Increased hippocampal excitability and altered learning dynamics mediate cognitive mapping deficits in human aging

1 # Increased hippocampal excitability and altered learning dynamics mediate
2 # cognitive mapping deficits in human aging

3 Running title: Cognitive mapping deficits in human aging

4 Nadine Diersch[1*], Jose P Valdes-Herrera[1], Claus Tempelmann[2] & Thomas Wolbers[1,2,3]

5 *[1]Aging & Cognition Research Group, German Center for Neurodegenerative Diseases (DZNE), Magdeburg, 39120,*
6 *Germany*
7 *[2]Department of Neurology, Otto-von-Guericke University Magdeburg, 39120, Germany*
8 *[3]Center for Behavioural Brain Sciences (CBBS), Otto-von-Guericke University Magdeburg, 39120, Germany*
9
10 *Corresponding Author:
11 Dr Nadine Diersch (nadine.diersch@dzne.de)
12
13 33 pages
14 8 figures, 1 table, 2 videos
15 Abstract: 156; Introduction: 649; Discussion: 1500 words
16 JoN Section: Behavioral/Cognitive
17

18 **CONFLICT OF INTEREST STATEMENT**
19 The authors declare no competing interests.

26 **AUTHOR CONTRIBUTIONS**
27 Conceptualization, N.D. and T.W.; Project administration, N.D.; Investigation, N.D.; Resources, T.W. and C.T.;
28 Methodology & Formal Analysis, N.D. and J.P.V.-H.; Visualization, N.D.; Writing – Original Draft, N.D.; Writing
29 – Review & Editing, N.D., J.P.V.-H., C.T., and T.W.; Supervision, T.W.; Funding Acquisition, T.W.

ABSTRACT

Learning the spatial layout of a novel environment is associated with dynamic activity changes in the hippocampus and in medial parietal areas. With advancing age, the ability to learn spatial environments deteriorates substantially but the underlying neural mechanisms are not well understood. Here, we report findings from a behavioral and a fMRI experiment where healthy human older and younger adults of either sex performed a spatial learning task in a photorealistic virtual environment. We modeled individual learning states using a Bayesian state-space model and found that activity in retrosplenial cortex/parieto-occipital sulcus and anterior hippocampus did not change systematically as a function learning in older compared to younger adults across repeated episodes in the environment. Moreover, effective connectivity analyses revealed that the age-related learning deficits were linked to an increase in hippocampal excitability. Together, these results provide novel insights into how human aging affects computations in the brain's navigation system, highlighting the critical role of the hippocampus.

SIGNIFICANCE STATEMENT

Key structures of the brain's navigation circuit are particularly vulnerable to the deleterious consequences of aging, and declines in spatial navigation are among the earliest indicators for a progression from healthy aging to neurodegenerative diseases. Our study is among the first to provide a mechanistic account about how physiological changes in the aging brain affect the formation of spatial knowledge. We show that neural activity in the aging hippocampus and medial parietal areas is decoupled from individual learning states across repeated episodes in a novel spatial environment. Importantly, we find that increased excitability of the anterior hippocampus might constitute a potential neural mechanism for cognitive mapping deficits in old age.

INTRODUCTION

Exploring our surroundings has always been one of the hallmarks of human identity. To do so, we need to rapidly generate spatial representations and flexibly retrieve them later. With advancing age, however, these abilities deteriorate considerably (Lester et al., 2017). Older adults are slower in learning novel environments and have problems to utilize this information later (Iaria et al., 2009). Moreover, learning landmark locations during exploratory navigation is more difficult for them (Yamamoto and DeGirolamo, 2012), whereas their spatial memory is relatively preserved for familiar environments (Rosenbaum et al., 2012). As a consequence, they may avoid unfamiliar places and become overwhelmed when confronted with changes in their environment.

Although core regions of the brain's navigation circuit in the medial temporal lobe are among the first to be affected during the progression from healthy aging to Alzheimer's disease (AD; Braak and Del Tredici, 2015), the neural mechanisms for age-related deficits in spatial learning are still poorly understood, even in healthy older adults. Studies in rodents and non-human primates showed that place cells in the CA3 subfield of the hippocampus exhibit higher firing rates in aged animals during navigation, and they fail to encode new information when rats encounter novel environments (Wilson et al., 2005; Thomé et al., 2016). Moreover, firing patterns of aged CA1 place cells are often unstable across repeated visits to the same environment (Barnes et al., 1997). In humans, in contrast, there is evidence for an age-related hypoactivation in the hippocampus and in medial parietal areas during spatial navigation (Moffat et al., 2006; Konishi et al., 2013).

However, whether activity changes in the aging brain are indicative of a compensatory mechanism or a correlate of aberrant processing is a long-standing issue in cognitive neuroscience research on aging (Grady, 2012; Morcom and Henson, 2018). Evidence from studies investigating age-related impairments in separating sensory input from mnemonic representations (i.e., pattern separation) suggests that hyperactivity in the dentate gyrus and CA3 may underlie memory deficits in healthy aging (Yassa et al., 2011; Reagh et al., 2018). Hippocampal hyperactivity has been further linked to preclinical markers for AD (Leal et al., 2017).

Age-related differences in neural activity may further depend on the point in time when activity is measured during task performance. Studies in younger adults showed that the engagement of the retrosplenial cortex (RSC) and the parieto-occipital sulcus (POS) together with the hippocampus changes over the course of learning (Wolbers and Büchel, 2005; Auger et al., 2015; Brodt et al., 2016; Patai et al., 2019). For example, Wolbers and Büchel (2005) showed that activity in the RSC/POS tracked the learning of relative landmark locations during spatial navigation and increased across learning sessions, whereas hippocampal activity reflected the amount of learning in a given session and decreased over time. Given the time course of its involvement during spatial learning, the RSC has been implicated in the retrieval of hippocampal-dependent memories. It receives inputs from CA1 and the subiculum (Kobayashi and Amaral, 2003; Bzdok et al., 2015) and is known to be involved in the integration of different spatial reference frames as well as in updating spatial representations (Epstein, 2008; Miller et al., 2014). The hippocampus, in turn, particularly its anterior portion, is known for its role in generating (spatial) representations (Zeidman and Maguire, 2016). Moreover, place-cell like activity in the RSC of mice relies on intact input from the hippocampus to support memory retrieval (Mao et al., 2018).

4

89 Here, we report findings from two experiments where we 1) characterized age-related
90 problems in learning a novel environment, and 2) investigated the underlying neural mechanisms using fMRI.
91 We focused on activity changes in the RSC/POS and the hippocampus and changes in effective connectivity
92 within and between the two regions. This allowed us to test whether age-related problems in retrieving newly
93 learnt information during spatial navigation is linked to a malfunctioning of the integration of hippocampal
94 input within RSC/POS and/or a corrupted hippocampal signal.

95 **MATERIAL AND METHODS**

96 *Participants*

97 In the behavioral experiment, 17 younger (9 female, mean age: 24.0 ± 1.66, age range: 21-28) and 17 older
98 adults took part (8 female, mean age: 66.4 ± 2.69, age range: 61-72). All of them were right-handed (LQ: 91.9
99 ± 11.0; Oldfield, 1971) and the older adults showed no signs of major cognitive impairment with scores higher
100 than 23 in the Montreal Cognitive Assessment (MoCA score: 26.9 ± 2.18; Nasreddine et al., 2005; Luis et al.,
101 2009).
102 To determine the required sample size for the fMRI experiment, we ran a power analysis with the
103 effect size that was obtained in the behavioral experiment for the interaction between age group and learning
104 blocks (η_p^2 = .188), using G*Power 3.1 (α = 0.05, 1-β = 0.95, 2 groups, 8 repeated measurements; Faul et al.,
105 2007). The power analysis further considered the most conservative correction for non-sphericity with
106 1/number of measurements - 1. This analysis indicated a requirement of 28 participants in total. We decided to
107 double this number and recruited a total of 64 participants (27 younger adults, 37 older adults). Three
108 participants (one younger and two older adults) were excluded from further analyses because they were
109 identified as outliers in the fMRI data quality checks. In addition, one younger and three older adults were
110 excluded due to problems in following task instructions and/or cybersickness. The final fMRI sample consisted
111 of 25 younger (13 female, mean age: 23.4 ± 2.18, age range: 20-26) and 32 older adults (17 female, mean age:
112 67.3 ± 4.80, age range: 58-75). They were all right-handed (LQ: 90.4 ± 12.1; Oldfield, 1971) and the older adults
113 did not show signs of major cognitive impairment (MoCA score: 27.6 ± 1.93, range: 25-31; Nasreddine et al.,
114 2005).
115 Across experiments, participants had normal or corrected-to-normal vision and none of them
116 reported a history of psychiatric or neurological diseases or use of medication that might affect task
117 performance or MRI scanning. In addition, most of the participants already participated in previous
118 virtual reality (VR) experiments and, hence, were familiar with navigating in these kinds of setups.
119 Participants provided informed consent and were paid for their participation in accordance with the
120 local ethics committee.

121 *Virtual Environment*

122 Using 3ds Max (Autodesk, San Rafael, CA, USA), a novel virtual environment (VE) was developed, which
123 resembled a typical German historic city center consisting of town houses, shops and restaurants. The VE had
124 a square-like spatial layout with four interconnected 4-way intersections (Figure 1B). At two intersections, a

125   church and a town hall were placed at the end of one of the outgoing streets, whereas a 2D wall displaying a

126   photo texture of a street continuation bordered the remaining street ends. The VE was based on a 3D model of

127   the old city center of Tübingen. All of the participants confirmed to have never visited Tübingen before the

128   time of testing.

129   *Experimental Design and Procedure*

130   Vizard 5.0 (World Viz, Santa Barbara, CA, USA) was used to animate the experiments, which both started with

131   a familiarization phase during which the participants encountered the VE for the first time. Their task during

132   this phase was to collect tokens that were placed at the street ends by actively traveling the VE, using the four

133   arrow keys of a standard computer keyboard. This phase ended once every token was collected, ensuring that

134   they had visited every street at least once. It followed a short practice of the pointing task (8 trials) that was

135   used to measure navigational retrieval in the experiments. In this way, the VE and the task were introduced in

136   a step-wise manner to reduce the impact of different degrees of experience in handling VR setups on

137   task performance (Diersch and Wolbers, 2019).

138   In the behavioral experiment, eight learning blocks were implemented during which eight retrieval

139   phases alternated with seven encoding phases. One navigational retrieval phase consisted of 12 pointing

140   trials. A pointing trial started with participants being passively transported towards one of the intersections

141   starting from one of the four streets leading towards that intersection (Figure 1C, see Video 2 for an

142   example trial). Duration of this travel phase was fixed to 4 s corresponding to 20 virtual meters. The

143   movement stopped at the center of the intersection, a red crosshair appeared in the middle of the screen, and

144   participants were asked to point in the direction of one of the two target landmarks. Pointing was performed

145   by moving the crosshair to the left or right with the arrow keys of the keyboard. Once they believed to have

146   reached the correct position, they confirmed their response by pressing the space bar. Participants were asked

147   to respond as fast and accurately as possible with a time-out of 12 s (corresponding to 1½ 360° turns in the

148   VE). The ITI, showing a fixation cross, was fixed to 1.5 s. Throughout each trial, a picture cue of the target

149   landmark was displayed at the bottom of the screen and the background was obscured by fog to prevent

150   participants from seeing the street ends or target landmarks during pointing. The first seven retrieval phases

151   were followed by an encoding phase during which participants were passively transported around the whole

152   VE (without fog), starting from one of the two target landmarks in clockwise or counterclockwise order,

153   counterbalanced across the experiment (see Video 1 for a short segment of one tour). During encoding,

154   participants were instructed to pay close attention to the spatial layout of the VE and the location of the target

155   landmarks. Passive transportation instead of self-controlled traveling was chosen to ensure that every

156   participant experienced the VE for the same amount of time (duration: 2.88 min per tour). In total, participants

157   performed 96 navigational retrieval trials (4 intersections x 4 directions x 2 target landmarks x 3 repetitions) in

158   a pseudo-randomized order, with the restriction that each intersection/target landmark combination was

159   encountered starting from two of the four possible directions in the first half of the experiment. In the second

160   half of the experiment, divided by a self-timed break, the respective other two directions were used,

161   counterbalanced across participants. This allowed us to examine how experiencing familiar locations from a

162   novel viewpoint affects pointing performance.

