

Hippocampal integrity predicts individual differences in memory- training responsiveness in healthy ageing

Raquel Lezama

`raquel.lezamacantero@unich.it`

University "G. d'Annunzio" of Chieti- Pescara

Rosanna Pinto

University "G. d'Annunzio" of Chieti- Pescara

Davide Di Censo

University "G. d'Annunzio" of Chieti- Pescara

Gabriella Monopoli

University "G. d'Annunzio" of Chieti- Pescara

Mauro Gianni Perrucci

University "G. d'Annunzio" of Chieti- Pescara

Beth Fairfield

University Federico II

Stefano Delli Pizzi

University "G. d'Annunzio" of Chieti- Pescara

Giorgia Committeri

University "G. d'Annunzio" of Chieti- Pescara

Massimo Caulo

University "G. d'Annunzio" of Chieti- Pescara

Richard Geoffrey Wise

University "G. d'Annunzio" of Chieti- Pescara

Carlo Sestieri

University "G. d'Annunzio" of Chieti- Pescara

Research Article

Keywords: Memory training, healthy ageing, hippocampus, Method of Loci (MoL), Soma and Neurite Density Imaging (SANDI)

Posted Date: October 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11141574/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: Competing interest reported. R.P. is a doctoral candidate whose PhD fellowship is co-funded by the University "G. d'Annunzio" of Chieti–Pescara and by Maggioli S.p.A. The Method of Loci training platform used in this study was developed by Maggioli S.p.A. under a paid service agreement, at the request of and to specifications set by the research team, specifically for this study; it is not a publicly available commercial product. The other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Abstract

Episodic memory decline is among the most prevalent consequences of healthy ageing, and Method of Loci (MoL) training offers a promising avenue for enhancement. Yet variability in responsiveness remains poorly understood. Prior work points to hippocampal integrity before training as a predictor, but its spatial organization and tissue-level nature remain unresolved. Here, thirty healthy older adults (65–75 years) underwent high-resolution structural and diffusion MRI before a 4-week adaptive home-based MoL intervention. Replicating prior findings, left but not right hippocampal volume predicted training responsiveness, with significant associations in CA2/3 and CA4/dentate gyrus. Extending these predictors along the hippocampal long axis, significant subregions also included the body and tail, with effects concentrated posteriorly. Microstructural analyses using Soma and Neurite Density Imaging (SANDI) showed that only neurite signal fraction (f_{neurite}), not soma signal fraction (f_{soma}), predicted training benefit, again in the posterior but not anterior hippocampus. Both effects generalized to held-out participants under leave-one-out cross-validation. These results spatially and biologically characterize the role of left hippocampal integrity in MoL training responsiveness, implicating neurite-related properties as a substrate of individual differences in cognitive plasticity and informing personalized cognitive interventions in older adults.

Introduction

Episodic memory decline is among the most prevalent consequences of healthy ageing, reducing the ability to encode and retrieve personally experienced events together with their associated spatial and temporal contexts (Tulving and Thomson 1973) from approximately 60 years of age (Nyberg et al. 2024). At the same time, age-related decline varies considerably across individuals, making the factors underlying cognitive plasticity and training responsiveness a key focus in healthy ageing research (Baykara et al. 2021; Gallen and D'Esposito 2019; Mendonça et al. 2022; Tucker-Drob and Salthouse 2011; Zehnder et al. 2009).

Memory training interventions improve memory performance in healthy older adults by teaching strategies that facilitate encoding and retrieval (Gross et al. 2012). Among these, the Method of Loci (MoL) requires the generation of spatially organized mnemonic representations and sequential retrieval of associated items, processes that rely heavily on hippocampal functions like spatial-associative memory, navigation and sequential retrieval (Twomey and Kroneisen 2021; Ngo et al. 2016). A growing body of evidence supports the effectiveness of the MoL in cognitive training programs (Ondřej 2025). However, substantial heterogeneity in training gain has been consistently reported, with evidence suggesting that this variability is driven more by individual differences among participants within studies than by methodological differences between studies (Guye et al. 2017; Jaeggi et al. 2014; Traut et al. 2021). This within-group variability is particularly relevant in older adults, in whom memory training holds the greatest applied value. Characterizing the factors that account for these individual differences may therefore be more informative than contrasting younger and older groups.

