hz

291   temporal high-pass filter removed slow-frequency signals (Woolrich et al., 2001). Voxel

292   time series were renormalized to have a mean of 100 to allow similar scaling of

293   voxelwise regression coefficients across run and subject.

294

295   Runs where framewise displacement exceeded 10% of total volumes had a framewise

296   displacement of ≥0.9mm, or subjects where peak displacement exceeded 5 mm at any

297   point during the scan were excluded from all analyses. This led to a total of 549 usable

298   runs, with 11 excluded from GLM analysis. There were no subjects for which both runs

299   were excluded due to motion. No runs were excluded for deconvolved signal analyses

300   (described below).

301

302   Voxelwise general linear models (GLM) for BOLD data were estimated in FSL (v6.0.1;

303   S. M. Smith et al., 2004) using FSL FEAT (v6.0). All models included regressors for

304   clock and feedback events, calculated by convolving boxcar regressors for clock and

305   feedback events with a double-gamma canonical hemodynamic response function

306   (HRF). In addition, parametric BOLD modulation by either V_{max} or entropy of the value

307   distribution was modeled by convolving corresponding trial onset-aligned SCEPTIC-

308   derived signals with the HRF. Confound regressors from ventricles and cerebral white

309   matter, as well as derivatives of these signals, were also included. At the second

310   (subject) level, parameter estimates were combined from each run using a weighted

311   fixed effects model in FEAT. Level 3 group analyses were then performed using the

312   FSL FLAME 1+2 approach with automatic outlier deweighting (Woolrich, 2008). All

313   group analyses included age and sex as covariates.

314

315   Familywise error correction for voxelwise statistical parametric maps was carried out

316   using Probabilistic Threshold-free Cluster Enhancement method at $p < 0.05$ (Spisák et

317   al., 2019).

318

319   *FMRI ACQUISITION AND PREPROCESSING – EXPERIMENT 2*

320   The imaging sessions were conducted at the University of Pittsburgh's Magnetic

321   Resonance Research Center (RRID:SCR_025215). We used a Siemens Magnetom

322   Prisma scanner with the following T2*-optimized sequence: TR=600ms, TE=27ms, flip

323   angle = 45°, multiband acceleration factor = 5, voxel size = $3.1mm^3$. T1 anatomical

324   scans for coregistration were acquired with voxel size = $1.0mm^3$, TR=2.3s, TE=3.35ms,

325   and GRAPPA 2x acceleration. A gradient echo fieldmap TE=4.47ms and 6.93ms was

326   acquired to correct for magnetic field inhomogeneity.

327

328 The fMRI data were then preprocessed using FMRIPREP version 20.1.1 (Roche, 2011),

329 a Nipype-based (Gorgolewski et al., 2011) tool. Each T1w (T1-weighted) volume was

330 corrected for INU (intensity non-uniformity) using N4BiasFieldCorrection v2.1.0

331 (Tustison et al., 2010) and skull-stripped using antsBrainExtraction.sh v2.1.0 (using the

332 OASIS template). Brain surfaces were reconstructed using recon-all from FreeSurfer

333 v6.0.1 (Dale et al., 1999) and the brain mask estimated previously was refined with a

334 custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived

335 segmentations of the cortical gray-matter of Mindboggle (Klein et al., 2017). Spatial

336 normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c (Fonov

337 et al., 2009) was performed through nonlinear registration with the antsRegistration tool

338 of ANTs v2.1.0 (Avants et al., 2008) using brain-extracted versions of both T1w volume

339 and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter

340 (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast FSL

341 v5.0.9 (Zhang et al., 2001).

342

343 Functional data was slice-time corrected using 3dTshift from AFNI v16.2.07 (Cox, 1996)

344 and motion corrected using mcflirt (FSL v5.0.9; Jenkinson et al., 2002). Distortion

345 correction was performed using fieldmaps processed with fugue (Jenkinson, 2003).

346 This was followed by co-registration to the corresponding T1w using boundary-based

347 registration (Greve and Fischl, 2009) with six degrees of freedom, using bbregister

348 (FreeSurfer v6.0.1). Motion correcting transformations, field distortion correcting warp,

349 BOLD-to-T1w transformation and T1w-to-template (MNI) warp were concatenated and

350   applied in a single step using antsApplyTransforms (ANTs v2.1.0) using Lanczos

351   interpolation.

352

353   Frame-wise displacement (Power et al., 2014) was calculated for each functional run

354   using the implementation of Nipype. ICA-based Automatic Removal Of Motion Artifacts

355   (AROMA) was used to generate non-aggressive noise regressors (Pruim et al., 2015).

356   Components identified as noise were regressed out of the data using FSL regfilt (non-

357   aggressive regression approach). ICA-AROMA has performed very well in head-to-head

358   comparisons of alternative strategies for reducing head motion artifacts (Ciric et al.,

359   2017). We then applied a .008 Hz temporal high-pass filter to remove slow frequency

360   signal changes (Woolrich et al., 2001); the same filter was applied to all regressors in

361   MEDuSA analyses (see below). Finally, we renormalized each voxel time series to have

362   a mean of 100 to provide similar scaling of voxelwise regression coefficients across

363   runs and participants.

364

365   Many internal operations of FMRIPREP use Nilearn (Abraham et al., 2014) principally

366   within the BOLD-processing workflow.

367

368   *MULTILEVEL EVENT-RELATED DECONVOLVED SIGNAL ANALYSIS (MEDuSA)*

369   The standard GLM approach using canonical HRFs is useful for testing hypotheses

370   about session-level linear associations between model-based signals and regional

371   activity. However, GLMs assume 1) that the correct amplitude and duration of task-

372   related neural signals are specified in the model; and 2) that a canonical HRF describes

373   BOLD activity associated with a model-based signal. In addition, GLMs do not allow the

374   experimenter to probe the time courses of signals before or after trial-level events of

375   interest. Therefore, we employed Multilevel Event-related Deconvolved Signal Analysis

376   (MEDuSA), which operates through voxelwise deconvolution of BOLD activity to

377   functionally align the signal to the event of interest (Dombrovski et al., 2020). In

378   addition, MEDuSA operates within a multilevel framework, providing the experimenter

379   with an appropriate statistical analysis of hierarchically structured data. We briefly

380   outline the MEDuSA pipeline below.

381

382   After first applying the preprocessing pipelines described above, a leading

383   hemodynamic deconvolution algorithm was applied voxelwise across the entire run to

384   estimate neural population activity (Bush and Cisler, 2013). Then, the deconvolved

385   signal was extracted from -4s to +4s around trial onset. Linear interpolation was

386   performed to align the deconvolved data onto a TR-consistent grid (1.0s and 0.6s,

387   respectively, for studies 1 and 2). This operation merely resampled the data and did not

388   introduce additional information into the signal. Activity estimates that fell into

389   subsequent trials were censored to ensure analyses only included the time within the

390   current trial. As our behavioral data had a nested structure (trials nested within run,

391   nested within subject), we used multilevel regression models to compute effects.

392   Multilevel models were estimated using restricted maximum likelihood in the *lme4*

393   package (Bates et al., 2015) in R 4.2.1 (R Core Team, 2021). Estimated *p* values for

394   predictors in the models were computed using Wald chi-square tests and degrees of

395   freedom were based on the Kenward-Roger approximation.

396

397   To examine entropy and V_{max} modulation of vPFC and hippocampus, deconvolved

398   signal in each structure was regressed on entropy and value using a multilevel

399   regression framework. In addition, these analyses included several control variables:

400   prior ITI, current response time, scaled inverted (negative) trial number, outcome, age

401   and gender. Random intercepts of run nested within subject were used, and separate

402   models were computed at each time point in the within-trial grid (-4s to +4s) and each

403   structure. MEDuSA can also be used to estimate functional connectivity between

404   different regions by including one region as an independent variable in a regression

405   analysis predicting activity in another region. Thus, for functional connectivity MEDuSA

406   analyses, multilevel models included vPFC deconvolved signal as the dependent

407   variable and main effects and interactions of within-run change in aligned hippocampal

408   activity and entropy and V_{max} (as well as several confound regressors) as independent

409   variables. In all analyses, false-discovery rate multiple comparison was applied

410   (Benjamini and Yekutieli, 2001). Additionally, no runs were removed due to framewise

411   displacement, as in the last step of preprocessing for Experiment 1 GLMs.

412

413   For the analysis in Fig. 5 and Fig. 6 (Hippocampal signal predicting vPFC activity and

414   Effects on explore/exploit), for both Experiment 1 and 2, activity in the hippocampus

415   was z-scored within subject and run (8 runs in Experiment 1, and 2 runs in Experiment

416   2) to emphasize relative activity change within a run. The average hippocampal activity

417   for each subject and run was regressed out as a nuisance effect in the multi-level

418   analysis, which controls for BOLD quality effects and conflation of high global signal

419   with the desired functional connectivity signal. The goal of this approach was to capture

420   trial-level fluctuations in hippocampal activity.

421

422   To examine how neural activity predicted behavior, the relationship between entropy or

423   and network activity for each participant was extracted from random (participant-

424   varying) effects estimates of the vPFC neural activity models. $x_{H(\mathbf{w})}x_{HC}$ case where HC

425   is hippocampal activity and $H(\mathbf{w})$ is entropy from Eq. 8. The outcome term Y_{vPFC} is

426   computed for each subject s and trial t. The random slope of the interaction of

427   hippocampal activity with entropy term for each subject is u_s which gives the subject-

428   level deviation of the regression coefficient from the group regression coefficient γ_0 (Eq.

429   9). *u_s is the key variable we use as a moderator for model-derived effects on behavior*.

430   The overall intercept is γ_0, and the level 2 regression intercepts are u_{0s}. The level 1

431   residual variance is r_{st}.

432

$$Y_{vPFC\,st} = \gamma_0 + (\gamma_t + u_s)x_{H(\mathbf{w})_t}x_{HC_t} + u_{0s} + r_{st} \qquad (9)$$

433

434

435   This subject-level random effect from Eq. 9, u_s, was then averaged across the time

436   window and used as a fixed effect to predict upcoming participant response times in a

437   second multilevel model with the lagged response time (exploration term; Eq. 10) or the

438   lagged response time of the V_{max} (RT_{Vmax}, exploitation term). See Fig. 2D for intuition

439   on how correlations between lagged RT with current RT and lagged RT_{Vmax} with current

440   RT translate into exploration and exploitation. For clarity that it is a fixed effect

441   predicting $Y_{RT\,t}$ in Eq. 10, we rename $u_s = x_s$. $x_{RT(t-1)_t}$ is the lagged RT and x_R is the

442   last outcome, which is a binary variable (reward or omission). The coefficients reported

443   in Fig. 6 represent $Y_{RT\,t}$ from Eq. 10, estimated $x_s = \pm 1$ standard deviations from the

444   mean using a simple slopes analysis (carried out using the *emtrends* package in R).