163   The fMRI experiment also consisted of eight learning blocks during which eight retrieval phases

164   alternated with seven encoding phases (see Figure 1A for the structure of the fMRI experiment). fMRI

165   scanning started after a familiarization phase outside of the scanner with the same structure as in the

166   behavioral experiment and a short practice phase during structural imaging. One retrieval phase consisted of 8

167   navigational retrieval trials, which were followed by 4 control trials. These control trials also started with a 4 s

168   travel phase towards an intersection, followed by a pointing phase with a crosshair on screen. Here, cued by a

169   corresponding picture, however, participants were instructed to indicate which of the four corner buildings at

170   the intersection had changed its color and was shaded in blue. Their responses in the control task were

171   classified as correct if they were within ± 25° from the middle of the respective building, approximately

172   corresponding to its outline. Participants moved the crosshair with their index and middle finger for left and

173   right turns and confirmed their responses with their right thumb on a 5-key Lumitouch response box. Again,

174   participants were asked to respond as fast and accurately as possible with a time-out of 12 s. The ITIs had a

175   variable duration of 1-5 s with a mean of 3 s. During retrieval trials, an additional jittered interval of 0.5-1.5 s

176   duration with a mean of 1 s was included after the travel phase/before the crosshair appeared. The structure of

177   the respective encoding tours was the same as in the behavioral experiment (passive traveling with a constant

178   duration of 2.88 min per tour). In total, participants performed 64 navigational retrieval trials (4 intersections x

179   4 directions x 2 target landmarks x 2 repetitions) without the change of directions from the first to the second

180   half of the experiment as in the behavioral experiment. The change in directions was omitted in the fMRI

181   environment to eliminate the potential influence of approaching the intersections from novel viewpoints and

182   to accommodate a reduced number of trials due to the inclusion of the control task. In total, participants

183   performed 32 control trials (4 intersections x 4 directions x 2 repetitions). fMRI scanning consisted of 3 runs

184   that were divided by short breaks with 24 navigational retrieval trials, 12 control trials and 2 encoding tours in

185   the 1st run; 24 navigational retrieval trials, 12 control trials and 3 encoding tours in the 2nd run; and 16

186   navigational retrieval trials, 8 control trials and 2 encoding tours in the 3rd run.

187   *--- insert Figure 1 here ---*

188   *Bayesian Modeling of Performance Data*

189   In both experiments, subject-specific improvements in navigational performance were estimated by

190   using a Bayesian implementation of a state-space model that is similar to a local level model where

191   the trial outcomes, y, correspond to the observed level, and the state level represents the hidden

192   learning state, μ (Figure 2; Commandeur and Koopman, 2007). The hidden learning state, μ, is

193   following a random walk such that the actual block learning state depends on the learning state from

194   the previous block. Similar state-space models (e.g., Smith et al., 2007) have been used in previous

195   studies to estimate spatial learning (Wolbers and Büchel, 2005; Auger et al., 2015). However, these

196   studies modeled binary data on a trial-by-trial basis, whereas the present study used continuous

197   performance outcomes and focused on estimating spatial learning block-wise instead of trial-wise. To

198   model the learning state block-wise, an intermediate level accounts for the effects of the responses, η,

199   and shrinks the effects of individual trials within a block towards the block-wise learning state. In this

200   way, the model accounts for the fact that we can only measure behavioral performance but not the

201 effect of learning or navigational improvement, which we expected to change from one encoding
202 phase to the next but not necessarily from trial to trial. Introducing this intermediate level additionally
203 allowed us to incorporate potential missing trials into the response effects, η. In case of missing trials,
204 we estimated η ~ HalfNormal (log($\bar{y}_b$),1), i.e., using the log of the block mean as location parameter.
205 The model was implemented using the Python interface to Stan, PyStan (Carpenter et al., 2017; Stan
206 Development Team, 2017; see Figure 2-1 for the Stan code). To account for the substantial between-
207 and within subject variability of the data, weakly informative priors were chosen to provide vague
208 guidance for effective sampling. The model was fit for each participant using four chains each with
209 4000 iterations, of which 2000 correspond to the warm up period, totaling 8000 post-warm-up draws.
210 After inference, convergence of the chains was checked by means of the effective sample size and
211 the potential scale reduction factor (Rhat), confirming that our chains mixed well (Gelman and Shirley,
212 2011).

213   *--- insert Figure 2 here ---*

214 To determine the fit of our model to the data, we performed a posterior predictive check that compares the
215 observed data with simulated data using samples from the posterior distribution. In Figure 2-2A-E, the
216 posterior predictive samples distribution y_{rep} is plotted together with the observed data y for representative
217 individuals from different learning sub-groups in the fMRI experiment (see Performance Clustering section)
218 showing that our model was adequate to capture the observed data. We further compared our model to an
219 alternative, simpler model where η was removed (i.e., learning was estimated trial-wise instead of block-wise).
220 Using a leave-one-out (LOO) cross-validation (Vehtari et al., 2017), point-wise out-of-sample prediction
221 accuracies were estimated for both models. Comparing them confirmed that the model incorporating the
222 intermediate layer accounting for the response effects, η, provided better fit to the data, as evidenced by
223 positive LOO differences across the whole sample (sample mean = 1209, SE = 242; see Figure 2-2F for a
224 histogram showing the individual LOO difference values).

225 *fMRI Acquisition Parameters*

226 Scanning was performed on a 3T Magnetom Prisma scanner (Siemens Healthcare, Erlangen, Germany) with a
227 20-channel head coil. High-resolution T1-weighted anatomical images were acquired using a MPRAGE
228 sequence (1 mm isotropic resolution; TE = 2,82 ms; TR = 2500 ms; flip angle = 7°). In three functional runs,
229 whole-brain T2*-weighted echo planar images with BOLD contrast were acquired in interleaved bottom-up
230 order (36 slices, 3 mm isotropic resolution; TE = 30 ms; TR = 2000 ms; FoV = 216 mm; 72 x 72 image matrix; flip
231 angle = 90°).

232 *Behavioral and fMRI statistical analyses*

233 Behavioral Analyses
234 Absolute pointing errors (i.e., the deviation of the subject's response from the direction towards the respective
235 target landmark) served as performance measures in both experiments. In the behavioral experiment, we
236 additionally analyzed response times given the change in directions from which the intersections were

237   approached after the first half of the experiment. Where appropriate, analyses of variance (ANOVA) were

238   performed across learning blocks with age-group (younger adults, older adults) as between-subjects' variable.

239   In a control analysis, we checked for potential biases in pointing behavior by applying circular statistics on the

240   signed pointing error data relative to each target landmark for every intersection-direction combination, using

241   the CircStat toolbox in MATLAB (Berens, 2009). In general, a threshold of p < 0.05 was considered significant

242   (with correction for the number of tests where applicable).

243   <u>Logistic Regression Model</u>

244   With respect to the behavioral experiment, we were interested in whether two features that characterized

245   age-related differences in performance could be used to predict the age group of our participants. The first

246   feature was the mean amount of learning across all learning blocks, which was calculated based on the

247   differences between individual learning state estimates, derived from the Bayesian state-space model, from

248   two consecutive learning blocks. The estimates from the first learning block after the familiarization phase,

249   during which participants encountered the VE for the first time, were subtracted from chance level

250   performance (90°). In this way, learning related improvements in performance were considered that already

251   took place during the familiarization phase, resulting in pointing errors well below chance level in the first

252   learning block for some participants. The second feature were the changes in response times after the

253   directions changed from which the intersections were approached after the first half of the experiment. These

254   two features were normalized and then fed into a logistic regression model as implemented in Scikit-learn

255   (Pedregosa et al., 2011), with age group as target vector. The regularization parameter was set using a 10-fold

256   nested cross-validation, and the performance of the model was assessed by computing the average area

257   under the curve (AUC) for all folds. In this way, the probability of each individual belonging to the younger or

258   the older age group could be estimated. The resulting probabilities are interpreted in terms of individual

259   performance: those participants with a higher probability of belonging to the younger age group show better

260   performance on the task while a higher probability of being in the older age group relates to poorer

261   navigational performance.

262   <u>Performance Clustering</u>

263   In the analysis of the behavioral data from the fMRI experiment, we assessed whether subjects could be

264   clustered into different learning sub-groups based on their performance. This allowed us to investigate

265   learning-related differences in neural activity at the between-subjects level. For each participant, we created a

266   distribution based on the difference of the latent state distributions of the last and first learning block to

267   capture the overall amount of learning across the experiment. The mean and the standard deviation

268   parameters of this difference distribution were obtained by fitting it to a normal distribution using SciPy

269   (Jones et al., 2001). In this way, the clustering provides a richer source of information to distinguish between

270   different learning sub-groups. For example, taking only the steepness of the curve across learning blocks into

271   account, would not capture differences between very good learners, who learned most of the spatial layout

272   already during familiarization, and very bad learners, with both groups exhibiting flat learning curves.

273   However, they may differ in the uncertainty of their judgments, which is captured by the dispersion of the

274   difference distribution. We used a K-means clustering algorithm as implemented in Scikit-learn (Pedregosa et

275 al., 2011) to identify the centers of a pre-determined number of clusters based on their distances to the data

276 points. To obtain the optimal number of clusters to input into the K-means, we varied the number of possible

277 clusters from 3 to 7 and computed the mean Silhouette Coefficient of all samples per cluster as a measure for

278 the distance between the resulting clusters with values ranging from -1 to 1 (negative values would indicate

279 wrong cluster assignments and values near zero overlapping clusters). We found 5 to be the best choice for the

280 number of learning sub-groups in our sample (respective silhouette scores per tested cluster number: 3: 0.259,

281 4: 0.395, 5: 0.457, 6: 0.429, 7: 0.410). One should note that the results of this data-driven approach to

282 characterize the heterogeneity within the two age groups are specific to our sample and cannot be

283 generalized to the whole population. Different samples of younger and older adults might result in different

284 learning clusters due to different performance levels.

285 <u>fMRI Image Quality Control and Preprocessing</u>

286 The imaging data were first transformed into the Brain Imaging Data Structure (BIDS) format (Gorgolewski et

287 al., 2016). MRIQC (Version 0.9.3; Esteban et al., 2017) was used for checking the quality of the MRI data.

288 MRIQC utilizes tools from different software packages such as FSL or Advanced Normalization Tools (ANTs)

289 to extract image quality metrics (IQMs) and generates visual reports at the individual and group level. This

290 allows the evaluation of different characteristics of the structural and functional MR images, for example,

291 SNR/tSNR, sharpness, and presence of artifacts. Data from one younger adult and two older adults were

292 consequently excluded from further analyses due to strong task-related movement and/or artifacts in several

293 functional runs resulting in low-quality IQMs (e.g., high Ghost-to-Signal ratio, low tSNR). Next, motion

294 correction, slice timing, co-registration, and normalization of the images was performed using fMRIprep

295 version 1.0.0-rc5 (Esteban et al., 2019) that also draws on different software packages to provide the optimal

296 implementation for different stages of preprocessing. For example, normalization to MNI space was

297 performed using ANTs as a state-of-the-art medical image registration and segmentation toolkit. Finally, the

298 data were smoothed with a 6 mm full-width at half maximum isotropic Gaussian kernel using SPM 12

299 (Wellcome Department of Imaging Neuroscience, London, UK).

300 <u>ROI Definition</u>

301 Based on results from previous studies (Wolbers and Büchel, 2005; Auger et al., 2015; Mao et al., 2018), we

302 defined two regions of interest (ROI), namely, the RSC/POS and the hippocampus. The single ROIs were

303 created based on each participant's T1 structural scan using a semiautomated anatomic reconstruction and

304 labeling procedure as implemented in FreeSurfer v6.0.0 (http://surfer.nmr.mgh.harvard.edu; Dale et al., 1999;

305 Fischl et al., 1999). In each hemisphere, labels corresponding to the posterior-ventral part of the cingulate

306 gyrus (area 10) and the parieto-occipital sulcus (area 65) from the Destrieux Atlas and the hippocampus from

307 the subcortical segmentation were extracted (Fischl et al., 2002; Destrieux et al., 2010). The two cortical labels

308 were combined into one RSC/POS ROI. Each ROI was next transformed to MNI space. Hemispheres were

309 combined to one bilateral ROI, thresholded at 0.5, and finally resampled to correspond to the resolution of our

310 functional images. The ROIs were subsequently used in the univariate analysis and for the volumes of interest

311 (VOI) extraction in the effective connectivity analysis. In our definition of the hippocampus ROI, we did not

312 separate between anterior and posterior hippocampus because previous fMRI studies have not reported a

313 clear dissociation along the hippocampal long axis during navigation in healthy aging. Hence, strong a-priori

314 hypotheses about a potential anterior-posterior dissociation seemed unwarranted. Distinctions between

315 anterior or posterior hippocampus in the results presentation refer to the location of the clusters we obtained

316 in our analyses with foci at or anterior to $y = -21$ mm in MNI space being regarded as belonging to the anterior

317 hippocampus (Poppenk et al., 2013).