Regarding cognitive factors that may explain this variability, educational attainment has been the most widely adopted and extensively validated proxy of cognitive reserve in healthy ageing (Elkana et al. 2025; Simfukwe et al. 2025; Stern et al. 2020) and has been consistently associated with a lower risk of cognitive decline and dementia (Meng and D'Arcy 2012; Nelson et al. 2021). Although cognitive proxies explain some variability in training outcomes, neural markers may more directly capture the biological substrates underlying successful cognitive training (Baykara et al. 2021). The hippocampus is a particularly strong candidate region, as the MoL relies on spatially organized representations and sequential retrieval.

Hippocampal integrity before training has indeed emerged as a consistent predictor of MoL-related benefit. Engvig et al. (2012) found that, in older adults with subjective memory complaints, greater left hippocampal volume, and larger left CA2/3 and CA4/DG volumes specifically, were associated with greater verbal recall improvement following an 8-week MoL intervention. This regional specificity is consistent with evidence that hippocampal subfields support complementary mnemonic operations: whereas CA2/3 and CA4/DG are central to pattern separation, sequence representation, and the formation of detailed spatial associations, CA1 contributes primarily to contextual reinstatement and temporal order coding (Dimsdale-Zucker et al. 2018; Hainmueller and Bartos 2020; Liu et al. 2022; Rolls 2018). Fimbria integrity supports hippocampal-dependent spatial, associative, and temporal-context memory, while the subiculum participates in efficient recollection and the presubiculum is related to spatial navigation and contextual organization processes, although human evidence for these subregions in memory remains limited (Eldridge et al. 2005; Gazes et al. 2019; Read et al. 2023; Zammit et al. 2017). The MoL places particularly heavy demands on the former, i.e., binding distinct objects to specific, non-overlapping spatial locations and maintaining their sequential order, which may explain why CA2/3 and CA4/DG volume is a stronger predictor of training responsiveness.

In a cognitively healthy sample including young and older participants, Bråthen et al. (2018) subsequently confirmed that hippocampal volume before training predicts memory gains within a multimodal framework including cortical thickness and resting-state functional connectivity, with hippocampal volume emerging as a unique predictor of training-related improvement selectively in older adults. However, while subfields were not investigated in this latter study, the subfield results of Engvig et al. (2012) were obtained using an early segmentation protocol that has since been substantially refined, leaving open whether their specificity pattern holds under current delineation methods. Moreover, neither study examined whether predictive findings hold under cross-validation, a particularly important consideration given the modest samples typical of this literature.

Another important question unaddressed by prior studies concerns the degree of training responsiveness along the longitudinal hippocampal axis. Anterior and posterior hippocampal regions differ systematically in both their extrinsic connectivity and the granularity of the representations they support (Schlichting et al. 2015). Whereas the anterior hippocampus is associated with coarse, gist-like representations, the posterior hippocampus supports fine-grained spatial and contextual information, which is critical for the ordered retrieval of episodic sequences (Evensmoen et al. 2013; Grady 2020;

Poppenk et al. 2013; Strange et al. 2014; Zeidman and Maguire, 2016). These functions closely match the cognitive demands of the MoL, which relies on constructing a detailed mental route and retrieving item–location associations in a specific sequence. Moreover, longitudinal-axis parcellation has been shown to provide a more anatomically robust characterization of hippocampal organization than subfield segmentation (Daugherty et al. 2026; Ding and Van Hoesen 2015; but see Whelan et al. 2016). However, to our knowledge, no study has examined whether the benefits of MoL training in older adults are differentially associated with anterior and posterior hippocampal regions.