445   The significant p_{FDR} values reported in the text are derived from main effects of the

446   multilevel model.

$$Y_{RT\,t} = \gamma_0 + u_{0s} + \gamma_t^1\, x_s x_{RT(t-1)_t} x_{R_t} + \gamma_t^2 x_s x_{RT_{Vmax}(t-1)_t} + r_{st} \qquad (10)$$

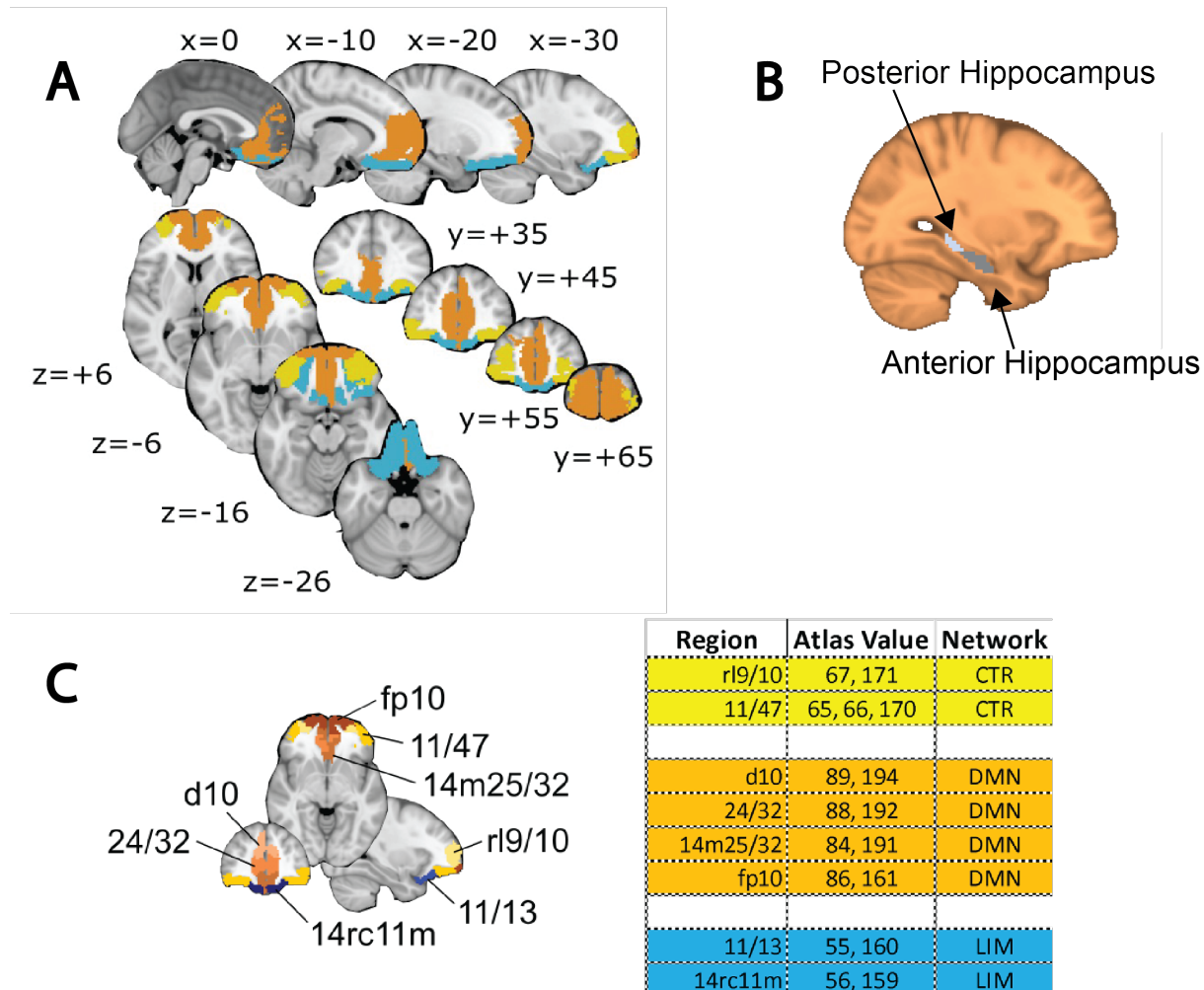| Region	Atlas Value	Network
rl9/10	67, 171	CTR
11/47	65, 66, 170	CTR
d10	89, 194	DMN
24/32	88, 192	DMN
14m25/32	84, 191	DMN
fp10	86, 161	DMN
11/13	55, 160	LIM
14rc11m	56, 159	LIM

Fig. 1: vPFC mask and hippocampus mask. A. The vPFC mask consists of 17 regions grouped into 3 resting-state functional networks as in Yeo et al., 2011, the Default Mode Network (DMN; orange), the Control Network (CTR; yellow), and the Limbic Network (LIM; blue). **B.** The bilateral hippocampal mask is from the Harvard-Oxford subcortical atlas. The posterior 1/3 of the hippocampus was grouped into the Posterior Hippocampus region (PH; light gray) and the anterior 2/3 of the hippocampus was grouped into the Anterior Hippocampus region (AH; dark gray). **C.** The vPFC mask is comprised of 17 parcels of the Schaefer 2018 mask derived from resting state functional connectivity. We labeled these parcels in AFNI (Cox, 1996) based on

457   approximate overlap with the Eickhoff-Zilles Cytoarchitectonic MPM atlas 2.2. Because

458   adjacent regions of the DMN in vPFC included what is clearly dorsal PFC, we included

459   parcels d10 (peach shading), although it is not strictly vPFC. The vPFC mask was

460   grouped by rough left-right symmetry into 8 distinct groups. Models were run separately

461   with each region as a dependent variable to complement the main analyses separating

462   vPFC into three resting state networks (Frontoparietal Control Network - CTR; yellow,

463   Default Mode Network - DMN; orange, and Limbic Network - LIM, blue) in the 7-network

464   solution. Since the parcellation is not strictly symmetric, one of the groups is not a pair
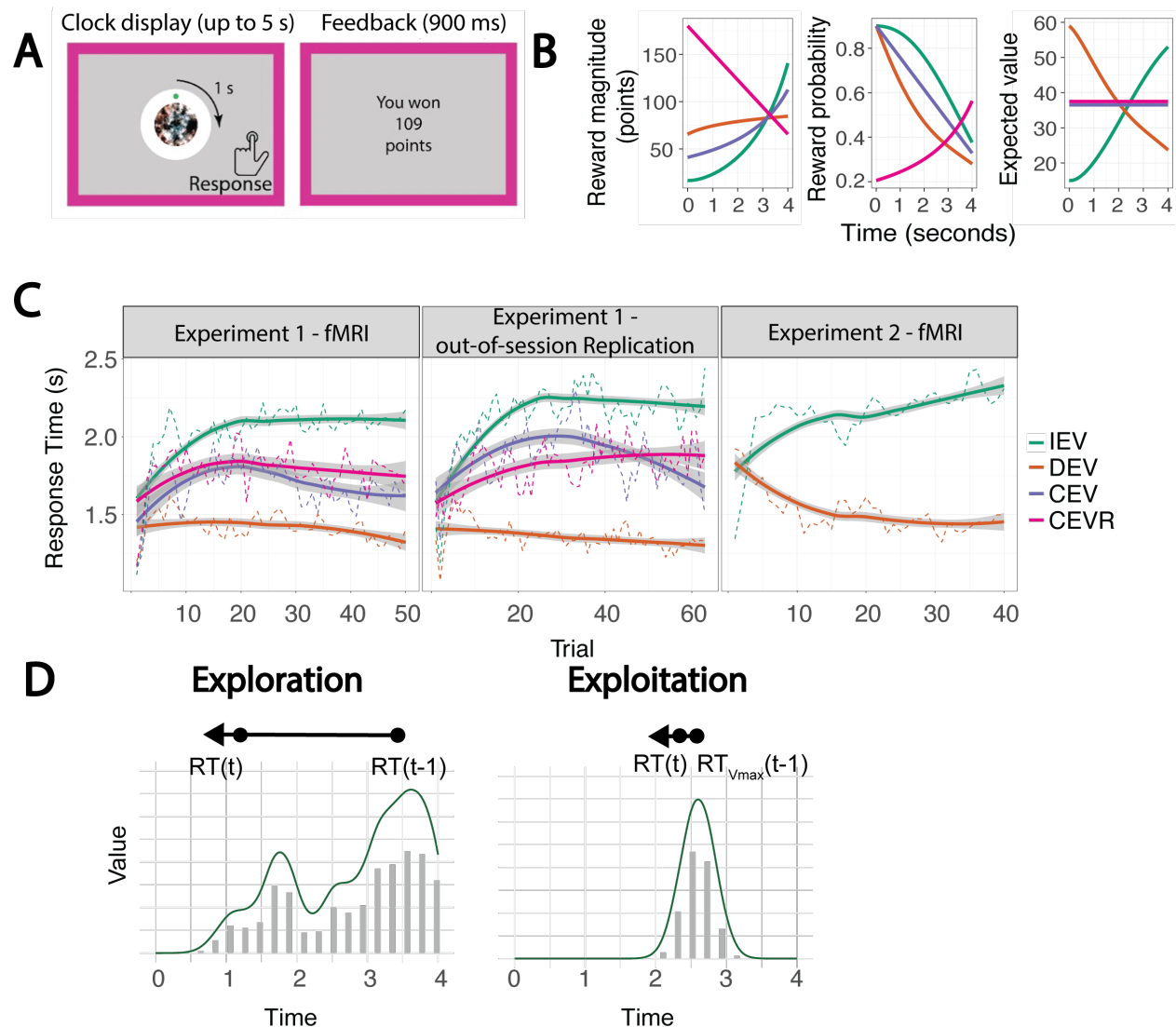
465   (11/47).

466

Fig. 2 Clock task and operationalizing exploration vs. exploitation. A. Each trial on

the clock task begins with a green dot that rotates clockwise from the 12 o'clock position

around a pseudo-clock face with a central image (here a phase-scrambled image) in 4s

(Experiment 1) or 5s (Experiment 2). On each trial, participants press a button to stop

the green dot to earn points in each time-varying contingency. Pressing the button

immediately triggers 900ms of feedback where the reward associated with that choice is

displayed. **B.** Over the course of a session, the participant learns about the underlying

reward contingencies and adjusts their responses accordingly. There are 4

476   contingencies on the clock task: increasing expected value (IEV, green), decreasing

477   expected value (DEV, orange), constant expected value (CEV, blue) and constant

478   expected value reversed (CEVR, pink). During IEV trials (green) subjects learn that

479   responding later in the interval results in a higher overall payoff (greater magnitude but

480   lower probability). In CEV (blue) and CEVR (magenta) trials, the expected value

481   (reward probability times reward magnitude) is constant across the interval. In DEV

482   (orange) trials, subjects learn that responding earlier in the interval results in a higher

483   overall payoff (higher probability and marginally lower magnitude). **C)** Average behavior

484   on the clock task in Experiment 1 (left panel), the out-of-session Replication session of

485   Experiment 1 (middle panel) and Experiment 2 (right panel). The average response

486   time of subjects in IEV (green) increases, indicating that subjects learn to respond later.

487   During DEV trials (orange) subjects learn to respond earlier in the interval with shorter

488   RTs. Responses in the Average responses in the CEV and CEVR conditions lie in

489   between the IEV and DEV conditions. In Experiment 2, the CEV and CEVR conditions

490   were omitted and there were unsigned reversals between IEV and DEV every 40 trials.

491   **D)** On this task, shifts toward the best-known option (RT_{Vmax}) are interpreted as

492   exploitative (right panel) whereas other shifts in RT (left panel) are interpreted as

493   exploration. Note that the left panel has a more complex value distribution

494   ("landscape") which would result in a higher entropy, whereas the right panel has a

495   single prominent value peak which would result in a lower entropy.

496

497

498

Results

VPFC RESPONSE TO ENTROPY OF THE VALUE FUNCTION

Entropy is an emergent property that captures the complexity of the spatially structured value function on a given trial: higher entropy indicates many smaller local value maxima (complex value landscape) and lower entropy typically indicates one prominent global value maximum (simple value landscape; Eq. 8). The higher the value entropy, the more global uncertainty there is about the location of the best option. In exploratory whole-brain voxelwise GLM analyses of Experiment 1 we mapped responses to the entropy of the value function. A cluster responsive to low value entropy in right vPFC (Fig. 3A, black arrow; $p_{FDR} < 0.05$) was centered on the DMN region 24/32, extending into fp10.

To investigate regional responses offline (during the inter-trial interval) vs. online (during navigation and choice) we aligned deconvolved BOLD responses to trial onset and examined the within-trial temporal dynamics (Dombrovski et al., 2020). Extending the GLM results, this analysis revealed divergent representations in DMN vs. CTR and no detectable response in LIM (Fig. 3B). DMN subregions responded positively to lower value entropy (orange, Fig. 3B, $p_{FDR} < 0.001$) while CTR subregions responded positively to higher value entropy (yellow, Fig. 3B; $p_{FDR} < 0.05$). In Experiment 2, we replicated the findings of low value entropy responses in DMN (orange, Fig. 3C, 5 points $p_{FDR} < 0.001$) and additionally detected similar responses in LIM (blue, Fig. 3C, $p_{FDR} < 0.01$). High entropy responses in CTR did not replicate in Experiment 2.