318 <u>fMRI Univariate Analysis</u>

319 At the single-subject level, a general linear model (GLM) was specified with six regressors of interest for each

320 learning block using a high-pass filter of 100 Hz. For the navigational as well as control retrieval trials, we

321 created regressors for the 4 s travel phase and the pointing phase. For the encoding phases, regressors

322 modeled periods when participants were located within 20 m of the intersection centers (corresponding to the

323 area covered during the retrieval travel phases) as well as outside of these areas. Finally, the time of the

324 button press was modeled as regressor of no interest. All regressors were convolved with the standard

325 canonical hemodynamic response function (HRF) in SPM12. In addition, we included motion parameters, the

326 frame-wise displacement (FD) and aCompCor values (Behzadi et al., 2007), as obtained from fMRIprep

327 preprocessing, in the GLM to control for physiological and movement confounds. In aCompCor, significant

328 principal components are derived from noise regions-of-interest (ROI) in which the time series data are

329 unlikely to be modulated by neural activity. In this way, potential confounding effects of physiological

330 fluctuations that may differ between age groups, such as cardiac pulsations and respiration-induced

331 modulations, are removed from the fMRI time-series.

332 We focused on interaction effects between conditions of interest and age group that are unlikely to

333 be driven by group differences in neurovascular coupling, unlike main effects of age (Rugg and Morcom, 2005;

334 see also Grinband, Steffener, Razlighi, & Stern, 2017). First, we contrasted navigational retrieval trials to

335 control trials to identify general activation patterns in the RSC/POS and the hippocampus during spatial

336 navigation in our complex real-world environment, similar to previous studies investigating age-group

337 differences in spatial navigation (Moffat et al., 2006). We additionally contrasted the travel phases towards

338 the intersections during navigational retrieval trials to the corresponding periods when participants

339 encountered the same areas during the encoding tours. The within-subject effects of learning were assessed

340 by using the normalized differences between learning state estimates (i.e., the outputs from the Bayesian

341 state-space model) from consecutive learning blocks (i.e., amount of learning) as contrast weights over the

342 regressors modeling each travel phase during navigational retrieval per learning block (cf., Wolbers and

343 Büchel, 2005). At the group level, the resulting individual contrast images were entered into two-sample t-

344 tests to assess interactions with age group. Finally, in order to check in which regions activity changes across

345 learning blocks are modulated by the overall learning ability of the individual, we ran an additional analysis in

346 which learning sub-group was added as covariate in the age-group comparisons at the second-level. All

347 contrasts were evaluated at $p < .001$ (uncorrected) and we report activations that survived the FWE-correction

348 for multiple comparisons using a threshold of $p < 0.05$ at the cluster level.

349 In a control analysis, we checked whether learning-related changes within the two ROIs could

350 alternatively be driven by spatial computations in which older and younger adults engage in differently over

351   the course of the experiment. We fitted a separate finite impulse response (FIR) GLM for each participant with

352   the same regressors of interest as in our main GLM described above. The FIR model was set up with eight time

353   bins (2 s duration each, total time window:16 s) as a basis function for the HRF, and FIR time courses (percent

354   signal change per time bin) were extracted within both the hippocampus and the RSC/POS ROIs for the

355   regressors modeling the eight travel phases using MarsBar (Brett et al., 2002). For each participant, we then

356   determined the time bin when the HRF reached its peak, separately for the beginning of learning (first four

357   learning blocks) and the end of learning (last four learning blocks). This approach allowed us to test (i) whether

358   the HRF reached its peak at different time points in the first vs. second half of the experiment, and (ii)

359   whether this time-to-peak differed between age groups.

360   <u>Effective Connectivity Analysis</u>

361   Effective connectivity within and between the hippocampus and the POS was examined using the parametric

362   empirical Bayesian (PEB) approach in the context of Dynamic Causal Modeling (DCM) as implemented in

363   SPM12 (Friston et al., 2016).

364   *GLM and VOI Selection.* For the DCM analysis, we created a GLM in which the time-series from our three

365   functional runs were concatenated and added regressors that modeled the mean signal for each run. The

366   amount of learning per learning block was included as parametric modulation of the regressor modeling the

367   travel phase during navigational retrieval trials for each participant. All other regressors were the same as in

368   the first GLM although they were not modeled separately for each learning block. The sanity check of the

369   concatenated GLM revealed that activity in the right anterior hippocampus (27, -9, -15, Z = 3.97; 27 voxels)

370   decreased and activity in the left POS (-12, -63, 31, Z = 3.56; 42 voxels) increased with the amount of learning

371   in younger adults ($p < 0.05$, FWE-corrected for the respective ROI). No additional activations emerged

372   elsewhere in the brain. When testing for interactions between learning-related activity changes and age group

373   within our ROIs, one cluster within bilateral POS extending to RSC was revealed (15, -66, 44, Z = 4.23; 12, -57,

374   4, Z = 4.21; -6, -66, 24, Z = 3.94; -15, -60, 28, Z = 3.61; 369 voxels). Thus, activity in this region increased with

375   learning in younger adults but less so in older adults. The slight differences of these results to the ones from

376   the first GLM are likely related to differences in the design of the two GLMs. Whereas the first GLM was

377   optimized to capture our experimental design as precisely as possible by modeling all regressors of interest

378   separately for each learning block, the concatenated GLM was optimized for the DCM analysis that relies on

379   single-run time-series.

380   BOLD time-series were extracted for each individual using a t-contrast over the regressors modeling

381   the travel phase during navigational retrieval and the amount of learning with a liberal threshold of $p < 0.1$

382   (Note that this threshold was only used for VOI selection, but not in the final DCM statistics). The principal

383   eigenvariate was extracted around the group peak coordinates within the hippocampus and POS as obtained

384   in the univariate analysis of the concatenated GLM and was allowed to vary as an 8 mm sphere centered on

385   the subject-specific maximum constrained by a 24 mm sphere centered on the group maximum and the

386   respective ROI mask. In this way, variation between individuals in the exact location of the effect was

387   considered, given the high heterogeneity in our sample and slightly different peak voxels in the two GLMs.

388   The extractions were corrected using an F-contrast that retained the effects of interest (navigational as well as

389 control retrieval phases, encoding phases, button press) while partitioning out task-unrelated variance caused

390 by head motion, for example. For participants for which no supra-threshold voxels were identified (three

391 younger adults and one older adult), the threshold was lowered to p < 0.5 to extract BOLD time-series (cf.,

392 Zeidman et al., 2019b).

393 *First-Level DCM Specification.* We specified a bilinear, one-state DCM for each participant by setting the

394 regressor modeling the travel phase during navigational retrieval trials as driving input entering the cortical

395 network via the POS. The amount of learning per learning block was included as modulatory input on the

396 bidirectional connections between hippocampus and POS. All inputs were mean-centered so that the A-

397 matrix of the DCM represents the average connectivity across experimental conditions. We used stochastic

398 DCM that seeks to improve model estimation by modeling random fluctuations and hidden neuronal causes in

399 the differential equations of the neuronal states (Li et al., 2011; Daunizeau et al., 2012). In this way, the impact

400 of potential confounding effects of variations in BOLD response caused by age is reduced. Bayesian group

401 inversion was performed, providing estimates of the connection strength parameters that best explained the

402 observed data per participant. Critically, within DCM PEB, at each iteration of the within subject inversion, the

403 individual priors are updated using the group average connection strengths as priors. Inspection of the single

404 DCMs after inversion confirmed that our full model provided good fit to the observed data with an average of

405 44.5 ± 3.22% variance explained.

406 *Second-Level PEB Model.* Next, we created a second level PEB model over the parameters that included the

407 group mean and age group as covariates to identify differences between younger and older adults. We further

408 included learning sub-group and its interaction with age group as covariates in the model. The interaction

409 term was modelled as the two main effects of age and learning group element-wise multiplied with the main

410 effects being mean-centered and coded in a way that low/negative values represent younger or better

411 performing individuals. A search over nested PEB models was performed by using Bayesian model comparison

412 (BMC) that explores a space of models under the assumption that different combinations of the connections

413 may exist across participants (Zeidman et al., 2019a). To search over hundreds of nested models incorporating

414 different combinations of connections and group differences, Bayesian model reduction (BMR) was used that

415 iteratively prunes parameters from the full model until model-evidence decreases. To reduce dilution of

416 evidence, we separately checked for group differences in the A-matrix (average connectivity across

417 experimental conditions) and the B-matrix (within-subject modulatory input of the amount of learning per

418 block). We further performed a LOO cross-validation to check whether the model parameter that differed

419 between older and younger adults could be used to predict the participants' age group.

420 *Data Availability*

421 Source data files for the main results figures and tables are stored at https://osf.io/fjbxu/. We additionally

422 provide a key resources table listing all the software packages that were used in the current study. The Stan

423 code of the Bayesian state-space model can be found in Figure 2-1.

424 **RESULTS**

425   Findings are reported from two separate samples comprising healthy younger and older adults who

426   performed a spatial learning task either purely behaviorally (17 younger adults and 17 older adults) or in a

427   combined fMRI-behavioral experiment (25 younger adults and 32 older adults). In both experiments, following

428   an initial familiarization phase before testing/outside of the scanner, eight learning blocks were implemented

429   during which eight retrieval phases alternated with seven encoding phases. We used the angular deviation of

430   the participants' response from the respective target landmark (i.e., absolute pointing errors) to measure

431   performance improvements across learning blocks. However, performance in these kinds of tasks can be

432   corrupted by various noise sources and, hence, might not accurately reflect the actual learning state of the

433   participant. Therefore, subject-specific improvements in navigational performance were estimated by using a

434   Bayesian implementation of a state-space model that disambiguated learning from random trial-by-trial

435   fluctuations in performance. We used the outputs of the model in the analysis of the fMRI data to examine

436   intra- and interindividual differences in learning.

437   ***Behavioral experiment***

438   <u>Lower performance and reduced learning in older adults</u>

439   An ANOVA with learning block (1-8) as repeated measures variable and age-group (younger adults, older

440   adults) as between-subjects variable on the average absolute pointing errors showed significant main effects

441   of learning block, $F(7, 224) = 19.5$, $p < .001$, $\eta_p^2 = .379$, and age group, $F(1, 32) = 85.2$, $p < .001$, $\eta_p^2 = .727$. This

442   was modulated by a significant interaction between the two factors, $F(7, 224) = 7.40$, $p < .001$, $\eta_p^2 = .188$. At the

443   beginning, both age groups performed around chance level (90°), even though older adults had spent

444   significantly more time than younger adults ($M_{old} = 534$ s $\pm$ 161 s; $M_{young} = 218$ s $\pm$ 41.4 s; $t(16.9) = -7.63$, $p < .001$,

445    $d = 2.69$) in the initial familiarization phase of the experiment, during which they encountered the VE for the

446   first time. Over the course of the experiment, however, older adults showed lower performance and less

447   improvement compared to younger adults (Figure 3A). The change in direction from the first to the second

448   half of the experiment did not have a major effect on this pattern of results as implied by a non-significant

449   interaction between learning block and age group when directly comparing the fourth and fifth learning block,

450    $F(1, 32) = 1.96$, $p = .171$, $\eta_p^2 = .058$. A separate ANOVA within the older age group on pointing performance per

451   learning block confirmed that older adults generally improved on the task over time as evidenced by a

452   significant main effect of learning, $F(7, 112) = 2.58$, $p = .017$, $\eta_p^2 = .139$. According to the outputs of the

453   Bayesian state-space model (Figure 3C, see Figure 3-1A-B for average pointing errors per learning block for

454   each participant), most of the younger adults learned the spatial layout of the VE very fast, reaching ceiling

455   performance already after the first few learning blocks. The older adults, in contrast, differed more widely in

456   their ability to learn.

457   To investigate potential biases in pointing behavior that might differ between the age groups, such as

458   an increased tendency to point along streets, circular statistics were applied on the signed pointing error data

459   relative to each target landmark for every intersection-direction combination. From the 32 age-group

460   comparisons (4 intersections x 4 directions x 2 target landmarks), only 7 reached significance as determined by

461   a Watson-Williams test, all $p \le .047$. Older adults showed larger deviation from the correct angle than younger

462 adults in 6 of the 7 instances. The direction of the deviations in pointing (e.g., to the left or right relative to the

463 target landmark), however, varied and none of the effects survived when correcting for multiple comparisons.