A key limitation of the existing literature is that hippocampal volume provides only a macroscopic measure of tissue integrity, leaving the biological processes underlying age-related atrophy and individual differences in plasticity uncharacterized. In contrast, diffusion-weighted imaging (DWI) offers greater microstructural sensitivity and has been shown to predict memory training benefit and detect training-related changes (de Lange et al. 2016, 2017). However, previous findings were limited to white matter because conventional diffusion tensor metrics lack cellular specificity in grey matter (de Lange et al. 2017). Soma And Neurite Density Imaging (SANDI; Palombo et al. 2020) overcomes this limitation by explicitly modelling soma and neurite compartments, yielding separable fractional estimates of cell body density (f_{soma}) and dendritic and axonal compartments (f_{neurite}). To our knowledge, no previous study has examined SANDI-derived measures as predictors of mnemonic training benefit. This approach enables the assessment of whether hippocampal microstructural properties provide predictive information beyond conventional macrostructural measures, and whether hippocampal reserve is better captured by the soma-related or the neurite-related signal fraction.

The present study addresses these gaps in a sample of healthy older adults who completed a 4-week tablet-based MoL training, with training benefit operationalized as the maximum level attained on the training platform (MaxLevel). We combined complementary macrostructural and microstructural MRI measures to characterize the anatomical and biological substrates of individual differences in training responsiveness, using cross-validation to assess the generalizability of all primary findings. We hypothesized that greater left hippocampal volume before training would predict training benefit beyond established sociodemographic and cognitive predictors, replicating Engvig et al. (2012) in a cognitively healthy sample with a refined segmentation protocol. Building on this, we examined three main questions: whether the volumetric association is preferentially driven by specific hippocampal subfields; whether training benefit is more closely related to anterior or posterior hippocampal structure; and whether the predictive value of hippocampal integrity is better explained by neurite- or soma-related microstructural properties.

Methods

Participants

A total of 34 community-dwelling older adults were recruited. Exclusion criteria included neurological or psychiatric disorders, previous stroke, and MRI contraindications. Informed consent was obtained from

all participants. All participants scored ≥ 26 on the MMSE ($M = 29$, $SD = 1.3$, range = 26–30). Four participants withdrew during training for unrelated reasons, yielding a final sample of 30 older adults (20 women; age range = 65–75 years; $M = 68.67$, $SD = 3.00$). All participants provided written informed consent. The study was approved by the departmental research review board [CReREU, protocol #42024].

General procedure

The study followed a pre-post design comprising two assessment sessions and two scan sessions, one of each conducted before and after a four-week memory-training intervention. Participants first completed a neuropsychological battery after enrollment and underwent the first MRI scan a median of 12 days after the neuropsychological battery (IQR = [2, 29.75]). The extended range reflects scheduling constraints and was not associated with training outcome or baseline cognitive performance. Immediately after the pre-training MRI session, participants received instructions on the MoL and familiarized themselves with the training platform on a tablet. They then completed the home-based MoL training program over the subsequent 28 days, with instructions to complete 30 minutes of training a day. The second MRI scan was conducted 28 days after the first one (median time between scans: 28 days; IQR [28, 28]), followed by a second neuropsychological battery, administered a median of 1 day after the second MRI scan (IQR = [0, 2.75]). The present analyses focus on baseline MRI data and behavioral measures acquired at both timepoints, consistent with the predictive design.

Materials

Neuropsychological battery

The battery assessed associative and episodic memory, working memory, processing speed, and psychological wellbeing. The Object-Position Combination task (described in detail below; Mitchell et al. 2000) assessed associative binding, the process most proximal to the MoL training. The Rey Auditory Verbal Learning Test (RAVLT; Carlesimo et al. 1996) assessed verbal episodic memory. The forward and backward Digit Span, the forward and backward Corsi Block-Tapping Task (Monaco et al. 2013), the Dot Matrix task (Miyake et al. 2001) assessed working- and short-term memory. The Digit Symbol Coding task (Orsini and Pezzuti 2013) and the Pattern Comparison and Letter Comparison (Salthouse and Babcock 1991) tasks assessed processing speed. Psychological well-being was assessed with the Geriatric Depression Scale (GDS, Galeoto et al. 2018), the Geriatric Anxiety Scale (GAS, Picconi et al. 2023), and the 36-Item Short Form Health Survey (SF-36, Apolone and Mosconi 1998), which assessed depressive symptomatology, anxiety, and health-related quality of life, respectively.