521

522   We then examined the specific contributions of each of the 17 subregions across the

523   three networks (Fig. 1C) derived from the 200 subregion parcellation of Yeo and

524   Schaefer (Schaefer et al., 2018; see Methods). This analysis of Experiment 1 revealed

525   similar responses to low value entropy across DMN subregions 14m25/32 (orange, Fig.

526   3D; p_{FDR} < 0.05), fp10 (p_{FDR} < 0.05), 24/32 (p_{FDR} < 0.05) and d10 (p_{FDR} < 0.05 between -

527   1 and 2s). Responses in CTR were more heterogeneous, with dorsal frontopolar region

528   rl9/10 but not orbitofrontal 11/47 responding to higher entropy (yellow, Fig. 3D; 3

529   timepoints at p_{FDR} < 0.05). In LIM, a response to low value entropy was observed only

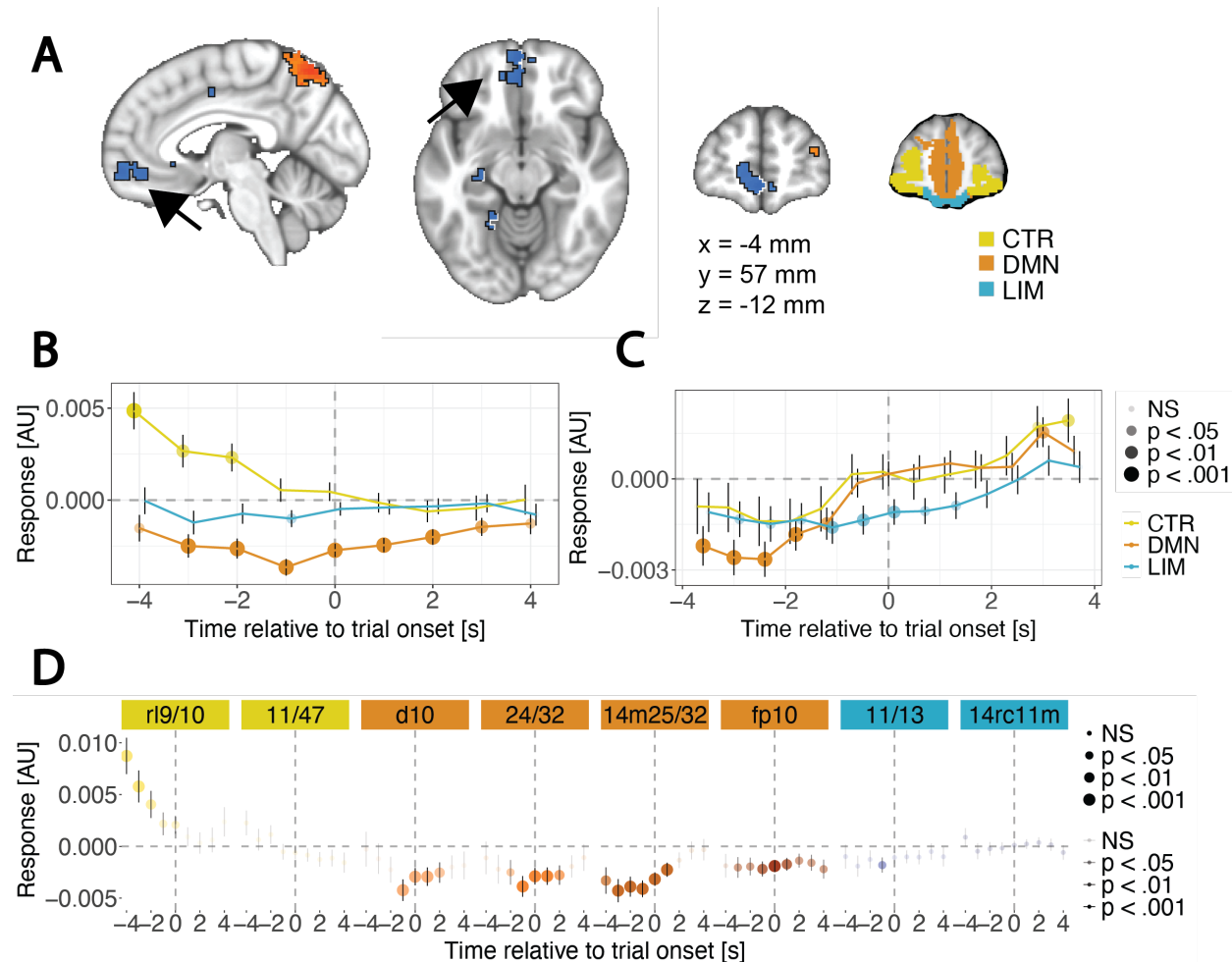530   at a single timepoint in region 11/13.

531

Fig. 3: Responses of the vPFC to value function entropy signals. A. A whole-brain

GLM with a parametric entropy regressor centered on trial onset revealed activation in

vPFC (black arrow; DMN region 24/32 extending into fp10). **B.** Entropy was used as a

predictor of deconvolved vPFC activity grouped by resting-state network using the

MEDuSA method. The DMN (orange) increased activity on lower-value entropy trials,

while the CTR (yellow) network increased its activity (FDR-adjusted p-values indicated

via opacity and marker size). Timepoints < 0 s represent the inter-trial interval beginning

at the offset of previous trial (variable), excluding any volumes from the previous trial.

C. DMN decreases in response to high entropy (orange) was replicated in Experiment

542   2. The increased activity of CTR to entropy was observed at a different segment of the

543   window, and an additional strong response to low entropy in LIM was observed. **D.** All

544   regions in the DMN showed response to low value entropy. Only the rl9/10 sub-region

545   of the CTR network increased its activity to higher value entropy.

546

547   *VPFC RESPONSE TO HIGHEST EXPECTED VALUE (V_{max})*

548   While the preceding analysis focused on the global landscape of the value function,

549   scalar expected value signals are commonly reported in vPFC (Hare et al., 2008; Bartra

550   et al., 2013; Smith et al., 2014; Lopez-Gamundi et al., 2021; Knudsen and Wallis, 2022).

551   Thus, to relate our observations to more traditional scalar value analyses of discrete

552   choice tasks, we examined responses to the model-derived expected value of the best

553   option (V_{max}, Eq. 2, scaled within run to minimize global signal and other run- and

554   subject-level confounds). Our whole-brain voxelwise GLM analysis of Experiment 1

555   revealed V_{max} responses in prefrontal DMN (axial slice, region 24/32; Fig. 4A).

556   Significant clusters also included left ventromedial PFC and hippocampus (Table 1).

557

558   As above, we followed up by examining within-trial time courses across regions using

559   MEDuSA analyses aligned to trial onset. DMN and LIM vPFC networks responded

560   positively to V_{max} (Fig. 4B; DMN: orange line, $p_{FDR} < 0.001$; LIM: blue line, $p_{FDR} < 0.05$)

561   and this pattern replicated in Experiment 2 (Fig. 4C). Responses in CTR were absent in

562   Experiment 1 (Fig. 4B, yellow line) and positive in Experiment 2 (Fig. 4C, yellow line).

563

564   Among vPFC subregions in Experiment 1 (Fig. 5D), DMN 14m25/32 responded most

565   robustly *to* V_{max} ($p_{FDR} < 0.05$). Both LIM subregions 11/3 ($p_{FDR} < 0.05$) and 14rc11m

566   ($p_{FDR} < 0.05$) displayed positive responses. CTR regions displayed divergent responses

567   to V_{max} with region 11/47 responding to positive signals ($p_{FDR} < 0.05$) and region rl9/10

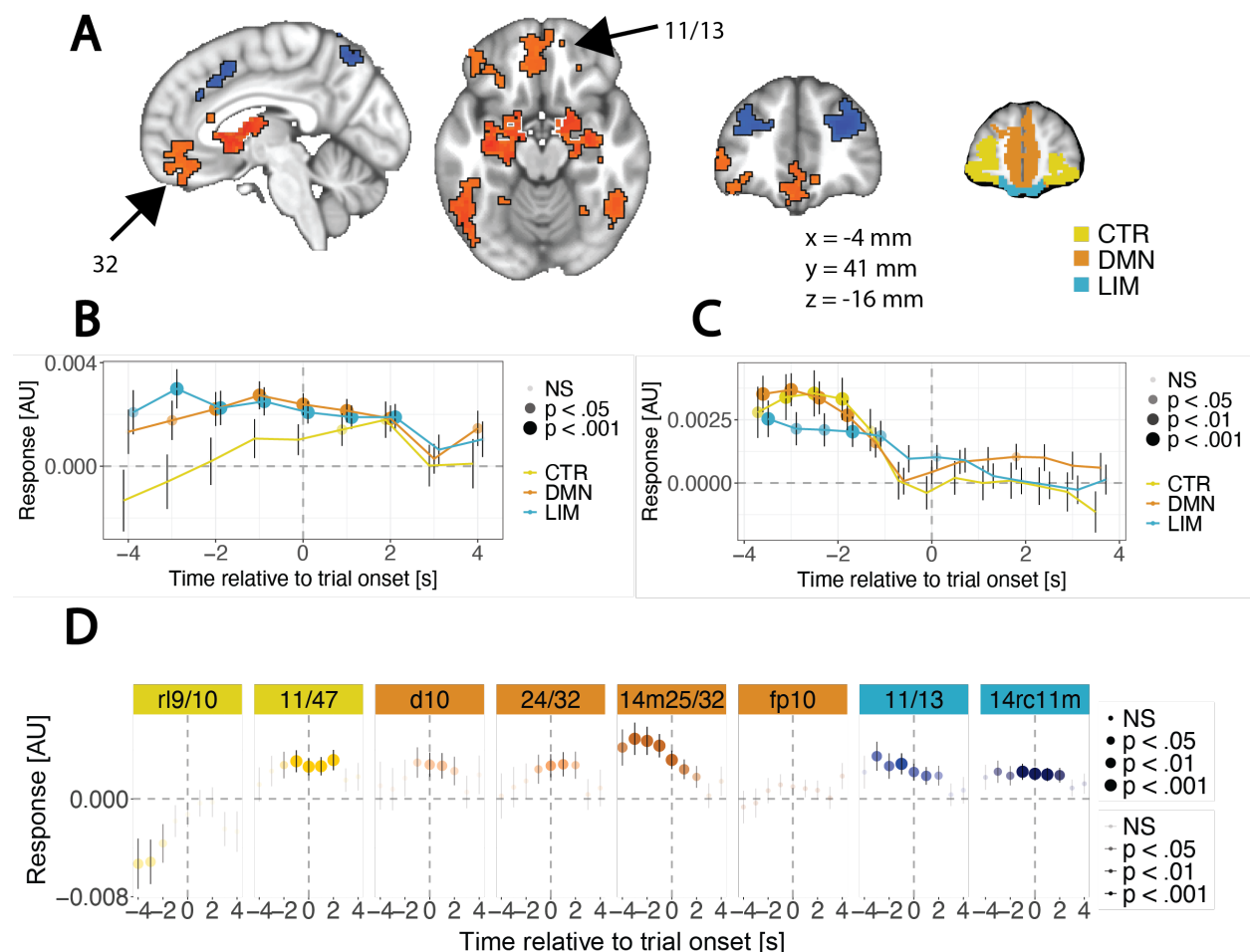568   responding to negative value signals ($p_{FDR} < 0.05$).

569

571   **Fig. 4: Responses of the vPFC to model-predicted value maximum (V_{max}) signals.**

572   **A.** A whole-brain GLM with a parametric value maximum regressor centered on trial

573 onset revealed significant activation at p_{FDR} < 0.05 in both LIM (axial slice arrow, region

574 11/13) and DMN network (sagittal slice arrow, multiple regions centered on region 32)

575 visible in the axial slice. **B.** Model-derived V_{max} was used as a predictor of deconvolved

576 vPFC activity grouped by network using the MEDuSA method. DMN and LIM showed

577 strong responses to scalar value maximum signals. **C.** We replicated our main result in

578 a separate Experiment. DMN (p_{FDR} < 0.01) and LIM (p_{FDR} < 0.05) increased activity to

579 higher value maxima, but these values occurred earlier in the window than in

580 Experiment 1. **D.** V_{max} predicted vPFC activity grouped by region in Experiment 1. In

581 the DMN, region 14m25/32 responded strongly with regions 24/32 and d10 responding

582 less prominently. Both LIM regions responded strongly to value signals. In the CTR

583 network, only region 11/47 increased activity to higher V_{max}, while region rl9/10

584 increased in response to lower model-derived V_{max}.