464 Higher uncertainty when viewpoints are changing in older adults

465 An ANOVA with learning block (1-8) as repeated measures variable and age group (younger adults, older

466 adults) as between-subjects variable on the response time data confirmed significant main effects of learning

467 block, $F_{(7, 224)} = 9.26$, $p < .001$, $\eta_p^2 = .225$, and age group, $F_{(1, 32)} = 10.5$, $p = .003$, $\eta_p^2 = .247$. Compared to

468 older adults, younger adults responded quicker and showed a steeper decline in response times over time as

469 revealed by a significant interaction between learning block and age group, $F_{(7, 224)} = 4.29$, $p = .001$, $\eta_p^2 =$

470 .118. Notably, when comparing the fourth and fifth learning blocks, a significant interaction between learning

471 block and age group was obtained, $F_{(1, 32)} = 9.34$, $p = .004$, $\eta_p^2 = .226$. Older but not younger adults showed a

472 substantial increase in response times in the fifth learning block when the intersections were encountered

473 from novel directions (Figure 3B). This result cannot be explained by a confound between pointing

474 performance and required turning at the intersections because the required amount of turning to perform

475 accurately on the task varied from trial to trial, depending on the specific intersection-direction-target

476 landmark combination. Moreover, it was kept constant across experiment halves and participants (average

477 turning direction: 135°). When considering the fourth and fifth block only, the correct turning angle did not

478 differ between blocks, age groups, or varied between age groups as a function of learning block, all $F \le 3.23$, p

479 $\ge .082$, $\eta_p^2 \le .092$. Thus, older adults' representations of the spatial layout of the environment seem to be

480 more rigidly tied to the sensory input encountered at the beginning of learning, leading to a temporary

481 uncertainty when viewpoints are suddenly changing.

482 *--- insert Figure 3 here ---*

483 Performance in older adults is partly influenced by their facing direction

484 Age-related differences in pointing performance depending on the nature of the trials during navigational

485 retrieval (i.e., respective intersection-direction-target landmark combination) were further analyzed by means

486 of an ANOVA on the absolute pointing errors with intersection (I1-I4), direction (D1-D4), and target landmark

487 (town hall, church) as repeated measures variables and age group (younger adults, older adults) as between-

488 subjects variable. A significant interaction between the four factors suggested that the performance of the

489 age groups was modulated by the respective intersection-direction-target landmark combination

490 encountered in the VE during retrieval, $F_{(9, 288)} = 2.05$, $p = .034$, $\eta_p^2 = .060$. Therefore, follow-up ANOVAs

491 were conducted within the two age groups separately. In younger adults, a significant main effect of

492 intersection, $F_{(3, 48)} = 6.18$, $p = .005$, $\eta_p^2 = .279$ (Greenhouse-Geißer corrected), showed that performance was

493 worse when they were located at I4 (M = 48.6° ± 21.9°) as compared to I1 (M = 25.0° ± 19.7°) or I2 (M = 30.6° ±

494 25.7°), all $p \le .010$ (Bonferroni-corrected). This was modulated by a significant interaction between

495 intersection and target landmark, $F_{(3, 48)} = 11.4$, $p < .001$, $\eta_p^2 = .416$. Pointing errors were smaller in this age

496 group when they pointed towards the town hall (M = 13.5° ± 14.5°) as compared to the church (M = 36.6° ±

497 29.3°) at I1, which was the intersection adjacent to the town hall, and vice versa at I3, which was the

498 intersection adjacent to the church (town hall: M = 45.5° ± 35.5°; church: M = 25.8° ± 26.9°), all $t \ge 3.33$, $p \le$

499   .004, d ≥ 0.807. The directions from which the intersections were approached did not have an influence on

500   performance in this age group, all $F \leq 1.80$, $p \geq .133$, $\eta_p^2 \leq .101$. In older adults, there was also an interaction

501   between intersection and target landmark, $F(3, 48) = 3.38$, $p = .026$, $\eta_p^2 = .174$. When located at I3, pointing

502   errors were smaller when the target landmark was the adjacent church (M = 73.2° ± 27.4°) as compared to the

503   town hall (M = 94.2° ± 19.9°), $t(16) = 2.73$, $p = .015$, d = 0.662. The corresponding comparison for I1 did not

504   reach significance, $t(16) = 1.12$, $p = .281$, d = 0.271. In addition, there was a significant interaction between

505   direction and target landmark, $F(3, 48) = 3.75$, $p = .039$, $\eta_p^2 = .190$ (Greenhouse-Geißer corrected). Post-hoc t-

506   tests indicated that pointing towards the town hall (M = 80.3° ± 18.1°) tended to be easier as compared to

507   pointing towards the church (M = 94.5° ± 27.1°) for the older adults when they approached the intersections

508   from D4 (i.e., facing east), $t(16) = 1.96$, $p = .068$, d = 0.475. In contrast, pointing towards the church (M = 69.9°

509   ± 19.6°) tended to be easier than pointing towards the town hall (M = 87.0° ± 23.3°) when they approached the

510   intersections from D2 (i.e., facing west), $t(16) = 1.99$, $p = .064$, d = 0.483. This was modulated by an interaction

511   between intersection, direction, and target landmark, $F(9, 144) = 2.25$, $p = .022$, $\eta_p^2 = .123$. Separate follow-up

512   ANOVAs for each intersection with direction (D1-D4) and target landmark (town hall, church) as repeated

513   measures variables revealed for I2 a main effect of direction, $F(3, 48) = 5.25$, $p = .003$, $\eta_p^2 = .247$, indicating that

514   pointing generally seemed to be easier from D2 (i.e., facing west; M = 68.8° ± 32.7°) as compared to D1 (M =

515   94.0° ± 24.5°) or D4 (M = 95.8° ± 37.1°), that is, when they were facing towards the dead-ends at this

516   intersection, all $p \leq .054$ (Bonferroni-corrected). At I3, a main effect of target landmark indicated that pointing

517   towards the adjacent church (M = 73.2° ± 27.4°) was easier for the older adults than pointing towards the town

518   hall (M = 94.2° ± 20.0°), $F(3, 48) = 7.47$, $p = .015$, $\eta_p^2 = .318$. This was modulated by an interaction between

519   direction and target landmark, $F(3, 48) = 4.44$, $p = .008$, $\eta_p^2 = .217$. Pointing towards the church was easier

520   when coming from D1 (i.e., facing south; town hall: M = 107.7° ± 31.5°; church: M = 65.6° ± 36.0°) or D2 (i.e.,

521   facing west; town hall: M = 98.0° ± 39.4°; church: M = 52.6° ± 40.8°), all $t \geq 3.19$, $p \leq .006$, d ≥ .773. Finally, at I4,

522   there was also an interaction between direction and target landmark, $F(3, 48) = 3.74$, $p = .017$, $\eta_p^2 = .190$.

523   Performance was better when participants pointed towards the church (M = 75.9° ± 30.8°) as compared to the

524   town hall (M = 106.6° ± 31.4°) when approaching the intersection from D2 (i.e., facing west), $t(16) = 2.59$, $p =$

525   .020, d = .629.

526   To summarize, the results of this analysis again demonstrate better navigational encoding in younger

527   adults and a higher reliance on the specific sensory input in older adults. The directions from which the older

528   adults were approaching the intersections partly seemed to have an impact on their performance, although

529   variability in performance was generally quite high.

530   <u>Individual learning state and response time increase after direction change predict age-group</u>

531   We next used a logistic regression model to check whether age-group can be determined based on two

532   features that characterized age-related performance differences in our task. The mean amount of learning

533   across the whole experiment (i.e., difference between individual learning state estimates from consecutive

534   learning blocks) and the change in response times from the 4th to the 5th learning block served as input

535   features. The model performed very well to estimate the probability of being classified as a younger adult with

536   an average area under the curve (AUC) of 0.99±0.02%. Thus, those participants with a higher probability of

537 belonging to the younger age group show better performance on the task while a higher probability of being

538 in the older age group relates to poorer navigational performance, i.e., a lower mean amount of learning

539 across blocks and a higher increase in response times when previously learned locations are encountered from

540 novel viewpoints (Figure 3D).

541 *fMRI experiment*

542 After pre-processing of the fMRI data using fmriprep (Esteban et al., 2019) and SPM12, we performed a

543 univariate regression analysis to identify age-related differences in neural activity in the RSC/POS and the

544 hippocampus during different phases of the experiment. We further examined the effects of learning at the

545 within- and between-subject level. Finally, we examined age- and learning-related differences in effective

546 connectivity within and between the two regions.

547 <u>Learning ability varies within the older age group</u>

548 As in the behavioral experiment, older compared to younger adults spent considerably more time in the initial

549 familiarization phase of the experiment outside of the scanner (M_{old} = 466 s ± 133 s; M_{young} = 258 s ± 57.5 s;

550 t(44.3) = -7.91, p < .001, d = 2.03. Moreover, we found significant main effects for learning block, F(7, 385) =

551 32.3, p < .001, η_p^2 = .370, and age group, F(1, 55) = 167, p < .001, η_p^2 = .752, together with a significant

552 interaction between the two factors for the average absolute pointing errors, F(7, 385) = 11.0, p < .001, η_p^2 =

553 .166. This indicates that younger compared to older adults again showed better performance on the task and

554 stronger improvement across learning blocks (Figure 4A). Older adults, however, did show learning at the

555 group level as confirmed by a separate ANOVA within this age group, F(7, 217) = 3.58, p = .001, η_p^2 = .103.

556 Accuracy for the control trials was at ceiling across the whole sample (mean proportion of correct responses =

557 0.97 ± 0.05). In contrast to the behavioral experiment, the change in directions from the first to the second half

558 of the experiment was omitted here due to the reduced number of trials per learning block. Thus, we did not

559 expect changes in response times from the first to the second half of the experiment.

560 Individual learning state estimates as obtained from the state-space model again showed that

561 participants varied substantially in their ability to learn the spatial layout of the VE (Figure 4B, see also Figure

562 3-1C-D for average pointing errors per learning block for each participant). To determine how neural activation

563 patterns were modulated by the individuals' overall amount of learning across the experiment, we used a K-

564 means clustering algorithm to identify learning sub-groups based on the difference between the latent state

565 distributions of the last and first learning block. The estimated optimal number of clusters in our sample

566 turned out to be five (Figure 4C): A group of *top learners* (n = 9), consisting of seven younger adults and two

567 older adults, already learned the layout of the VE after the familiarization phase resulting in a small difference

568 in learning between the first and the last learning block. The second cluster exclusively consisted of younger

569 adults, categorized as *good learners* (n = 14). They typically reached ceiling performance during the first half of

570 the experiment with a low variance in their difference distribution. A group of *intermediate learners* (n = 9),

571 consisting of a three younger and six older adults, were still improving in the second half of the experiment

572 and consequently exhibited the largest difference in their hidden learning state from the beginning to the end

573 of the experiment and a relatively high variance. Individuals belonging to the fourth cluster were categorized

574 as *weak learners* (n = 12) who showed only a small improvement across the whole experiment and a high

575 variability. This cluster consisted of older adults only. Finally, 12 older adults and one younger adult did not

576 show considerable improvement across the learning blocks and were consequently categorized as *non-*

577 *learners* (n = 13) in the context of our experiment. Difference distributions for representative individuals from

578 each learning sub-group, together with learning estimates and behavioral data per learning block, can be

579 found in Figure 4-1. Although these results are specific to our sample, the same clustering analysis within the

580 behavioral experiment yielded comparable results in terms of the number of learning sub-groups,

581 underscoring the validity of its results in the context of our task (see Figure 4-2).

582 These between-subject differences in learning demonstrate that our task was neither too easy nor

583 too difficult for one of the age groups per se. In addition, the learning sub-groups within each age group were

584 comparable with respect to the factors age, sex or signs of major cognitive impairment (Figure 4-3).

585 *--- insert Figure 4 here ---*

586 <u>Age-related hyperactivation in the hippocampus and RSC/POS during navigational retrieval</u>

587 First, to identify overall age-related differences in activation patterns within the RSC/POS and the

588 hippocampus, irrespective of learning, we contrasted navigational retrieval trials to control trials. Activity in

589 medial parts of the RSC/POS was increased in older compared to younger adults for this comparison. This age-

590 related activity increase was also observed in the left anterior hippocampus (Table 1A). An age-related activity

591 reduction was found in a small cluster in the superior right POS and also in a more lateral cluster in the right

592 POS (Table 1B). Second, we tested for interactions between age group and activation differences during

593 navigational retrieval versus encoding to check whether these two phases of the experiment were differently

594 affected by age. In one cluster of the right POS as well as two clusters in the right and left anterior

595 hippocampus, activity was increased in older adults compared to younger adults during navigational retrieval

596 versus encoding (Table 1C). There were no clusters within our ROIs where activity was reduced in older adults.

597 Activations outside of our ROIs for these two comparisons and the corresponding results for the whole sample

598 can be found in Table 1-1A-G.

599 <u>Learning-related activity changes in anterior hippocampus and RSC/POS are less pronounced in older adults</u>

600 By using the amount of learning per block as contrast weights in our GLM, we assessed learning-related age-

601 group differences in the time-course of hippocampal and RSC/POS involvement during navigational retrieval.

602 First, we found that activity in the anterior portion of the right hippocampus decreased in younger but less so

603 in older adults as a function of learning (Table 1D, Figure 5A). This suggests that hippocampal activity

604 reflected the amount of spatial knowledge that was acquired after each encoding tour in the younger age

605 group. In older adults, in contrast, hippocampal activation did not change systematically across learning

606 blocks. Second, we also found several clusters within the RSC/POS ROI where activity increased over the

607 course of the experiment more in younger than in older adults (Table 1E, Figure 5B). This concerned the whole

608 extent of the left POS from its superior parts to the left RSC, a cluster in the right RSC/POS, and a more lateral

609 cluster in the right POS. Activity in these clusters paralleled changes in performance across learning blocks in

610 the younger age group. Older adults' individual learning curves, in contrast, were again decoupled from

611 activity changes in these regions. Activations outside of our ROIs for these comparisons and the

612 corresponding results for the whole sample can be found in Table 1-1H-I.

613   *--- insert Figure 5 here ---*

614   It is possible that the decreasing BOLD responses in the hippocampus and the increasing responses in

615   RSC/POS, which we observed in younger adults, could have been driven by younger adults becoming quicker

616   with self-localization, allowing them to compute the direction towards the target landmark at progressively

617   earlier time points. Under this scenario, one would predict a temporal shift of the BOLD response, in particular

618   for the RSC (assuming a role of the RSC/POS in deriving directional relationships between one's position and

619   landmarks). To directly test this hypothesis, we performed a control analysis using a FIR model of the

620   hemodynamic response but did not find any indications that the time bin of the peak of the HRF changed

621   between the first and the second half of the experiment, neither in the hippocampus nor in the RSC/POS ROI.

622   In the hippocampus, the median time bin was 3.75 (IQR = 1.25) in the first and 4.25 (IQR = 1.50) in the second

623   half of the experiment in younger adults (out of 8 time bins that were modeled with a 2 s duration each). In

624   older adults, the median in the first half was 4.50 (IQR = 1.44) and 4.50 (IQR = 0.75) in the second half. A two-

625   sample Wilcoxon test confirmed that these differences were not significant in any of the experiment halves (all

626    $p \geq .106$, Bonferroni corrected, all effect sizes $r \leq .262$). Similar results were obtained in the RSC/POS ROI with

627   a median of 4.75 (IQR = 1.00) in the first and 4.50 (IQR = 1.00) in the second half within the younger age group

628   and 4.50 (IQR = 1.19) in the first and 4.75 (IQR = 1.25) in the second half within the older age group (all $p \geq 1.00$,

629   Bonferroni corrected, all effect sizes $r \leq .049$). This suggests that the differential hippocampal and RSC/POS

630   dynamics in the two age groups are unlikely to be driven by changes in the onset/duration of the spatial

631   computations carried out in the two regions.

632   Learning-related activity changes across blocks are modulated by inter-individual differences in learning

633   within older adults

634   Behavioral performance of older adults varied substantially, with some of them showing hidden learning

635   states similar to younger adults while others showed very little performance improvements. Therefore, we

636   next included the individual's learning sub-group as covariate in the second-level analysis to examine in which

637   regions learning-related activity changes across blocks differed as a function of the overall learning ability of

638   the individual. In younger adults, no activations emerged within our ROIs or elsewhere in the brain. In older

639   adults, however, we found that activity changes in several regions across the entire brain, including visual

640   cortices, the cerebellum, temporal and frontal cortices, as well as the parahippocampal cortex (PHC)

641   extending to the anterior hippocampus, were more strongly related to the individual learning curves in better

642   performing groups (i.e., decreased across learning blocks, Figure 6, Figure 6-1). The learning curves of those

643   older adults who were less able to learn the layout of the environment, in contrast, were decoupled from

644   activity changes in these regions. No activations survived our correction for multiple comparisons within our

645   ROIs or across the whole-brain when testing for the interactions between age group and learning sub-group.

646   *--- insert Figure 6 here ---*

647   Age-related reduction in the inhibitory self-connection of the anterior hippocampus

648   To check whether age-related problems in spatial learning are related to changes in the intrinsic excitability of

649   the anterior hippocampus and the RSC/POS or in the coupling between the two regions, we used DCM PEB

650 (Friston et al., 2016). DCM has been successfully used to determine effective connectivity changes in the

651 hippocampus and related regions during memory processing (Gluth et al., 2015). Moreover, DCM PEB offers

652 several advantages over classical DCM variants in terms of model selection and second-level group

653 comparisons. First, instead of specifying several models at the first level and comparing their evidence, a full

654 model is estimated for each participant incorporating all parameters of interest, and Bayesian model

655 reduction (BMR) is performed to obtain posterior estimates of nested models in which parameters that do not

656 contribute to the model evidence are pruned. Second, first-level DCMs are equipped with empirical priors that

657 shrink parameter estimates towards a group mean. In this way, each subject's contribution to the group PEB

658 result is weighted by their precision. Third, applying classical inference methods to examine whether certain

659 parameters differ between groups after model inversion ignores within-subject uncertainty (i.e., variance of

660 the posterior distributions). This is circumvented in PEB by using the full posterior density over the parameters

661 from each participant's DCM to draw inferences about group level effects.

662 For each participant, we first specified and estimated a DCM between the anterior hippocampus and

663 the POS using peak coordinates from the corresponding univariate analysis. Navigational retrieval phases

664 were modeled as driving input into the network via the POS. The amount of learning per block was modeled

665 as modulatory input on the bidirectional connections between the two regions (Figure 7A). In the second-level

666 PEB model, we included age group, learning sub-group, and their interaction as covariates to determine their

667 relative influence on the connection strengths. The left panels in Figure 7 show the group mean of the average

668 connection strength before (Figure 7B) and after BMR (Figure 7D), indicating that all four parameters were

669 necessary to explain our data.

670 With respect to age group differences in connectivity, only one parameter survived BMR (second

671 panels of Figure 7B and 7D). Specifically, older compared to younger adults had a reduced inhibitory self-

672 connection strength in the anterior hippocampus, i.e., a relative disinhibition in this region. Note that for self-

673 connections in the DCM framework, parameters are expressed as log scaling parameters and that the

674 regressor representing age group was coded in a way that the resulting parameter is the amount that needs to

675 be added to the group mean to obtain the older adults' connection strength (the group mean is obtained by

676 calculating -0.5Hz * exp(-0.33698) = -0.357Hz and for older adults -0.5Hz * exp(-0.33698 + -0.039719) = -

677 0.3431Hz). Thus, our model provides evidence that the aging hippocampus seems to be more readily excited

678 by afferent activity from other regions during spatial learning. The interaction between age group and

679 learning sub-group in this model parameter also survived BMR (right panels in Figure 7B and 7D), indicating

680 that the hippocampal self-connection strength was more strongly modulated by the overall learning ability of

681 the individual in the older age group. Inspection of the fourth panel in Figure 7D indicates that the age-related

682 disinhibition in this region was attenuated in better performing individuals (see also Figure 7C for posterior

683 probabilities of each parameter). We did not find any modulatory effects of the (within-subject) amount of

684 learning per block.

685 We further performed a leave-one-out (LOO) cross-validation using the model parameter denoting

686 the self-connection strength in the anterior hippocampus to test whether this effect would be large enough to

687 predict the participants' age group. In this analysis, all but one subject were used to estimate the model

688 parameter, which was then used to evaluate the posterior belief of the model parameter in a left-out (test)

689 subject. The predicted and actual between-subject effect for each test subject was then compared to derive an

690 independent out-of-sample correlation, which was 0.29 in the current sample (p=0.01434, Figure 7E). Thus,

691 the estimated intrinsic connection strength in the anterior hippocampus during spatial learning was large

692 enough to predict the age group of a new subject above chance level.

693 *--- insert Figure 7 here ---*

694 ***Summary of the key findings***

695 At the behavioral level, we found in two separate experiments that performance improvements were

696 considerably reduced in healthy older compared to younger adults, when they were asked to retrieve the

697 spatial layout of an initially unfamiliar environment. Older adults further showed a higher uncertainty when

698 familiar locations were experienced from novel viewpoints during learning, as evidenced by a temporary

699 increase in response times. At the neural level, activity in the anterior hippocampus and RSC/POS changed

700 dynamically as a function of learning in younger adults, whereas this was not the case in older adults.

701 Importantly, a DCM PEB analysis revealed that the inhibitory self-connection of the anterior hippocampus was

702 reduced in older adults and was modulated by the overall learning ability of the individual as evidenced by an

703 interaction between age group and learning sub-group (see Figure 8 for a graphical summary of the results).

704 *--- insert Figure 8 here ---*

705 **DISCUSSION**

706 In two experiments, we show that healthy older adults, on average, have substantial problems in learning to

707 orient themselves in a novel, city-like virtual environment, in line with previous findings (Iaria et al., 2009;

708 Yamamoto and DeGirolamo, 2012). At neural levels, we could replicate earlier findings showing that activity in

709 RSC/POS increased while activity in the anterior hippocampus decreased as a function of learning in younger

710 adults (Wolbers and Büchel, 2005; Auger et al., 2015; Brodt et al., 2016), which shows that our task is suitable

711 to measure spatial learning, while using a complex photorealistic VE. In older adults, activity in these two

712 regions was decoupled from the amount of learning and did not change systematically across repeated

713 episodes in the environment. Importantly, we provide the first evidence that an increased excitability of the

714 anterior hippocampus might constitute a potential neural mechanism for cognitive mapping deficits in older

715 adults.

716 In the behavioral experiment, we additionally found that older adults had problems when locations

717 are encountered from novel directions during learning. This might be related to age-related deficits in

718 distinguishing novel from familiar input (Yassa et al., 2011; Vieweg et al., 2015) and to impairments in

719 allocentric processing, because Wiener et al. (2013) observed age-related performance declines when

720 locations were approached from novel directions during route learning. Given that viewpoint transformations

721 in spatial memory involve hippocampal computations (King et al., 2002), the behavioral results already point

722 to impaired information processing within the aging hippocampus that affects navigational learning. This

723 extends findings showing that a reduced sensitivity to changes in the environment might be linked to age-

724 related impairments in object-location binding and spatial perspective taking (Muffato et al., 2019; Segen et

725 al., 2020).

726   In the fMRI experiment, performance relied on the knowledge about the relation between the

727 participant's own position and the position of the target landmarks, while the change in viewpoints was

728 omitted. What neural mechanisms can account for the cognitive mapping deficits in older adults? The

729 learning-related activity decrease in the anterior hippocampus of younger adults was absent in older adults,

730 leading to an overall hippocampal hyperactivity. Similar effects have been observed in studies investigating

731 age-related deficits in pattern separation (Yassa et al., 2011; Reagh et al., 2018), as well as in rodent and non-

732 human primate studies on age-related changes in spatial navigation (Wilson et al., 2005; Thomé et al., 2016).

733 By examining effective connectivity, we were able to show, for the first time, that an age-related reduction in

734 the inhibitory self-connection strength of the anterior hippocampus might constitute the underlying neural

735 mechanism for the elevated signal in this region. Within the context of DCM, self-connection parameters

736 capture, at a macroscopic level, condition specific changes in the excitatory-inhibitory balance (Friston et al.,

737 2017). Because effective connectivity as inferred using DCM for fMRI is typically polysynaptic, we cannot

738 determine which class of cells or synapses underlie these effects. In memory-impaired monkeys, increased

739 firing rates in CA3 place cells have been linked to a reduced number of GABAergic inhibitory interneurons

740 (Thomé et al., 2016). Whether this is similarly the case in humans and how this is related to AD pathogenesis

741 are important questions for future research (Bi et al., 2020).

742   The age effect on the hippocampal self-connection strength was modulated by the learning ability of

743 the individual, suggesting that an increased hippocampal excitability might impair the formation of spatial

744 knowledge. Specifically, aberrant activity in the hippocampus could have affected the spatial resolution of the

745 emerging cognitive maps in older adults, in line with findings showing that (i) hippocampal lesion patients and

746 healthy older adults are impaired in forming high-resolution spatial representations when navigating novel

747 environments (Kolarik et al., 2016; Kolarik et al., 2018; Nilakantan et al., 2018), and (ii) that reducing

748 hippocampal hyperactivity with an anti-epileptic drug that targets excitatory neurotransmission improves

749 memory performance in amnestic patients (Bakker et al., 2015; see also Koh et al., 2013; Robitsek et al., 2015

750 for related findings in rodents). Critically, in the context of our task, imprecise cognitive maps will not only

751 affect self-localization but also the ability to compute allocentric vectors to the target landmarks. The latter

752 process has also been linked to computations in sub-regions of the MTL (Chadwick et al., 2015; Shine et al.,

753 2019; see also Wang et al., 2018; Høydal Ø et al., 2019).