585

		MNI Coordinates of Peak Activity		
	N voxels	X	Y	Z
Areas Sensitive to High Global Value Maxima (V_{max})				
bilateral hippocampus, left insula, thalamus	1445	32.8	12.4	-18.2
Left inferior temporal gyrus (Area PH and fusiform face complex)	507	48.9	60.7	-11.3
Left ventromedial prefrontal cortex (areas 10, 24, 32)	190	9.8	-38.2	-15.9
Right inferior temporal gyrus (Areas PH, TE2)	167	-47.7	56.1	-15.9
Left inferior frontal gyrus (IFS and area 45)	130	46.6	-35.9	2.5
Areas Sensitive to Low Global Value Maxima (V_{max})				
Bilateral superior frontal gyrus (Area 6)	2031	-24.7	7.8	50.8
Bilateral medial intraparietal area (IPA) and Lateal Area 7P	1397	-20.1	67.6	55.4
Right middle frontal gyrus (area 46)	552	-29.3	-40.5	25.5

586

587 **Table 1: Whole-brain GLM clusters significantly activated by the global value**

588 **maximum in Experiment 1.** Clusters of voxels (voxel size = 2.3mm^3) significantly

589   modulated by model-derived value maxima (V_{max}) in whole-brain analyses of

590   Experiment 1. Note: Clusters of significant voxels were identified based on a voxelwise

591   threshold of p < 0.001, with familywise error correction of p < 0.05. Clusters relevant to

592   the present study are in bold.

593

594   *HIPPOCAMPAL VALUE RESPONSE AND HIPPOCAMPAL-vPFC CONNECTIVITY*

595   We previously reported that the hippocampus, particularly its anterior divisions,

596   responded to low-entropy, easy-to-exploit value landscapes, and this response scaled

597   with behavioral exploitation (Dombrovski et al., 2020). Here, we further consider

598   spatiotemporal patterns of hippocampal activity and connectivity with vPFC, including

599   how connectivity is modulated by entropy and V_{max}, processes that control the explore-

600   exploit balance.

601

602   In whole-brain GLM (Table 1) and deconvolved signal analyses aligned to trial onset

603   (Fig. 5A) both the anterior and posterior hippocampus responded to V_{max} (all p_{FDR} <

604   0.001). This pattern was strongly replicated in Experiment 2 (Fig. 5B).

605

606   To assess functional connectivity, we regressed vPFC activity on contemporaneous

607   hippocampal activity (i.e., at the same moment within the trial). In Experiment 1, AH

608   (dark gray panel) was more strongly functionally coupled to DMN compared to PH (light

609   gray panel), and a clear gradient of hippocampal coupling was observed with DMN >

610   CTR > LIM (Fig. 4C). In Experiment 2, results were qualitatively similar, though

611   functional coupling between hippocampus and vPFC-CTR increased following trial

612   onset (Fig. 6D).

613

614   Activity patterns (Fig. 4B-C) led us to expect stronger AH-DMN connectivity during

615   periods of low entropy. Contrary to this expectation, hippocampal-DMN functional

616   coupling was increased during periods of higher entropy in both AH (dark gray) and PH

617   in both experiments (light gray; Fig. 5E-F). Further, in Experiment 2 only, connectivity

618   modulation was seen only during the online portion of the task, post trial onset and both

619   CTR- and LIM-hippocampal connectivity also scaled up with entropy (Fig. 5F). V_{max}

620   modulated hippocampal connectivity in a similar pattern across both experiments (Fig.

621   5G-H).

622

623   In summary, the highest levels of connectivity with hippocampus were seen in prefrontal

624   DMN regions, in anterior hippocampus, and when decisions were difficult, as indicated
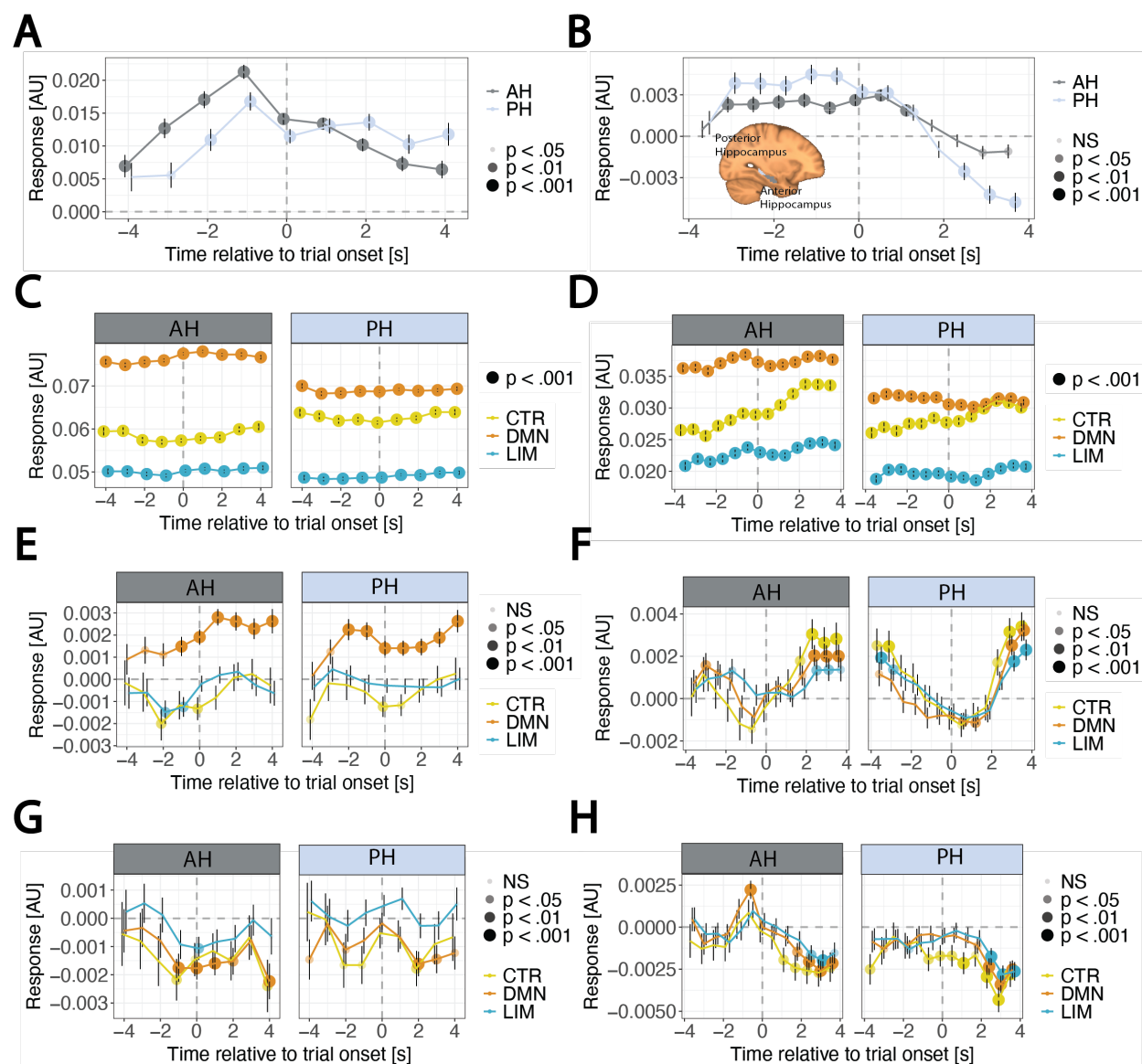
625   by higher entropy or lower V_{max}.

626

627

Fig. 5: Hippocampus responds to Vmax and is selectively synchronized with DMN-vPFC, particularly during periods of higher entropy. **A.** Model-derived V_{max} was used as a predictor of deconvolved hippocampal activity in anterior (dark grey) and posterior (light gray) hippocampus using the MEDuSA method. Both hippocampal sub-regions responded to V_{max}. **B.** V_{max} responses in hippocampus replicated out-of-sample in Experiment 2. **C.** Hippocampal activity was averaged within region (AH or PH), scaled within subject and run, and then used as a regressor to predict vPFC activity in

636   Experiment 1. AH had higher connectivity with DMN compared to PH. Connectivity

637   generally followed a gradient of DMN > CTR > LIM. **D.** Hippocampal activity was used

638   as a regressor to predict vPFC activity in Experiment 2 following the same procedure.

639   Overall, hippocampal-vPFC connectivity was diminished, but the same gradient of DMN

640   > CTR > LIM remained and AH connectivity was greater than PH connectivity. **E.**

641   Interacting hippocampal signals with entropy for Experiment 1 suggests that DMN is

642   most strongly influenced by hippocampus at times when entropy is higher. **F.** Anterior

643   Hippocampal (AH) activity affects DMN activity more strongly when entropy is higher in

644   Experiment 2 following trial onset, replicating the results out-of-sample. In addition, in

645   this population, hippocampal activity affects CTR more strongly when entropy is higher.

646   **G.** Interacting hippocampal signals with V_{max} in Experiment 1 suggests that DMN is most

647   strongly influenced by anterior hippocampus at times when V_{max} is lower. **H.** In

648   Experiment 2, anterior hippocampus influences DMN more strongly during trials when

649   V_{max} is lower, replicating this finding out-of-sample. In addition, in Experiment 2, the

650   entire vPFC is synched with hippocampus activity at lower V_{max}, notably during a later

651   time in the window.

652

653   *HIPPOCAMPAL-vPFC INTERACTIONS AND BEHAVIORAL EXPLORATION*

654   Normatively, high entropy of the value function favors exploration. Surges in

655   hippocampal-prefrontal DMN connectivity during difficult high-entropy and low-value

656   epochs could thus support behavioral exploration, which we previously found to be

657   related to posterior hippocampal learning signals (Dombrovski et al., 2020). Hence, we

658   examined whether value-related increases in PH-prefrontal DMN connectivity were

659   selectively related to exploration. The temporal resolution of fMRI is insufficient for

660   testing momentary connectivity-behavior relationships, so we used session-level

661   estimates of connectivity modulation to predict behavioral exploration both within-

662   session and out-of-session. We derived session-level estimates of connectivity

663   modulation from the models described above (Eq. 9, Fig. 3, Fig. 5) by extracting

664   individual random slopes for the hippocampal response-by-entropy interaction effect on

665   prefrontal activity and averaged them across timepoints within trial to reduce multiple

666   comparisons. These predictors, one for each subject and each region, were then

667   entered into the behavioral model predicting exploration (Eq. 10). As in previous

668   studies, exploration of the continuous space was measured by large departures from a

669   previous response ("RT swings", Fig. 2D, left panel (Badre et al., 2012; Hallquist and

670   Dombrovski, 2019; Hallquist et al., 2023). While this analysis controlled for long-term

671   value (exploitation) effects, one must also account for the strong win-stay/lose-shift

672   response tendency prominent in humans and other species (Zhang et al., 2023; Randall

673   and Zentall, 1997). Thus, post-reward RT swings provide a purer measure of

674   exploration as defined in RL.