754   In addition to hippocampal hyperactivation, older adults also exhibited a lack of learning related

755 dynamics in RSC/POS. Medial parietal cortex undergoes significant changes during aging, including increased

756 atrophy and enhanced tau deposition (Jockwitz et al., 2017; Harrison et al., 2019). Moreover, the increased

757 excitability of the aging hippocampus may impact on information processing in RSC/POS, given the close

758 reciprocal interactions between both regions. For example, Mao et al. (2018) found that bilateral hippocampal

759 lesions suppress the gradual emergence of a spatial code in the RSC. In the present study, given that RSC/POS

760 is assumed to support the anchoring of cognitive maps to external landmarks (Epstein et al., 2017), a deficient

761 anchoring may compromise older adults' ability to precisely recover their facing direction and to orient their

762 cognitive maps when approaching the intersections. Together with the imprecision in the cognitive maps,

763 both deficits are likely to contribute to the compromised pointing performance in older adults.

764 Moreover, this anchoring process should occur in parallel to self-localization in our task, because as
765 soon as an intersection was visible during navigational retrieval, participants could use the local buildings
766 and/or the geometric layout to recover both their position and their facing direction. This could explain why –
767 particularly in younger adults – the latency of the BOLD response in RSC/POS did not change over the course
768 of the experiment, because the process of reorientation could start immediately at the beginning of a trial.

769 Our univariate results differ from Moffat et al. (2006) who measured brain activity during encoding of
770 a virtual maze and reported an age-related hypoactivation in the RSC and the hippocampus. This discrepancy
771 might be related to the timepoint when activity was measured, because if younger adults were still acquiring
772 knowledge about the VE, our findings would also predict stronger hippocampal effects compared to older
773 adults. More generally, this discrepancy highlights the need to track the learning status of an individual when
774 interpreting differences in (hippocampal) BOLD responses between groups. In addition, it is important to note
775 that we focused on hemodynamic changes during retrieval in our study. Thus, overall task demands could be
776 another factor that might have contributed to our findings, because we also observed an age-related activity
777 increase in RSC/POS and hippocampus when contrasting retrieval to encoding.

778 Performance in our task was highly variable. While some older adults learned the layout of the
779 environment as quickly as younger adults, others showed continuous learning, learned very slowly, or were
780 not able to retrieve relevant information to perform the task. During MRI scanning, the amount of exposure in
781 the VE was kept constant for all participants. This allowed us to replicate earlier findings in younger adults and
782 to use this as a baseline against which we could compare the results of the older adults. Therefore, we cannot
783 determine whether low-performing older adults would just need more time for learning. However, it seems
784 unlikely that all of them would have reached the same performance level as younger adults if provided with
785 more time in the VE, because older adults already spent considerably more time in the initial familiarization
786 phase of the experiments. Using machine learning methods on MRI data of hundreds of older adults, Eavani et
787 al. (2018) described multiple phenotypes of brain agers that are characterized by specific functional and
788 structural changes. The authors described one phenotype that displays atrophy in the hippocampus,
789 decreased coherence in posterior medial parietal cortex, and an increased connectivity in the MTL. Thus, older
790 adults who show an increased excitability of the anterior hippocampus might be particularly impaired in
791 memorizing novel spatial environments.

792 Finally, by forming sub-groups of learners based on their estimated learning states and by including
793 this information in the fMRI analysis, we found that activity changes in several brain regions were decoupled
794 from the individual learning curves in those older adults who had more problems to learn. Although these
795 results should be interpreted with caution given the small sample sizes of our groups, they provide further
796 indications that hyperactivity in the aging brain does not seem to support task performance (Morcom and
797 Henson, 2018). We did not find any indications that the learning differences within older adults were related to
798 their age, sex, or their cognitive screening scores. Thus, future studies should apply additional measures, for
799 example preclinical markers for AD, to further characterize age-related deficits in spatial learning and,
800 specifically, why these abilities are preserved in some older adults.

801   Taken together, increased excitability of the anterior hippocampus, together with aberrant RSC/POS

802   functioning, provides a novel explanation why older adults experience problems with forming accurate spatial

803   representations of a novel environment. In addition, our findings add to a growing body of evidence

804   associating hyperactivity in the hippocampus to memory impairments in aging.

REFERENCES

Auger SD, Zeidman P, Maguire EA (2015) A central role for the retrosplenial cortex in de novo environmental learning. eLife 4:e09031. doi: 10.7554/eLife.09031

Bakker A, Albert MS, Krauss G, Speck CL, Gallagher M (2015) Response of the medial temporal lobe network in amnestic mild cognitive impairment to therapeutic intervention assessed by fMRI and memory task performance. NeuroImage: Clinical 7:688-698.

Barnes CA, Suster MS, Shen J, McNaughton BL (1997) Multistability of cognitive maps in the hippocampus of old rats. Nature 388:272-275.

Behzadi Y, Restom K, Liau J, Liu TT (2007) A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37:90-101.

Berens P (2009) CircStat: A MATLAB toolbox for circular statistics. J Stat Softw 31:1-21.

Bi D, Wen L, Wu Z, Shen Y (2020) GABAergic dysfunction in excitatory and inhibitory (E/I) imbalance drives the pathogenesis of Alzheimer's disease. Alzheimer's & Dementia 16:1312-1329.

Braak H, Del Tredici K (2015) The preclinical phase of the pathological process underlying sporadic Alzheimer's disease. Brain 138:2814-2833.

Brett M, Anton J-L, Valabregue R, Poline JB (2002) Region of interest analysis using an SPM toolbox. Neuroimage 16.

Brodt S, Pöhlchen D, Flanagin VL, Glasauer S, Gais S, Schönauer M (2016) Rapid and independent memory formation in the parietal cortex. Proc Natl Acad Sci USA 113:13251-13256.

Bzdok D, Heeger A, Langner R, Laird AR, Fox PT, Palomero-Gallagher N, Vogt BA, Zilles K, Eickhoff SB (2015) Subspecialization in the human posterior medial cortex. Neuroimage 106:55-71.

Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P, Riddell A (2017) Stan: A probabilistic programming language. J Stat Softw 76:1-32.

Chadwick MJ, Jolly AEJ, Amos DP, Hassabis D, Spiers HJ (2015) A goal direction signal in the human entorhinal/subicular region. Curr Biol 25:87-92.

Commandeur J, Koopman SJ (2007) Introduction to state space time series analysis. Oxford: Oxford University Press.

Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 9:179-194.

Daunizeau J, Stephan KE, Friston KJ (2012) Stochastic dynamic causal modelling of fMRI data: Should we care about neural noise? Neuroimage 62:464-481.

Destrieux C, Fischl B, Dale AM, Halgren E (2010) Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 53:1-15.

Diersch N, Wolbers T (2019) The potential of virtual reality for spatial navigation research across the adult lifespan. The Journal of Experimental Biology 222. doi: 10.1242/jeb.187252

Eavani H, Habes M, Satterthwaite TD, An Y, Hsieh M-K, Honnorat N, Erus G, Doshi J, Ferrucci L, Beason-Held LL, Resnick SM, Davatzikos C (2018) Heterogeneity of structural and functional imaging patterns of advanced brain aging revealed via machine learning methods. Neurobiol Aging 71:41-50.

Epstein RA (2008) Parahippocampal and retrosplenial contributions to human spatial navigation. Trends Cogn Sci 12:388-396.

845 Epstein RA, Patai EZ, Julian JB, Spiers HJ (2017) The cognitive map in humans: Spatial navigation and beyond.
846 Nat Neurosci 20:1504-1513.

847 Esteban O, Birman D, Schaer M, Koyejo OO, Poldrack RA, Gorgolewski KJ (2017) MRIQC: Advancing the
848 automatic prediction of image quality in MRI from unseen sites. PLOS ONE 12:e0184661. doi:
849 10.1371/journal.pone.0184661

850 Esteban O, Markiewicz C, Blair RW, Moodie C, Isik AI, Erramuzpe Aliaga A, Kent J, Goncalves M, DuPre E,
851 Snyder M, Oya H, Ghosh SS, Wright JD, Durnez J, Poldrack RA, Gorgolewski KJ (2019) FMRIPrep: A robust
852 preprocessing pipeline for functional MRI. Nature Methods 16:111-116.

853 Faul F, Erdfelder E, Lang A-G, Buchner A (2007) G*Power 3: A flexible statistical power analysis program for
854 the social, behavioral, and biomedical sciences. Behavior Research Methods 39:175-191.

855 Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis: II: Inflation, flattening, and a surface-
856 based coordinate system. Neuroimage 9:195-207.

857 Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D,
858 Klaveness S, Montillo A, Makris N, Rosen BR, Dale AM (2002) Whole brain segmentation: Automated labeling
859 of neuroanatomical structures in the human brain. Neuron 33:341-355.

860 Friston KJ, Preller KH, Mathys C, Cagnan H, Heinzle J, Razi A, Zeidman P (2017) Dynamic causal modelling
861 revisited. Neuroimage. doi: 10.1016/j.neuroimage.2017.02.045

862 Friston KJ, Litvak V, Oswal A, Razi A, Stephan KE, van Wijk BCM, Ziegler G, Zeidman P (2016) Bayesian model
863 reduction and empirical Bayes for group (DCM) studies. Neuroimage 128:413-431.

864 Gelman A, Shirley K (2011) Inference from simulations and monitoring convergence. In: Handbook of Markov
865 Chain Monte Carlo (Brooks S, Gelman A, Jones GL, Meng X-L, eds), pp 163-174. Boca Raton, FL: Chapman
866 Hall.

867 Gluth S, Sommer T, Rieskamp J, Büchel C (2015) Effective connectivity between Hippocampus and
868 ventromedial Prefrontal Cortex controls preferential choices from memory. Neuron 86:1078-1090.

869 Gorgolewski KJ et al. (2016) The brain imaging data structure, a format for organizing and describing outputs
870 of neuroimaging experiments. Scientific Data 3:160044. doi: 10.1038/sdata.2016.44

871 Grady CL (2012) The cognitive neuroscience of ageing. Nat Rev Neurosci 13:491-505.

872 Grinband J, Steffener J, Razlighi QR, Stern Y (2017) BOLD neurovascular coupling does not change
873 significantly with normal aging. Hum Brain Mapp 38:3538-3551.

874 Harrison TM, La Joie R, Maass A, Baker SL, Swinnerton K, Fenton L, Mellinger TJ, Edwards L, Pham J, Miller
875 BL, Rabinovici GD, Jagust WJ (2019) Longitudinal tau accumulation and atrophy in aging and alzheimer
876 disease. Ann Neurol 85:229-240.

877 Høydal Ø A, Skytøen ER, Andersson SO, Moser MB, Moser EI (2019) Object-vector coding in the medial
878 entorhinal cortex. Nature 568:400-404.

879 Iaria G, Palermo L, Committeri G, Barton JJS (2009) Age differences in the formation and use of cognitive
880 maps. Behav Brain Res 196:187-191.

881 Jockwitz C, Caspers S, Lux S, Jütten K, Schleicher A, Eickhoff SB, Amunts K, Zilles K (2017) Age- and function-
882 related regional changes in cortical folding of the default mode network in older adults. Brain Struct Funct
883 222:83-99.

884 Jones E, Oliphant T, Peterson P (2001) SciPy: Open source scientific tools for Python. In: http://www.scipy.org.

885 King JA, Burgess N, Hartley T, Vargha-Khadem F, O'Keefe J (2002) Human hippocampus and viewpoint
886 dependence in spatial memory. Hippocampus 12:811-820.

887 Kobayashi Y, Amaral DG (2003) Macaque monkey retrosplenial cortex: II. Cortical afferents. The Journal of
888 Comparative Neurology 466:48-79.

889 Koh MT, Rosenzweig-Lipson S, Gallagher M (2013) Selective GABAA α 5 positive allosteric modulators
890 improve cognitive function in aged rats with memory impairment. Neuropharmacology 64:145-152.

891 Kolarik BS, Baer T, Shahlaie K, Yonelinas AP, Ekstrom AD (2018) Close but no cigar: Spatial precision deficits
892 following medial temporal lobe lesions provide novel insight into theoretical models of navigation and
893 memory. Hippocampus 28:31-41.

894 Kolarik BS, Shahlaie K, Hassan AS, Borders AA, Kaufman KC, Gurkoff G, Yonelinas AP, Ekstrom AD (2016)
895 Impairments in precision, rather than spatial strategy, characterize performance on the virtual Morris Water
896 Maze: A case study. Neuropsychologia 80:90-101.

897 Konishi K, Etchamendy N, Roy S, Marighetto A, Rajah N, Bohbot VD (2013) Decreased functional magnetic
898 resonance imaging activity in the hippocampus in favor of the caudate nucleus in older adults tested in a
899 virtual navigation task. Hippocampus 23:1005-1014.

900 Leal SL, Landau SM, Bell RK, Jagust WJ (2017) Hippocampal activation is associated with longitudinal amyloid
901 accumulation and cognitive decline. eLife 6:e22978. doi: 10.7554/eLife.22978

902 Lester AW, Moffat SD, Wiener JM, Barnes CA, Wolbers T (2017) The aging navigational system. Neuron
903 95:1019-1035.

904 Li B, Daunizeau J, Stephan KE, Penny WD, Hu D, Friston K (2011) Generalised filtering and stochastic DCM for
905 fMRI. Neuroimage 58:442-457.

906 Luis CA, Keegan AP, Mullan M (2009) Cross validation of the Montreal Cognitive Assessment in community
907 dwelling older adults residing in the Southeastern US. Int J Geriatr Psychiatry 24:197-201.

908 Mao D, Neumann AR, Sun J, Bonin V, Mohajerani MH, McNaughton BL (2018) Hippocampus-dependent
909 emergence of spatial sequence coding in retrosplenial cortex. Proc Natl Acad Sci USA. doi:
910 10.1073/pnas.1803224115

911 Miller AMP, Vedder LC, Law LM, Smith DM (2014) Cues, context, and long-term memory: The role of the
912 retrosplenial cortex in spatial cognition. Front Hum Neurosci 8:1-15. doi: 10.3389/fnhum.2014.00586

913 Moffat SD, Elkins W, Resnick SM (2006) Age differences in the neural systems supporting human allocentric
914 spatial navigation. Neurobiol Aging 27:965-972.

915 Morcom AM, Henson RNA (2018) Increased Prefrontal activity with aging reflects nonspecific neural responses
916 rather than compensation. J Neurosci 38:7303-7313.

917 Muffato V, Hilton C, Meneghetti C, De Beni R, Wiener JM (2019) Evidence for age-related deficits in object-
918 location binding during place recognition. Hippocampus 29:971-979.

919 Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H
920 (2005) The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. J Am
921 Geriatr Soc 53:695-699.

922 Nilakantan AS, Bridge DJ, VanHaerents S, Voss JL (2018) Distinguishing the precision of spatial recollection
923 from its success: Evidence from healthy aging and unilateral mesial temporal lobe resection.
924 Neuropsychologia 119:101-106.

925 Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia
926 9:97-113.

927 Patai EZ, Javadi A-H, Ozubko JD, O'Callaghan A, Ji S, Robin J, Grady C, Winocur G, Rosenbaum RS,
928 Moscovitch M, Spiers HJ (2019) Hippocampal and Retrosplenial goal distance coding after long-term
929 consolidation of a real-world environment. Cereb Cortex. doi: 10.1093/cercor/bhz044

930 Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R,
931 Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn:
932 Machine learning in python. J Mach Learn Res 12:2825-2830.

933 Poppenk J, Evensmoen HR, Moscovitch M, Nadel L (2013) Long-axis specialization of the human
934 hippocampus. Trends Cogn Sci 17:230-240.

935 Reagh ZM, Noche JA, Tustison NJ, Delisle D, Murray EA, Yassa MA (2018) Functional imbalance of
936 anterolateral entorhinal cortex and hippocampal dentate/CA3 underlies age-related object pattern separation
937 deficits. Neuron 97:1187-1198.

938 Robitsek J, Ratner MH, Stewart T, Eichenbaum H, Farb DH (2015) Combined administration of levetiracetam
939 and valproic acid attenuates age-related hyperactivity of CA3 place cells, reduces place field area, and
940 increases spatial information content in aged rat hippocampus. Hippocampus 25:1541-1555.

941 Rosenbaum RS, Winocur G, Binns MA, Moscovitch M (2012) Remote spatial memory in aging: all is not lost.
942 Front Ag Neurosci 4:1-10. doi: 10.3389/fnagi.2012.00025

943 Rugg MD, Morcom AM (2005) The relationship between brain activity, cognitive performance and aging: The
944 case of memory. In: Cognitive neuroscience of aging: Linking cognitive and cerebral aging (Cabeza R, Nyberg
945 L, Park DC, eds), pp 132-154. New York, NY: Oxford University Press.

946 Segen V, Avraamides MN, Slattery TJ, Wiener JM (2020) Age-related differences in visual encoding and
947 response strategies contribute to spatial memory deficits. PsyArXiv. doi: 10.31234/osf.io/4qdey

948 Shine JP, Valdés-Herrera JP, Tempelmann C, Wolbers T (2019) Evidence for allocentric boundary and goal
949 direction information in the human entorhinal cortex and subiculum. Nature Communications 10:4004. doi:
950 10.1038/s41467-019-11802-9

951 Smith AC, Wirth S, Suzuki WA, Brown EN (2007) Bayesian analysis of interleaved learning and response bias in
952 behavioral experiments. J Neurophysiol 97:2516-2524.

953 Stan Development Team (2017) PyStan: the Python interface to Stan. In, 2.16.0.0. Edition. http://mc-stan.org.

954 Thomé A, Gray DT, Erickson CA, Lipa P, Barnes CA (2016) Memory impairment in aged primates is associated
955 with region-specific network dysfunction. Mol Psychiatry 21:1257.

956 Vehtari A, Gelman A, Gabry J (2017) Practical Bayesian model evaluation using leave-one-out cross-validation
957 and WAIC. Statistics and Computing 27:1413-1432.

958 Vieweg P, Stangl M, Howard LR, Wolbers T (2015) Changes in pattern completion – A key mechanism to
959 explain age-related recognition memory deficits? Cortex 64:343-351.

960 Wang C, Chen X, Lee H, Deshmukh SS, Yoganarasimha D, Savelli F, Knierim JJ (2018) Egocentric coding of
961 external items in the lateral entorhinal cortex. Science 362:945-949.

962 Wiener JM, de Condappa O, Harris MA, Wolbers T (2013) Maladaptive bias for extrahippocampal navigation
963 strategies in aging humans. J Neurosci 33:6012-6017.

964 Wilson IA, Ikonen S, Gallagher M, Eichenbaum H, Tanila H (2005) Age-associated alterations of hippocampal
965 place cells are subregion specific. J Neurosci 25:6877-6886.

966   Wolbers T, Büchel C (2005) Dissociable Retrosplenial and Hippocampal contributions to successful formation
967   of survey representations. J Neurosci 25:3333-3340.

968   Yamamoto N, DeGirolamo G (2012) Differential effects of aging on spatial learning through exploratory
969   navigation and map reading. Front Ag Neurosci 4:1-7. doi: 10.3389/fnagi.2012.00014

970   Yassa MA, Lacy JW, Stark SM, Albert MS, Gallagher M, Stark CEL (2011) Pattern separation deficits associated
971   with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. Hippocampus
972   21:968-979.

973   Zeidman P, Maguire EA (2016) Anterior hippocampus: The anatomy of perception, imagination and episodic
974   memory. Nat Rev Neurosci 17:173-182.

975   Zeidman P, Jafarian A, Seghier ML, Litvak V, Cagnan H, Price CJ, Friston KJ (2019a) A guide to group effective
976   connectivity analysis, part 2: Second level analysis with PEB. Neuroimage 200:12-25.

977   Zeidman P, Jafarian A, Corbin N, Seghier ML, Razi A, Price CJ, Friston KJ (2019b) A guide to group effective
978   connectivity analysis, part 1: First level analysis with DCM for fMRI. Neuroimage. doi:
979   10.1016/j.neuroimage.2019.06.031

980 **FIGURE LEGENDS**

981 **Figure 1.** Spatial learning task.

982 (A) Procedure of the fMRI experiment. After a familiarization phase outside of the scanner, eight retrieval

983 phases, each comprising 8 navigational retrieval trials and 4 control trials, alternated with 7 encoding phases

984 during scanning. In the behavioral experiment, the structure was the same except that 12 navigational

985 retrieval trials per learning block were completed while the control trials were omitted. (B) Layout of the

986 virtual environment (VE). The VE resembled a typical German historic city center and consisted of four

987 interconnected intersections (I1-I4) that could be reached from 4 directions (D1-D4). At two intersections, a

988 town hall (T1) and a church (T2) were placed at the end of one of the outgoing streets that served as target

989 landmarks in the navigational retrieval trials. Yellow arrows exemplify one encoding tour that started from

990 one of the target landmarks in clockwise or counterclockwise direction (a short segment of one tour is shown

991 in Video 1). (C) Structure of one example navigational retrieval trial to measure spatial learning. After fixation,

992 participants were passively transported towards one of four intersections in the VE starting from one of the

993 four streets leading towards that intersection (see Video 2). Movement stopped at the center of the

994 intersection, a red crosshair appeared, and participants were asked to move the crosshair in the direction of

995 the respective target landmark. During the entire duration of the trial, a picture cue of the target landmark

996 was displayed at the bottom of the screen, and the background was obscured by fog to prevent seeing the

997 target landmarks. In the fMRI experiment, an additional jittered interval of 1 s (still phase) was added after the

998 travel phase/before the crosshair appeared on screen.

999 **Figure 2.** Bayesian state-space model to estimate the subject-specific hidden learning state per learning block

1000 (see Figure 2-1 for the model code). Results of the posterior predictive checks of the model for representative

1001 individuals from each learning sub-group in the fMRI experiment and a histogram of the individuals' loo

1002 differences for the comparison of the Bayesian state-space model to an alternative model that estimated the

1003 individuals' learning state trial-wise is depicted in Figure 2-2.

1004 **Figure 3.** Performance data in the behavioral experiment.

1005 (A) Average absolute pointing errors and (B) response times across the eight learning blocks in older (solid

1006 line) and younger adults (dashed line; highlighted in grey is the 4th and 5th learning block where the change in

1007 directions took place from which the intersections were approached). Error bars denote standard errors of the

1008 means (SE). See also Figure 3-1A-B for average pointing errors per learning block for each participant in each

1009 age group. (C) Mean estimated performance improvement (hidden learning state) of each participant in the

1010 older (orange) and younger (grey) age group, including the standard deviation (SD) of the posterior

1011 distributions (shaded area) across learning blocks. (D) Logistic regression results classifying age group

1012 membership based on two behavioral performance features, i.e., the mean amount of learning across the

1013 experiment and the increase in response times after the first half of the experiment. Shaded lines depict the

1014 probability of being classified as a younger adult.

1015 **Figure 4.** Performance data in the fMRI experiment.

1016 (A) Average absolute pointing errors across the eight learning blocks in older (solid line) and younger adults

1017 (dashed line). Error bars denote standard errors of the means (SE). See also Figure 3-1C-D for average pointing

1018 errors per learning block for each participant in each age group. (B) Mean estimated performance

1019 improvement (hidden learning state) of each participant in the older (orange) and younger (grey) age group,

1020 including the standard deviation (SD) of the posterior distributions (shaded area) across learning blocks. (C)

1021 Learning sub-groups as identified by a K-means clustering algorithm based on the individuals' overall amount

1022 of learning and its SD, as determined by the difference of the latent state distributions of the last and first

1023 learning block. See Figure 4-1 for difference distributions, learning state estimates, and performance data per

1024 learning block for representative individuals from each learning sub-group and Figure 4-2 for the results of the

1025 same clustering analysis within the sample of the behavioral experiment.

1026 **Figure 5.** Interaction effects between age group and the amount of learning per block during navigational

1027 retrieval. Age-related differences in (A) hippocampal activity decreases and (B) RSC/POS activity increases

1028 across the experiment. Activations are displayed on the 2009 nonlinear asymmetric MNI template that was

1029 used for normalization (p < 0.05, FWE-corrected for the respective ROI). Plots depict average parameter

1030 estimates of the respective peak voxels per learning block in selected clusters for each age group. Error bars

1031 indicate the across-subject standard error of the mean. See Table 1 for the spatial coordinates of the local

1032 maxima in the hippocampus and RSC/POS ROIs and Table 1-1 for significantly activated clusters elsewhere in

1033 the brain.

1034 **Figure 6.** Differential activity changes in relation to the amount of learning per block between learning sub-

1035 groups in the older age group. Activations are displayed on the 2009 nonlinear asymmetric MNI template that

1036 was used for normalization (p < 0.05, FWE-corrected). See Figure 6-1 for the spatial coordinates of the local

1037 maxima.