675

676   Indeed, stronger modulation of PH-DMN and PH-LIM connectivity by value entropy

677   scaled with exploratory post-reward RT swings (Fig. 6A) in-session (Experiment 1, p_{FDR}

678   < 10^{-5}), out-of-session (MEG replication, p_{FDR} = 0.03) and out-of-sample (Experiment 2,

679   p_{FDR} = < 10^{-5}). Main effects were significant for LIM-PH in Experiment 1 (p_{FDR} = 2.3e-

680   19), Experiment 1 out-of-session MEG replication (p_{FDR} = 1.2e-15), and Experiment 2

681   (p_{FDR} = 6.7e-4). Recall that, as Fig. 5E-F illustrates, only PH-DMN, but not PH-LIM

682   connectivity was consistently modulated by entropy across entire samples, qualifying

683   the interpretation of PH-LIM interaction behavioral correlates. Significant post-hoc

684   contrasts are indicated by non-overlapping confidence intervals and consistent effects

685   across all three experimental datasets are indicated by higher opacity. Here, p-values

686   with false discovery rate correction are reported from the entropy connectivity random

687   slope by lagged RT by last outcome term ($\gamma_t^1\ x_s x_{RT(t-1)_t} x_{R_t}$ in Eq. 10).

688

689   To ascertain the anatomical specificity of the relationship between PH-DMN connectivity

690   and exploration, we repeated the above analysis for AH-vPFC connectivity (Fig. 6B).

691   While modulation of AH-DMN by value entropy was not related to win-shift exploration,

692   it predicted smaller lose-shifts. Further, stronger modulation of AH-LIM connectivity by

693   value entropy predicted win-shift and, to a lesser degree, lose-shift responses.

694

695   Modulation of hippocampal-vPFC connectivity by scalar V_{max} showed largely non-

696   significant exploitation effects (not reported) suggesting altogether that the critical

697   hippocampal-vPFC interactions signal a spatially structured value map and not merely a

698   scalar V_{max}.

699

700   In summary, as suggested by connectivity analyses (Fig. 5E, F), value-related PH-DMN

701   scaled with behavioral exploration. Unexpectedly, the same was true for PH-LIM and

702   AH-LIM interactions.

703

HIPPOCAMPAL-vPFC INTERACTIONS AND BEHAVIORAL EXPLOITATION

Our prior findings (Dombrovski et al., 2020) led us to predict that value-related DMN and LIM activity and connectivity with AH would scale with behavioral exploitation. However, we failed to find a replicable general pattern, observing instead only one or the other study correlated with activation. Stronger negative entropy modulation of DMN activity scaled with behavioral exploitation in Experiment 1 (p_{FDR} = 0.01, Figure 6C), replicating out-of-session replication (p_{FDR} = 0.01), but not out-of-sample (p_{FDR} = 0.64). We found no evidence specifically implicating AH vs. PH prefrontal connectivity in exploitation using these analyses. P-values with false discovery rate correction are reported from the entropy connectivity random slope by lagged RT$_{Vmax}$ term ($\gamma_t^2 x_s x_{RT_{Vmax(t-1)_t}}$ in Eq. 10).
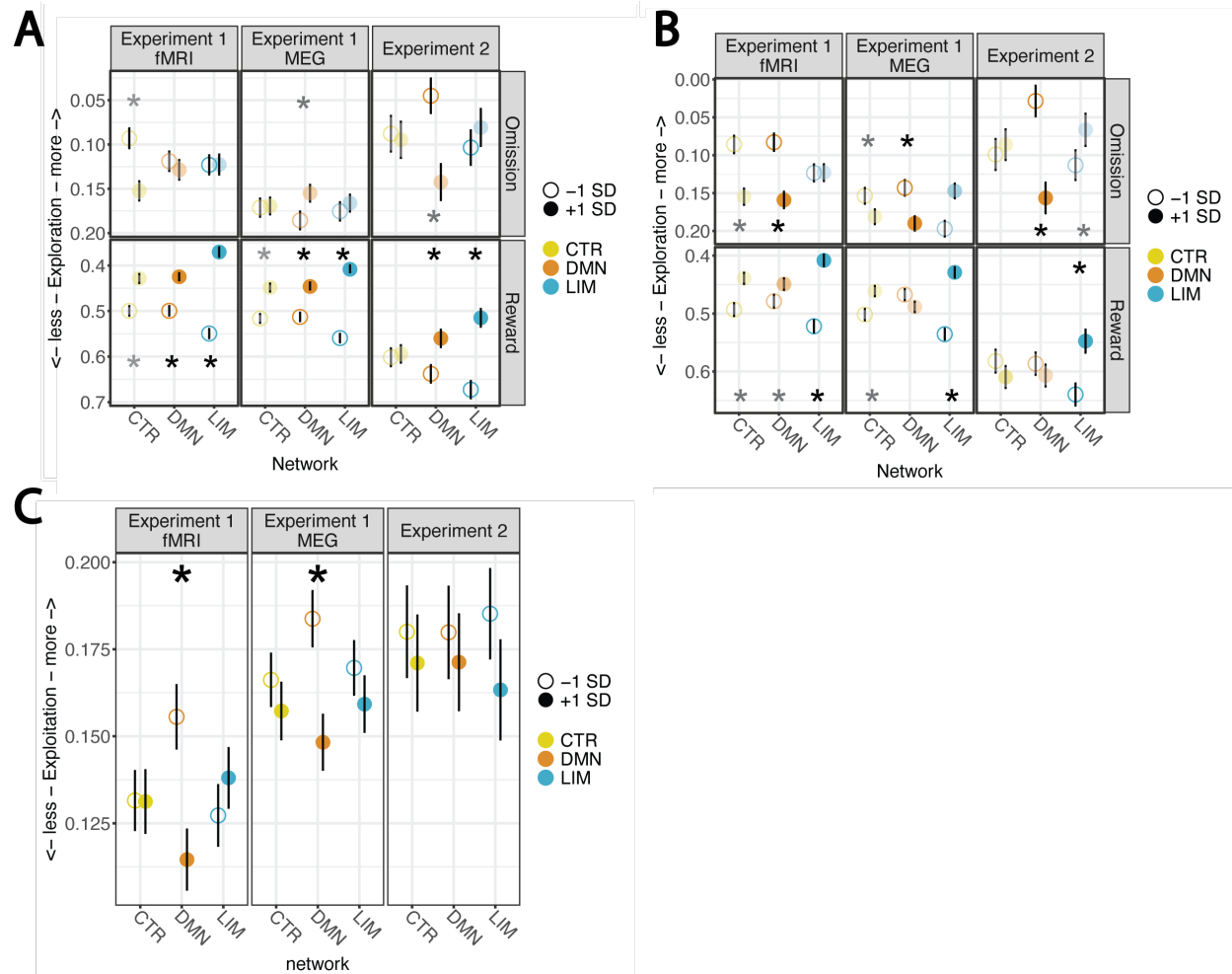
715

Fig. 6: Entropy modulation of vPFC-HC connectivity predicts exploration. The subject-level effects of entropy modulation of vPFC-hippocampus connectivity that predict exploratory behavior of subjects is shown for three datasets. Consistent effect directions across all three datasets are indicated by colored boxes, while significant main effects are indicated by asterisks (p_{FDR} reported in text). Higher entropy (one standard deviation above the mean) is indicated by filled colored circles, while lower entropy (one standard deviation below the mean) is indicated by open circles. Note that the y-axes are reversed such that lower autocorrelation between lagged RT and upcoming RT are higher (more Exploration is indicated by lower autocorrelation). For

726   panels A and B, effects that are consistent across all three Experiments are indicated

727   with full opacity, while those that do not replicate are indicated with partial opacity. **A.**

728   When entropy modulation of DMN- and LIM- posterior hippocampal connectivity was

729   higher following reward on the prior trial, subjects engaged in more exploratory

730   behaviors. This win-shift behavior had a significant main effect across all three

731   datasets; thus, we focus on this in the text and discussion. **B.** When entropy modulation

732   of DMN-anterior hippocampus activity was higher following omissions on the prior trial,

733   subjects tended to explore less. This lose-stay behavior was significant across all three

734   datasets. Additionally, when entropy modulation of LIM-anterior hippocampus activity

735   was higher following rewards, subjects engaged in more exploration. This win-shift

736   behavior was significant in Experiment 1 and out-of-session MEG replication session of

737   Experiment 1, but showed a non-significant trend in Experiment 2. **C.** The subject-level

738   effects of entropy modulation of vPFC that predict exploitative behavior of subjects is

739   shown for three datasets. Higher entropy (one standard deviation above the mean) is

740   indicated by filled colored circles, while lower entropy (one standard deviation below the

741   mean) is indicated by open circles. In Experiment 1 and the out-of-session MEG

742   replication session of Experiment 1, lower entropy modulation of DMN predicts

743   exploitation towards the RT$_{Vmax}$.

744

745   *SENSITIVITY ANALYSES*

746   Since we assigned the right fp10 region to DMN instead of LIM, we verified that this

747   departure from the previously published parcellation (Thomas Yeo et al., 2011) did not

748   change our results. Reassigning right fp10 back to LIM, findings of DMN activity

749   modulation by low value entropy (Fig. 4B) were qualitatively unchanged (Fig. 7A), as

750   were those for higher V_{max} (Fig. 3B, Fig. 7B). Similarly, connectivity patterns reported in

751   Fig. 5C remained qualitatively unchanged (Fig. 7C) and modulation of connectivity by

752   entropy (Fig. 5E) remained qualitatively unchanged (Fig.7D).

753

754   Finally, the literature is divided on whether the temporal resolution of fMRI is sufficient

755   for judging the direction of effective connectivity (Barnett and Seth, 2017; Stokes and

756   Purdon, 2017). Still, to ensure that our connectivity results were not explained by

757   temporally slower influences of one region on the other, we repeated our analyses of

758   vPFC-hippocampal connectivity, including both concurrent and *lagged* (by one TR)

759   hippocampal activity. The specific pattern of connectivity reported in Fig. 5C was

760   disrupted (Fig. 7E), but the general significance of connectivity across the time window

761   held. When interacting this divergent pattern of lagged connectivity modulation by value

762   entropy (Fig. 5E vs. Fig. 7F), the general pattern of higher connectivity with DMN during

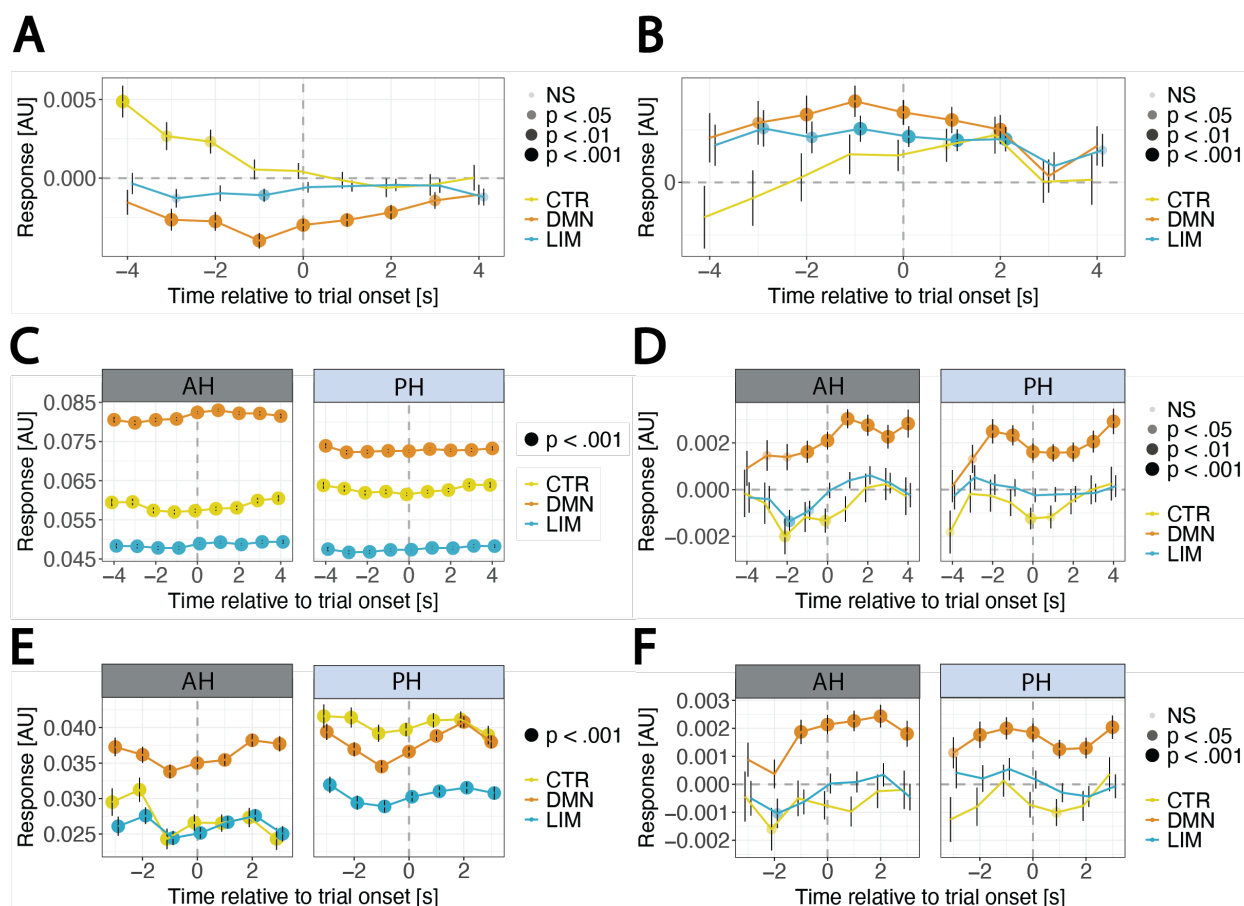763   higher value entropy held.