1038 **Figure 7.** Results of the DCM PEB analysis. (A) First-level DCM specification to determine average connectivity

1039 within and between anterior hippocampus and POS. Navigational retrieval phases were modeled as driving

1040 input entering the cortical network via the POS, and the amount of learning per block was included as

1041 modulatory input on the bidirectional connections between the regions. Estimated Parameters (1: self-

1042 connection POS, 2: POS – hippocampus connection, 3: hippocampus – POS connection, 4: self-connection

1043 hippocampus) (B) before and (D) after Bayesian model reduction (BMR) for each covariate (age group,

1044 learning group, interaction between age group and learning group) in the second-level PEB model. Grey bars

1045 represent parameter means and pink lines their 95% confidence intervals. The parameters for self-connections

1046 (parameter 1 and 4) are expressed as log scaling parameters that can be converted to Hz using $x\_Hz = -0.5 *$

1047 $exp(x)$ whereby x is the log scaling parameter and -0.5 Hz the prior. (C) Posterior probabilities per parameter

1048 for each second-level covariate after BMR, (E) Predicted age group of each participant as derived from a LOO

1049 cross-validation scheme based on the estimated self-connection strength in the anterior hippocampus.

1050 **Figure 8.** Key findings of the two experiments.

1051 **Video 1.** Exemplary segment of one of the tours through the virtual environment in the encoding phases of

1052 the two experiments.

1053   **Video 2.** Example for a pointing trial in the retrieval phases of the two experiments.

1054 **TABLES**

1055 **Table 1.** Spatial coordinates of the local maxima in the hippocampus and RSC/POS ROIs in the fMRI analyses

1056 on age-related differences in neural activation patterns (p < 0.05, FWE-corr.). See Table 1-1 for significantly

1057 activated clusters elsewhere in the brain.

Brain region	Cluster size	MNI coordinate			Z-score
		x	y	z	
A) Increased activity in older compared to younger adults during navigation vs. control					
L POS	83	-3	-60	34	4.56
L RSC		-6	-57	21	4.32
L Hippocampus	72	-21	-18	-12	4.76
		-30	-15	-22	4.21
B) Reduced activity in older compared to younger adults during navigation vs. control					
R POS	22	12	-69	54	4.00
		18	-69	57	3.94
		15	-75	51	3.25
R POS	55	27	-60	24	3.96
C) Increased activity in older compared to younger adults during retrieval vs. encoding					
R POS	27	12	-51	34	4.03
L Hippocampus	33	-30	-15	-15	4.08
R Hippocampus	20	24	-12	-12	3.97
D) Age-group differences in learning-related [1]activity decreases					
R Hippocampus	20	24	-18	-15	4.55
E) Age-group differences in learning-related [1]activity increases					
L POS	462	-9	-66	24	5.93
		-24	-72	47	5.37
L RSC		-18	-57	1	4.14
		-6	-63	11	3.98
R POS	148	24	-69	47	4.95
		21	-72	54	4.73
R RSC	205	9	-57	4	4.82
		9	-63	21	4.59

1058

1   **TABLE AND FIGURE LEGENDS**

2   **Table 1-1, related to Table 1 and Figure 5:** Spatial coordinates of the local maxima in the whole-brain fMRI analyses on
3   age-related differences in neural activation patterns (p < 0.05, whole-brain FWE-corr.).

4   **Figure 2-1, related to Figure 2 and <u>Bayesian Modeling of Performance Data</u> section:** Stan code of the Bayesian state-
5   space model.

6   **Figure 2-2, related to Figure 2 and <u>Bayesian Modeling of Performance Data</u> section:** Results of the posterior predictive
7   checks of the Bayesian state-space model for representative individuals from each learning sub-group (A: top learner
8   young – E: non-learner old; see <u>Performance Clustering</u> section; the posterior predictive samples distribution, y_{rep}, plotted
9   together with the observed data points, y, per learning block) and (F) histogram of the individuals' loo differences for the
10   comparison of the Bayesian state-space model incorporating the effects of the responses, η, to an alternative model that
11   estimated the individuals' learning state trial-wise. More positive values indicate a better fit of the first model.

12   **Figure 3-1, related to Figure 3 and 4:** Average absolute pointing errors per learning block for each participant in (A) the
13   younger and (B) the older age group in the behavioral experiment, and for each participant in (C) the younger and (D) the
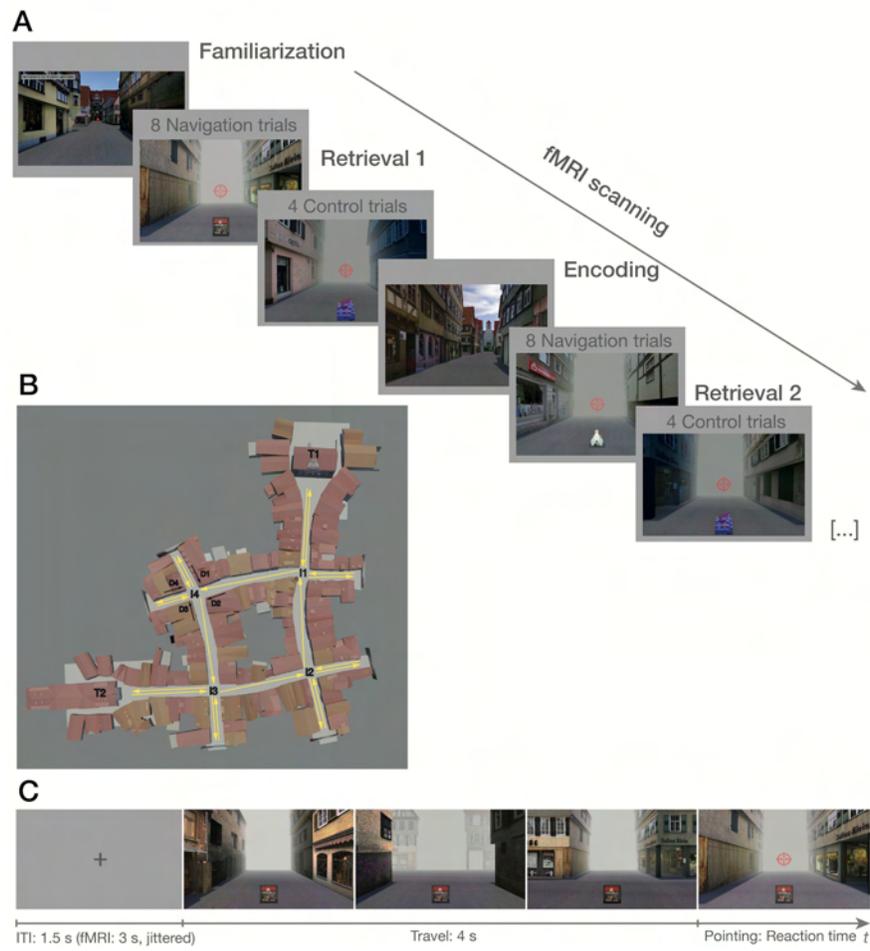14   older age group in the fMRI experiment.

15   **Figure 4-1, related to Figure 4 and <u>Performance Clustering</u> section:** Definition of learning sub-groups. Hidden learning
16   states (including SD) and trial-wise performance data per learning block (left) and the latent state distributions of the last
17   and first learning block plotted together with the difference distribution (right) from representative individuals from each
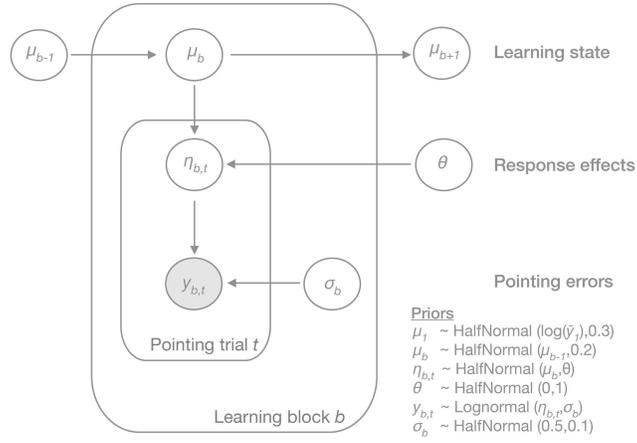18   learning sub-group in the fMRI experiment.

19   **Figure 4-2, related to Figure 4 and <u>Performance Clustering</u> section:** Learning sub-groups in the behavioral experiment as
20   identified by a K-means clustering algorithm based on the individuals' overall amount of learning and its SD, as
21   determined by the difference of the latent state distributions of the last and first learning block. Results are shown for (A) 5
22   and (B) 6 learning clusters that yielded similar silhouette scores (respective mean silhouette scores per tested cluster
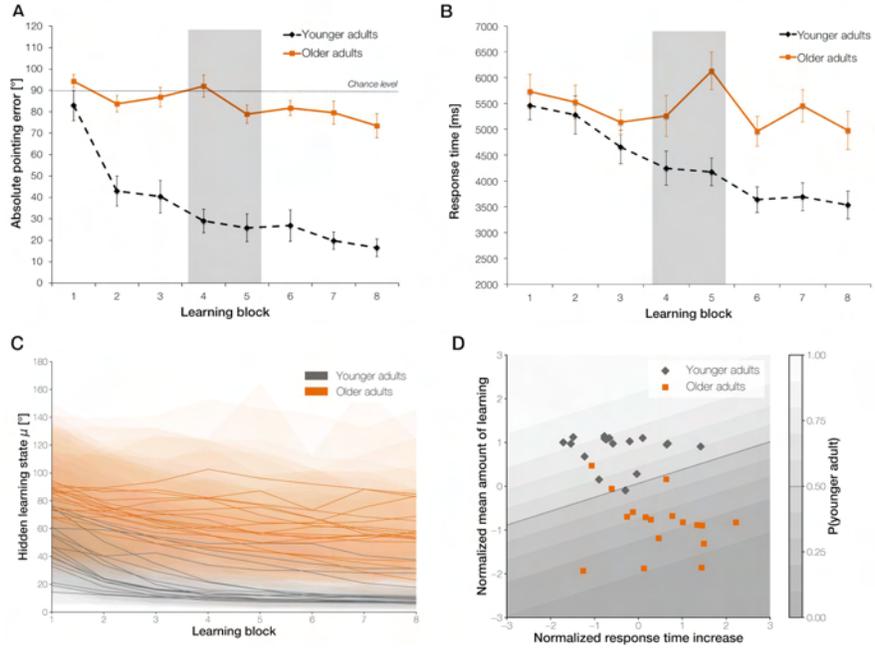23   number: 3: 0.232, 4: 0.293, 5: 0.400, 6: 0.404, 7: 0.370).

24   **Figure 4-3, related to Figure 4 and <u>Performance Clustering</u> section:** Key demographics of the learning sub-groups within
25   each age group in the fMRI experiment.

26   **Figure 5-1, related to Figure 5 and Table 1:** Spatial coordinates of the local maxima in the whole-brain fMRI analyses on
27   age-related differences in neural activation patterns (p < 0.05, whole-brain FWE-corr.).

28   **Figure 6-1, related to Figure 6:** Spatial coordinates of the local maxima in the fMRI analyses on inter-individual
29   differences in neural activation patterns across learning blocks within older adults (p < 0.05, FWE-corr.).
30

A Familiarization

8 Navigation trials

Retrieval 1

4 Control trials

fMRI scanning

Encoding

8 Navigation trials

Retrieval 2

4 Control trials

[...]

B

C

ITI: 1.5 s (fMRI: 3 s, jittered)   Travel: 4 s   Pointing: Reaction time *t*

μ_{b-1} μ_b μ_{b+1} Learning state

$\eta_{b,t}$ θ Response effects

$y_{b,t}$ σ_b Pointing errors

Pointing trial t

Learning block b

Priors

$\mu_1 \sim$ HalfNormal $(\log(\bar{y}_1), 0.3)$
$\mu_b \sim$ HalfNormal $(\mu_{b-1}, 0.2)$
$\eta_{b,t} \sim$ HalfNormal (μ_b, θ)
$\theta \sim$ HalfNormal $(0,1)$
$y_{b,t} \sim$ Lognormal $(\eta_{b,t}, \sigma_b)$
$\sigma_b \sim$ HalfNormal $(0.5, 0.1)$

A

Absolute pointing error [°]

- Younger adults
- Older adults

Chance level

Learning block (fMRI)

B

Hidden learning state μ [°]

Younger adults
Older adults

Learning block

C

SD of the learning state difference [°]

Learning state difference [°]

▲ top learner
● good learner
♦ intermediate learner
▼ weak learner
• non-learner
■ older adults
■ younger adults

A

x = 24

Hippocampus

B

x = -9   y = -63

POS   RSC

A

Estimated parameters:
1: self-connection POS
2: POS – hippocampus connection
3: hippocampus – POS connection
4: self-connection hippocampus

B **Full model**

C

D **Reduced model**

E corr(df:55) = 0.29: p = 0.01434