764

Fig. 7 Sensitivity Analyses. In the original 300 parcel, 7 network classification (Schaefer et al., 2018), the region we labeled 'right fp10' was assigned to the LIM network instead of the DMN. Here, we check whether re-assigning this parcel to LIM does not alter our major conclusions. **A.** Fig. 3B remains qualitatively unchanged (note a single LIM timepoint at -1s now survives correction at $p_{FDR} < 0.05$) when reassigning right fp10 from DMN to LIM. **B.** Fig. 4B remains qualitatively unchanged (note a slightly smaller effect of V_{max} on LIM) when reassigning right fp10 from DMN to LIM. **C.** In our functional connectivity model, LIM network correlations with AH were slightly increased and DMN network connectivity with AH was slightly decreased when fp10 was assigned to LIM (compare to Fig. 5C). **D.** Fig. 5E remains qualitatively unchanged after the fp10

776   parcel is moved to LIM. **E.** Autocorrelation of regressors in time series analysis can lead

777   to spurious significance being detected. To control for this possibility in our vPFC-

778   hippocampal connectivity and psychophysiological interaction analysis, we included the

779   lagged hippocampal signal in our models, controlling for contemporaneous activity of

780   hippocampus. The conclusions of Fig. 5C regarding the gradient of vPFC-hippocampal

781   connectivity are not preserved, but, critically, all points in the connectivity analysis

782   remain significant. **F.** Our finding that hippocampal modulation of DMN is stronger

783   when entropy is higher still holds when accounting for an autocorrelated hippocampal

784   signal, interacting the lagged hippocampal signal with entropy.

785   **Discussion** (1419 words)

786   Our study examined the evolving distribution of value across a space modulating vPFC

787   responses and interactions with hippocampus as humans explored and exploited a

788   continuous space. We combined an information-compressing reinforcement learning

789   (RL) model with fMRI to examine how responses to value information (entropy)

790   dynamics in the vPFC and hippocampus shaped the transition from exploration to

791   exploitation. Higher entropy indicates many smaller local value maxima; while lower

792   entropy typically indicates one prominent global value maximum. Entropy

793   representations were more richly captured in pericallosal DMN vPFC subregions than

794   LIM or CTR subregions. Unexpectedly, DMN vPFC-hippocampus functional

795   connectivity increased when people faced uncertain (high entropy, low V_{max}) value

796   landscapes, and this increase, primarily involving DMN-PH connectivity, scaled with

797   behavioral exploration. These two patterns were observed across sessions, samples

798   and experimental designs. In conclusion, vPFC and hippocampal BOLD activity

799 increases when values are higher and entropy is lower, favoring exploitation. In

800 contrast, hippocampal-vPFC DMN connectivity increases under conditions favoring

801 exploration, when many similarly-valued options compete.

802

803

804 *VALUE REPRESENTATIONS IN NON-HUMAN PRIMATE AND HUMAN VPFC*

805 Why might vPFC subregions of the DMN show greater sensitivity to complex

806 landscapes of learned value? The most straightforward conclusion of our study is that

807 human prefrontal DMN – rather than limbic OFC – contains richer spatial value

808 representations. Monkey lesion and electrophysiological studies of value

809 representations and prefrontal-hippocampal interactions during decision-making mostly

810 converge on the limbic cOFC/mOFC, particularly areas 11, 13 and 14 (Padoa-Schioppa

811 and Assad, 2008; Rolls, 2008; Rudebeck and Murray, 2011; Rich and Wallis, 2014;

812 Pastor-Bernier et al., 2019). By contrast, human imaging studies find value signals in

813 the pericallosal vPFC falling generally into DMN (Kable and Glimcher, 2007; Hare et al.,

814 2008; Bartra et al., 2013; Lim et al., 2013; Vassena et al., 2014; Lopez-Gamundi et al.,

815 2021), which has also become the focus of studies of prefrontal-hippocampal

816 connectivity during decision-making (Gerraty et al., 2014; Gluth et al., 2015; Schuck and

817 Niv, 2019).

818

819 Our computational model highlights a related factor potentially responsible for this

820 divergence: whether values are represented at option vs. entire environment level.

821   While responses to the highest-value option occurred throughout limbic OFC and

822   pericallosal DMN, responses to the entire value landscape were restricted to DMN,

823   particularly area 14m/24/25 and frontopolar dorsal area 10. Both areas, particularly the

824   DMN, synchronized with the hippocampus when value landscapes were complex.

825   Although human vPFC DMN lacks a direct non-hominoid homologue (Mantini et al.,

826   2011, 2012; Neubert et al., 2015; Garin et al., 2022), our results agree with studies

827   finding preferential medial frontal (Giarrocco and Averbeck, 2021) and DMN (Kaplan et

828   al., 2016) functional connectivity with the hippocampus. Strong scalar value responses

829   in the limbic OFC also dismiss a trivial explanation for the divergence between monkey

830   and human studies—that limbic OFC BOLD is obscured by signal dropout. Indeed,

831   mOFC DMN area 14m/25-32 is equally, if not more, affected by dropout compared e.g.

832   to limbic mOFC area 11/13. Overall, our results suggest that value representations in

833   vPFC DMN, as opposed to limbic OFC, are integrated across the entire environment

834   and more enriched with spatial information from the hippocampus.

835

836   *REWARD REPRESENTATIONS WITHIN HIPPOCAMPAL MAPS*

837   Whereas vPFC value representations have been studied in detail, less is known about

838   hippocampal value encoding. Adding to prior evidence of value responses in the human

839   hippocampus (Zeithamova et al., 2018; Bakkour et al., 2019) and our earlier observation

840   that the spatially structured value function selectively modulates AH vs. PH (Dombrovski

841   et al., 2020), here we describe scalar maximum value representations throughout the

842   hippocampal long axis (Fig. 5A-B). In line with our observations, Yun and colleagues

843   (2023) found that rodent dorsal (human posterior) hippocampus responses scaled with

844   discrete-option scalar value and ventral (human anterior) hippocampus represented the

845   maximum value option in a set during a cued licking task. Altogether, hippocampal

846   value representations appear to follow the same fine-coarse postero-anterior (dorso-

847   ventral) gradient as purely spatial representations. If so, one unanswered question is

848   whether this architecture extends to abstract, non-spatial problems. In a recent study,

849   navigation through an abstract space induced value representations in monkey

850   hippocampus (Knudsen and Wallis, 2021), providing evidence of a value-laden

851   cognitive map in a non-spatial context. While the rodent hippocampus is modulated by

852   value-related factors (Wikenheiser and Redish, 2015; Sun et al., 2020), rodent studies

853   have yet to provide evidence that hippocampal maps incorporate value signals in

854   abstract, non-spatial paradigms.

855

856   *INTERACTIONS BETWEEN HIPPOCAMPUS AND vPFC*

857   Whereas we anticipated that vPFC and hippocampus would synchronize when people

858   faced easy-to-exploit value landscapes, they instead synchronized when people faced

859   lower values and multiple competing options demanding exploration. Further, while

860   baseline functional connectivity with DMN was stronger in AH than in PH, consistent

861   with resting-state observations (Barnett et al., 2021), it was the synchronization of

862   vPFC-DMN with PH, rather than AH, that scaled with exploratory behavior. This aligns

863   with observations that PH (rodent dorsal hippocampus) represents stimuli during early

864   learning (Biane et al., 2023) and PH/DH learning signals scale with behavioral

865   exploration (Kheirbek et al., 2013; Dombrovski et al., 2020). Further, reward-selective

866   cells in deep DH layers project to the PFC (Harvey et al., 2023), and PH-prefrontal

867   interactions support human memory encoding (Spellman et al., 2015) and retrieval

868   (Rajasethupathy et al., 2015; Ezama et al., 2021).

869

870   Although our fMRI analyses do not speak to the electrophysiological mechanisms of

871   PH-prefrontal DMN interactions during exploration, rodent electrophysiological studies

872   offer a basis for speculation. Prefrontal-PH CA1 synchronization involves theta

873   sequences predicting upcoming locations during navigation ("look-ahead") and sharp

874   wave ripples (SWR) subserving trajectory replay offline, during the inter-trial period

875   (Tang et al., 2021; Wikenheiser and Redish, 2015; Buzsáki, 2015; Norman et al., 2019;

876   Liu et al., 2022). While care should be taken when relating BOLD to electrophysiological

877   responses (Ekstrom et al., 2009; Ekstrom, 2010; Hill et al., 2021), in our analyses of

878   deconvolved BOLD signal value-related PH-DMN interactions extend into the intertrial

879   period. This observation aligns with findings of replay-related interactions between

880   human hippocampus and vPFC interactions (Schuck and Niv, 2019) and broader

881   evidence that key DMN computations and interactions with hippocampus occur offline

882   and during recall (Kaplan et al., 2016; Norman et al., 2021). Thus, when we face

883   multiple competing options, our PH and prefrontal DMN may be jointly simulating

884   exploration trajectories, experienced or contemplated. It is more difficult to interpret the

885   general lack of relationships of value-driven activity and synchronization with

886   exploitation (although in Experiment 1, DMN entropy modulation did predict exploitation;

887   Fig. 6C). It is possible that contributions of hippocampal representations to behavioral

888   exploitation do not depend on coupling with vPFC, but rather on coupling with the dorsal

889   stream network, which we have found to facilitate exploitation on the clock task

890   (Hallquist et al., 2024).

891

892   Under the SCEPTIC model, infrequently sampled locations are compressed out of the

893   option set, decreasing value function entropy as one transitions from exploration to

894   exploitation (Hallquist and Dombrovski, 2019). Empirically, vPFC-HC connectivity

895   peaked when entropy was high (Fig. 5E-F), and this modulation scaled with exploratory

896   behavior (Fig. 6A-B). Thus, vPFC-HC coupling may support the compression of an

897   overwhelmingly complex value-laden map to a more manageable and exploitable one.

898   Such plasticity is consistent the broader contributions of the vPFC and hippocampus to

899   behavioral flexibility (Winocur and Olds, 1978; Hornak et al., 2004; Vilà-Balló et al.,

900   2017; McLaughlin and Redish, 2023).

901

902   *CONCLUSIONS*

903   Replication across sessions and studies—spanning samples of younger, middle-aged,

904   and older adults, tasks with and without reversals, and scanning sequences—

905   strengthens confidence in our main findings. At the same time, we cannot say which of

906   the differences between studies explain the divergences in the LIM response to entropy

907   (compare Fig. 3B, 3C), timing of V_{max} signals (compare Fig. 4B, 4C), and timing of

908   entropy modulation of hippocampal-DMN connectivity (compare Fig. 5E, 5F). Without

909   direct manipulation of hippocampal and prefrontal activity, we cannot be confident that

910   hippocampal-vPFC interactions are necessary for behavioral exploration, despite

911   controlling for behavioral confounds, inter-trial interval, age, and sex. The observational

912   study design and low temporal resolution of fMRI make it hard to determine the direction

913   of information transfer during hippocampal-vPFC interactions.

914

915   In summary, vPFC and hippocampal BOLD activity increases when the values of

916   dominant options are high, favoring exploitation. Conversely, hippocampal-vPFC DMN

917   connectivity rises under precisely the opposite conditions, when people explore multiple

918   competing options of relatively low value. Coordinated activity of the posterior

919   hippocampal-prefrontal DMN system updates detailed value-laden maps that guide

920   exploration. Given the connection between exploration in the laboratory and human

921   psychopathology, including suicidal behavior (Tsypes et al., 2024), our findings may

922   also inform our understanding of the role of PH-DMN and AH-LIM interactions in real-life

923   decisions at longer timescales and with higher stakes.

References

Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, Gramfort A, Thirion B, Varoquaux G (2014) Machine learning for neuroimaging with scikit-learn. Front Neuroinformatics 8 Available at: https://www.frontiersin.org/articles/10.3389/fninf.2014.00014 [Accessed May 5, 2023].

Badre D, Doll BB, Long NM, Frank MJ (2012) Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. Neuron 73:595–607.

Bakkour A, Palombo DJ, Zylberberg A, Kang YH, Reid A, Verfaellie M, Shadlen MN, Shohamy D (2019) The hippocampus supports deliberation during value-based decisions Kahnt T, Frank MJ, Fellows LK, Gluth S, eds. eLife 8:e46080.

Barnett AJ, Reilly W, Dimsdale-Zucker HR, Mizrak E, Reagh Z, Ranganath C (2021) Intrinsic connectivity reveals functionally distinct cortico-hippocampal networks in the human brain. PLOS Biol 19:e3001275.

Barnett L, Seth AK (2017) Detectability of Granger causality for subsampled continuous-time neurophysiological processes. J Neurosci Methods 275:93–121.

Bartra O, McGuire JT, Kable JW (2013) The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. Neuroimage 76:412–427.

943   Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models
944   Using lme4. J Stat Softw 67:1–48.

945   Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing
946   under dependency. Ann Stat 29:1165–1188.

947   Biane JS, Ladow MA, Stefanini F, Boddu SP, Fan A, Hassan S, Dundar N, Apodaca-
948   Montano DL, Zhou LZ, Fayner V, Woods NI, Kheirbek MA (2023) Neural
949   dynamics underlying associative learning in the dorsal and ventral hippocampus.
950   Nat Neurosci:1–12.

951   Brainard, D.H. (1997) The Psychophysics Toolbox. Spat Vis 10:433–436.

952   Bush K, Cisler J (2013) Decoding neural events from fMRI BOLD signal: A comparison
953   of existing approaches and development of a new algorithm. Magn Reson
954   Imaging 31:976–989.

955   Buzsáki, G. (2006) Rhythms in the Brain. Oxford University Press, USA. Available at:
956   https://academic.oup.com/book/11166.

957   Buzsáki G (2015) Hippocampal sharp wave-ripple: A cognitive biomarker for episodic
958   memory and planning. Hippocampus 25:1073–1188.

959   Chudasama Y, Doobay VM, Liu Y (2012) Hippocampal-Prefrontal Cortical Circuit
960   Mediates Inhibitory Response Control in the Rat. J Neurosci 32:10915–10924.

961   Costa VD, Mitz AR, Averbeck BB (2019) Subcortical Substrates of Explore-Exploit
962   Decisions in Primates. Neuron 103:533-545.e5.

963   Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic

964   resonance neuroimages. Comput Biomed Res 29:162–173.

965   Dale AM, Fischl B, Sereno MI (1999) Cortical Surface-Based Analysis: I. Segmentation

966   and Surface Reconstruction. NeuroImage 9:179–194.

967   Daunizeau J, Adam V, Rigoux L (2014) VBA: A Probabilistic Treatment of Nonlinear

968   Models for Neurobiological and Behavioural Data. PLOS Comput Biol

969   10:e1003441.

970   Dombrovski AY, Luna B, Hallquist MN (2020) Differential reinforcement encoding along

971   the hippocampal long axis helps resolve the explore–exploit dilemma. Nat

972   Commun 11:5407.

973   Domenech P, Rheims S, Koechlin E (2020) Neural mechanisms resolving exploitation-

974   exploration dilemmas in the medial prefrontal cortex. Science 369:eabb0184.

975   Eickhoff SB, Paus T, Caspers S, Grosbras M-H, Evans AC, Zilles K, Amunts K (2007)

976   Assignment of functional activations to probabilistic cytoarchitectonic areas

977   revisited. NeuroImage 36:511–521.

978   Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005) A

979   new SPM toolbox for combining probabilistic cytoarchitectonic maps and

980   functional imaging data. NeuroImage 25:1325–1335.

981   Ekstrom A (2010) How and when the fMRI BOLD signal relates to underlying neural

982   activity: The danger in dissociation. Brain Res Rev 62:233–244.

983   Ekstrom A, Suthana N, Millett D, Fried I, Bookheimer S (2009) Correlation Between

984   BOLD fMRI and Theta-Band Local Field Potentials in the Human Hippocampal

985   Area. J Neurophysiol 101:2668–2678.

986   Ezama L, Hernández-Cabrera JA, Seoane S, Pereda E, Janssen N (2021) Functional

987   connectivity of the hippocampus and its subfields in resting-state networks. Eur J

988   Neurosci 53:3378–3393.

989   Fonov V, Evans A, McKinstry R, Almli C, Collins D (2009) Unbiased nonlinear average

990   age-appropriate brain templates from birth to adulthood. NeuroImage 47:S102.

991   Garin CM, Hori Y, Everling S, Whitlow CT, Calabro FJ, Luna B, Froesel M, Gacoin M,

992   Ben Hamed S, Dhenain M, Constantinidis C (2022) An evolutionary gap in

993   primate default mode network organization. Cell Rep 39:110669.

994   Gerraty RT, Davidow JY, Wimmer GE, Kahn I, Shohamy D (2014) Transfer of Learning

995   Relates to Intrinsic Connectivity between Hippocampus, Ventromedial Prefrontal

996   Cortex, and Large-Scale Networks. J Neurosci 34:11297–11303.

997   Giarrocco F, Averbeck BB (2021) Organization of parietoprefrontal and

998   temporoprefrontal networks in the macaque. J Neurophysiol 126:1289–1309.

999   Gluth S, Sommer T, Rieskamp J, Büchel C (2015) Effective Connectivity between

1000   Hippocampus and Ventromedial Prefrontal Cortex Controls Preferential Choices

1001   from Memory. Neuron 86:1078–1090.

1002 Gorgolewski K, Burns C, Madison C, Clark D, Halchenko Y, Waskom M, Ghosh S

1003 (2011) Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data

1004 Processing Framework in Python. Front Neuroinformatics 5 Available at:

1005 https://www.frontiersin.org/articles/10.3389/fninf.2011.00013 [Accessed January

1006 23, 2023].

1007 Greve DN, Fischl B (2009) Accurate and robust brain image alignment using boundary-

1008 based registration. NeuroImage 48:63–72.

1009 Hallquist M, Hwang K, Luna B, Dombrovski AY (2023) Reinforcement-based option

1010 competition in human dorsal stream during exploration/exploitation of a

1011 continuous space. :2023.05.22.541828 Available at:

1012 https://www.biorxiv.org/content/10.1101/2023.05.22.541828v1 [Accessed May

1013 23, 2023].

1014 Hallquist MN, Dombrovski AY (2019) Selective maintenance of value information helps

1015 resolve the exploration/exploitation dilemma. Cognition 183:226–243.

1016 Hallquist MN, Hwang K, Luna B, Dombrovski AY (2024) Reward-based option

1017 competition in human dorsal stream and transition from stochastic exploration to

1018 exploitation in continuous space. Sci Adv.

1019 Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A (2008) Dissociating the Role

1020 of the Orbitofrontal Cortex and the Striatum in the Computation of Goal Values

1021 and Prediction Errors. J Neurosci 28:5623–5630.

1022   Harvey RE, Robinson HL, Liu C, Oliva A, Fernandez-Ruiz A (2023) Hippocampo-

1023   cortical circuits for selective memory encoding, routing, and replay.

1024   Neuron:S0896627323003008.

1025   Henssen A, Zilles K, Palomero-Gallagher N, Schleicher A, Mohlberg H, Gerboga F,

1026   Eickhoff SB, Bludau S, Amunts K (2016) Cytoarchitecture and probability maps

1027   of the human medial orbitofrontal cortex. Cortex 75:87–112.

1028   Hill PF, Seger SE, Yoo HB, King DR, Wang DX, Lega BC, Rugg MD (2021) Distinct

1029   Neurophysiological Correlates of the fMRI BOLD Signal in the Hippocampus and

1030   Neocortex. J Neurosci 41:6343–6352.

1031   Hornak J, O'Doherty J, Bramham J, Rolls ET, Morris RG, Bullock PR, Polkey CE (2004)

1032   Reward-related Reversal Learning after Surgical Excisions in Orbito-frontal or

1033   Dorsolateral Prefrontal Cortex in Humans. J Cogn Neurosci 16:463–478.

1034   Jenkinson M (2003) Fast, automated, N-dimensional phase-unwrapping algorithm.

1035   Magn Reson Med 49:193–197.

1036   Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved Optimization for the

1037   Robust and Accurate Linear Registration and Motion Correction of Brain Images.

1038   NeuroImage 17:825–841.

1039   Johnson A, Redish AD (2007) Neural Ensembles in CA3 Transiently Encode Paths

1040   Forward of the Animal at a Decision Point. J Neurosci 27:12176–12189.

1041   Johnson A, Varberg Z, Benhardus J, Maahs A, Schrater P (2012) The hippocampus

1042   and exploration: dynamically evolving behavior and neural representations. Front

1043   Hum Neurosci 6 Available at:

1044   https://www.frontiersin.org/articles/10.3389/fnhum.2012.00216 [Accessed

1045   February 7, 2024].

1046   Jones MW, Wilson MA (2005) Theta Rhythms Coordinate Hippocampal–Prefrontal

1047   Interactions in a Spatial Memory Task. PLoS Biol 3:e402.

1048   Joo HR, Frank LM (2018) The hippocampal sharp wave-ripple in memory retrieval for

1049   immediate use and consolidation. Nat Rev Neurosci 19:744–757.

1050   Kable JW, Glimcher PW (2007) The neural correlates of subjective value during

1051   intertemporal choice. Nat Neurosci 10:1625–1633.

1052   Kaplan R, Adhikari MH, Hindriks R, Mantini D, Murayama Y, Logothetis NK, Deco G

1053   (2016) Hippocampal Sharp-Wave Ripples Influence Selective Activation of the

1054   Default Mode Network. Curr Biol 26:686–691.

1055   Kheirbek MA, Drew LJ, Burghardt NS, Costantini DO, Tannenholz L, Ahmari SE, Zeng

1056   H, Fenton AA, Hen R (2013) Differential Control of Learning and Anxiety along

1057   the Dorsoventral Axis of the Dentate Gyrus. Neuron 77:955–968.

1058   Klein A, Ghosh SS, Bao FS, Giard J, Häme Y, Stavsky E, Lee N, Rossa B, Reuter M,

1059   Neto EC, Keshavan A (2017) Mindboggling morphometry of human brains. PLOS

1060   Comput Biol 13:e1005350.

1061 Kleiner, M, Brainard, D.H., Pelli, D., Ingling, A., Murray, R., Broussard, C. (2007) What's

1062   new in Psychtoolbox-3. Perception 36:1–16.

1063 Knudsen EB, Wallis JD (2021) Hippocampal neurons construct a map of an abstract

1064   value space. Cell 184:4640-4650.e10.

1065 Knudsen EB, Wallis JD (2022) Taking stock of value in the orbitofrontal cortex. Nat Rev

1066   Neurosci 23:428–438.

1067 Laureiro-Martínez D, Brusoni S, Canessa N, Zollo M (2015) Understanding the

1068   exploration–exploitation dilemma: An fMRI study of attention control and

1069   decision-making performance. Strateg Manag J 36:319–338.

1070 Lim S-L, O'Doherty JP, Rangel A (2013) Stimulus Value Signals in Ventromedial PFC

1071   Reflect the Integration of Attribute Value Signals Computed in Fusiform Gyrus

1072   and Posterior Superior Temporal Gyrus. J Neurosci 33:8729–8741.

1073 Liu AA et al. (2022) A consensus statement on detection of hippocampal sharp wave

1074   ripples and differentiation from other fast oscillations. Nat Commun 13:6000.

1075 Liu Y, Dolan RJ, Kurth-Nelson Z, Behrens TEJ (2019) Human Replay Spontaneously

1076   Reorganizes Experience. Cell 178:640-652.e14.

1077 Lopez-Gamundi P, Yao Y-W, Chong TT-J, Heekeren HR, Mas-Herrero E, Marco-

1078   Pallarés J (2021) The neural basis of effort valuation: A meta-analysis of

1079   functional magnetic resonance imaging studies. Neurosci Biobehav Rev

1080   131:1275–1287.

1081   Lu H, Zou Q, Gu H, Raichle ME, Stein EA, Yang Y (2012) Rat brains also have a default

1082   mode network. Proc Natl Acad Sci 109:3979–3984.

1083   Mantini D, Corbetta M, Romani GL, Orban GA, Vanduffel W (2012) Data-driven analysis

1084   of analogous brain networks in monkeys and humans during natural vision.

1085   NeuroImage 63:1107–1118.

1086   Mantini D, Gerits A, Nelissen K, Durand J-B, Joly O, Simone L, Sawamura H, Wardak

1087   C, Orban GA, Buckner RL, Vanduffel W (2011) Default Mode of Brain Function in

1088   Monkeys. J Neurosci 31:12954–12962.

1089   McLaughlin AE, Redish AD (2023) Optogenetic disruption of the prelimbic cortex alters

1090   long-term decision strategy but not valuation on a spatial delay discounting task.

1091   Neurobiol Learn Mem 200:107734.

1092   Mehlhorn K, Newell BR, Todd PM, Lee MD, Morgan K, Braithwaite VA, Hausmann D,

1093   Fiedler K, Gonzalez C (2015) Unpacking the exploration–exploitation tradeoff: A

1094   synthesis of human and animal literatures. Decision 2:191–215.

1095   Morici JF, Weisstaub NV, Zold CL (2022) Hippocampal-medial prefrontal cortex network

1096   dynamics predict performance during retrieval in a context-guided object memory

1097   task. Proc Natl Acad Sci 119:e2203024119.

1098   Moustafa AA, Cohen MX, Sherman SJ, Frank MJ (2008) A role for dopamine in

1099   temporal decision making and reward maximization in Parkinsonism. J Neurosci

1100   28:12294–12304.

1101 Neubert F-X, Mars RB, Sallet J, Rushworth MFS (2015) Connectivity reveals

1102   relationship of brain areas for reward-guided learning and decision making in

1103   human and monkey frontal cortex. Proc Natl Acad Sci 112:E2695–E2704.

1104 Norman Y, Raccah O, Liu S, Parvizi J, Malach R (2021) Hippocampal ripples and their

1105   coordinated dialogue with the default mode network during recent and remote

1106   recollection. Neuron 109:2767-2780.e5.

1107 Norman Y, Yeagle EM, Khuvis S, Harel M, Mehta AD, Malach R (2019) Hippocampal

1108   sharp-wave ripples linked to visual episodic recollection in humans. Science

1109   365:eaax1030.

1110 Padoa-Schioppa C (2007) Orbitofrontal cortex and the computation of economic value.

1111   Ann N Acad Sci 1121:232–253.

1112 Padoa-Schioppa C, Assad JA (2008) The representation of economic value in the

1113   orbitofrontal cortex is invariant for changes of menu. Nat Neurosci 11:95–102.

1114 Papale AE, Stott JJ, Powell NJ, Regier PS, Redish AD (2012) Interactions between

1115   deliberation and delay-discounting in rats. Cogn Affect Behav Neurosci 12:513–

1116   526.

1117 Pastor-Bernier A, Stasiak A, Schultz W (2019) Orbitofrontal signals for two-component

1118   choice options comply with indifference curves of Revealed Preference Theory.

1119   Nat Commun 10:4885.

Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE (2014) Methods to detect, characterize, and remove motion artifact in resting state fMRI. NeuroImage 84:320–341.

Pruim RHR, Mennes M, van Rooij D, Llera A, Buitelaar JK, Beckmann CF (2015) ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. NeuroImage 112:267–277.

R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/.

Rajasethupathy P, Sankaran S, Marshel JH, Kim CK, Ferenczi E, Lee SY, Berndt A, Ramakrishnan C, Jaffe A, Lo M, Liston C, Deisseroth K (2015) Projections from neocortex mediate top-down control of memory retrieval. Nature 526:653–659.

Randall CK, Zentall TR (1997) Win-stay/lose-shift and win-shift/lose-stay learning by pigeons in the absence of overt response mediation. Behav Processes 41:227–236.

Redish AD (2016) Vicarious trial and error. Nat Rev Neurosci 17:147–159.

Rich EL, Wallis JD (2014) Medial-lateral Organization of the Orbitofrontal Cortex. J Cogn Neurosci 26:1347–1362.

Rolls E (2008) Functions of the orbitofrontal and pregenual cingulate cortex in taste, olfaction, appetite and emotion. Acta Physiol Hung 95:131–164.

1140   Royer S, Sirota A, Patel J, Buzsáki G (2010) Distinct Representations and Theta

1141   Dynamics in Dorsal and Ventral Hippocampus. J Neurosci 30:1777–1787.

1142   Rudebeck PH, Murray EA (2011) Dissociable Effects of Subtotal Lesions within the

1143   Macaque Orbital Prefrontal Cortex on Reward-Guided Behavior. J Neurosci

1144   31:10569–10578.

1145   Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, Eickhoff SB, Yeo

1146   BTT (2018) Local-Global Parcellation of the Human Cerebral Cortex from

1147   Intrinsic Functional Connectivity MRI. Cereb Cortex 28:3095–3114.

1148   Schmidt B, Redish AD (2021) Disrupting the medial prefrontal cortex with DREADDs

1149   alters hippocampal sharp-wave ripples and their associated cognitive processes.

1150   Hippocampus 31:1051–1067.

1151   Schuck NW, Niv Y (2019) Sequential replay of nonspatial task states in the human

1152   hippocampus. Science 364:eaaw5181.

1153   Schultheiss NW, Schlecht M, Jayachandran M, Brooks DR, McGlothan JL, Guilarte TR,

1154   Allen TA (2020) Awake delta and theta-rhythmic hippocampal network modes

1155   during intermittent locomotor behaviors in the rat. Behav Neurosci 134:529–546.

1156   Skaggs WE, McNaughton BL (1996) Replay of Neuronal Firing Sequences in Rat

1157   Hippocampus During Sleep Following Spatial Experience. Science 271:1870–

1158   1873.

1159 Smith DV, Clithero JA, Boltuck SE, Huettel SA (2014) Functional connectivity with

1160   ventromedial prefrontal cortex reflects subjective value for social rewards. Soc

1161   Cogn Affect Neurosci 9:2017–2025.

1162 Spellman T, Rigotti M, Ahmari SE, Fusi S, Gogos JA, Gordon JA (2015) Hippocampal–

1163   prefrontal input supports spatial encoding in working memory. Nature 522:309–

1164   314.

1165 Stachenfeld KL, Botvinick MM, Gershman SJ (2017) The hippocampus as a predictive

1166   map. Nat Neurosci 20:1643–1653.

1167 Stokes PA, Purdon PL (2017) A study of problems encountered in Granger causality

1168   analysis from a neuroscience perspective. Proc Natl Acad Sci 114 Available at:

1169   https://pnas.org/doi/full/10.1073/pnas.1704663114 [Accessed March 21, 2023].

1170 Sun C, Yang W, Martin J, Tonegawa S (2020) Hippocampal neurons represent events

1171   as transferable units of experience. Nat Neurosci 23:651–663.

1172 Sutton RS, Barto AG (2018) Reinforcement Learning, second edition: An Introduction.

1173   MIT Press.

1174 Tang W, Jadhav SP (2019) Sharp-wave ripples as a signature of hippocampal-

1175   prefrontal reactivation for memory during sleep and waking states. Neurobiol

1176   Learn Mem 160:11–20.

1177 Tang W, Shin JD, Jadhav SP (2021) Multiple time-scales of decision-making in the

1178   hippocampus and prefrontal cortex. eLife 10:e66227.

1179   Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, Marcus DJ,

1180   Westerlund A, Casey B, Nelson C (2009) The NimStim set of facial expressions:

1181   judgments from untrained research participants. Psychiatry Res 168:242–249.

1182   Trudel N, Scholl J, Klein-Flügge MC, Fouragnan E, Tankelevitch L, Wittmann MK,

1183   Rushworth MFS (2021) Polarity of uncertainty representation during exploration

1184   and exploitation in ventromedial prefrontal cortex. Nat Hum Behav 5:83–98.

1185   Tsypes A, Hallquist MN, Ianni A, Kaurin A, Wright AGC, Dombrovski AY (2024)

1186   Exploration-Exploitation and Suicidal Behavior in Borderline Personality Disorder

1187   and Depression. JAMA Psychiatry Available at:

1188   https://doi.org/10.1001/jamapsychiatry.2024.1796 [Accessed August 6, 2024].

1189   Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC (2010)

1190   N4ITK: Improved N3 Bias Correction. IEEE Trans Med Imaging 29:1310–1320.

1191   Ursu S, Clark KA, Stenger VA, Carter CS (2008) Distinguishing expected negative

1192   outcomes from preparatory control in the human orbitofrontal cortex. Brain Res

1193   1227:110–119.

1194   Vassena E, Krebs RM, Silvetti M, Fias W, Verguts T (2014) Dissociating contributions of

1195   ACC and vmPFC in reward prediction, outcome, and choice. Neuropsychologia

1196   59:112–123.

1197   Vilà-Balló A, Mas-Herrero E, Ripollés P, Simó M, Miró J, Cucurell D, López-Barroso D,

1198   Juncadella M, Marco-Pallarés J, Falip M, Rodríguez-Fornells A (2017)

1199   Unraveling the Role of the Hippocampus in Reversal Learning. J Neurosci Off J

1200   Soc Neurosci 37:6686–6697.

1201 Weilbächer RA, Gluth S (2016) The Interplay of Hippocampus and Ventromedial

1202   Prefrontal Cortex in Memory-Based Decision Making. Brain Sci 7:4.

1203 Wikenheiser AM, Redish AD (2015) Hippocampal theta sequences reflect current goals.

1204   Nat Neurosci 18:289–294.

1205 Winocur G, Olds J (1978) Effects of context manipulation on memory and reversal

1206   learning in rats with hippocampal lesions. J Comp Physiol Psychol 92:312–321.

1207 Woolrich MW, Ripley BD, Brady M, Smith SM (2001) Temporal Autocorrelation in

1208   Univariate Linear Modeling of FMRI Data. NeuroImage 14:1370–1386.

1209 Yeo BTT, Krienen FM, Jorge, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL,

1210   Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The

1211   organization of the human cerebral cortex estimated by intrinsic functional

1212   connectivity. J Neurophysiol 106:1125–1165.

1213 Yun M, Hwang JY, Jung MW (2023) Septotemporal variations in hippocampal value and

1214   outcome processing. Cell Rep 42:112094.

1215 Zeithamova D, Gelman BD, Frank L, Preston AR (2018) Abstract Representation of

1216   Prospective Reward in the Hippocampus. J Neurosci 38:10093–10101.

1217   Zhang Y, Brady M, Smith S (2001) Segmentation of brain MR images through a hidden

1218   Markov random field model and the expectation-maximization algorithm. IEEE

1219   Trans Med Imaging 20:45–57.

1220   Zhang Y, Huynh TKT, Dyson BJ (2023) Deliberately making miskates: Behavioural

1221   consistency under win maximization and loss maximization conditions. Npj Sci

1222   Learn 8:1–7.

1223