Stimuli

Stimuli for both the Method of Loci task and the Object-Position Combination task were drawn from the Bank of Standardized Stimuli (BOSS; Brodeur et al. 2010, 2014). A total of 445 photographs depicting concrete, everyday objects were selected and divided into two non-overlapping subsets, one for each task, matched for nameability, familiarity, and visual complexity to avoid cross-task interference. Images spanned a broad range of semantic categories (e.g., kitchen utensils, tools, musical instruments, furniture, clothing, sports and leisure items).

Method of Loci training

Participants completed a home-based Method of Loci (MoL) training program (Eligence; Maggioli S.p.A. 2024) on a tablet for 30 min/day over 28 consecutive days. The program was specifically developed and customized for the present study, and participants received instructions immediately after the pre-training MRI. Training comprised two phases. First, participants learned a fixed route through a non-immersive virtual-reality city environment comprising 10 spatial locations (loci). They had to correctly recall the order of the loci on three consecutive attempts before proceeding to the training phase. In the subsequent level-based phase, objects were presented sequentially for 5 s, and participants were instructed to form vivid mental images binding each object to its corresponding locus. At retrieval, objects were presented simultaneously and participants indicated their order of presentation. Difficulty increased adaptively across levels by increasing the number of objects to be remembered (Level 1: 3 objects; Level 2: 4 objects; and so on). Participants advanced one level after two consecutive correct trials, moved back one level after three consecutive failures, and otherwise remained at the current level. When more than 10 objects were required, participants restarted the route from the first locus and superimposed additional images onto previously used loci. The primary training outcome was maximum level attained (MaxLevel), defined as the highest difficulty level reached by each participant across the training period. MaxLevel was used as an index of achieved performance under progressively increasing mnemonic demand.

Object-Position Combination task

The Object-Position Combination task was selected as the nearest transfer measure because it isolates object-location binding, the associative process most directly trained by MoL. The task consisted of an encoding phase followed by a recognition phase. During encoding, three objects were sequentially presented at different locations within a 3×3 grid (2 s per item). A cue (2 s) then indicated the upcoming recognition condition. The three conditions were memory object (Object), position memory (Position), and combination memory (Combination). Participants provided yes/no responses for all trials. Each condition comprised 16 trials, with equal numbers of target ($n = 8$) and non-target ($n = 8$) trials, yielding 48 recognition trials in total. This design enabled the separate assessment of object memory, location memory, and object–location binding memory.

MRI acquisition

MRI data were collected before and after training using a 3.0 Tesla Siemens MAGNETOM Prisma XR scanner (Siemens Healthineers, Erlangen, Germany) equipped with a 32-channel head coil. Acquisitions formed part of a comprehensive multimodal protocol that also included resting-state and task-based functional MRI, which will be reported separately.

Structural MRI

High-resolution, whole-brain T1-weighted (T1-w) images were acquired using a three-dimensional MP2RAGE sequence (Marques et al. 2010), which provides improved gray-white matter contrast and reduced sensitivity to B1 field inhomogeneities compared with conventional MPRAGE acquisitions, resulting in more accurate tissue classification and more reliable morphometric analyses and automated segmentation (Droby et al. 2021). Sequence parameters were: TR = 4,190 ms, TE = 3.49 ms, T11/T12 = 700/2,500 ms, flip angles = 4°/5°, 288 sagittal slices, isotropic voxel size = 0.6 mm³, FOV = 269 × 238 mm, acquisition matrix = 224 × 224, GRAPPA acceleration factor = 3; total acquisition time = 4 min 55 s. Following acquisition, all structural images were visually inspected for motion artifacts and gross anatomical abnormalities to ensure data quality prior to processing and analysis.

Diffusion-weighted MRI

Diffusion-weighted images were acquired using a multiband spin-echo EPI sequence (CMRR MB-EPI) with an anterior-to-posterior phase-encoding direction and the following parameters: TR = 3,000 ms, TE = 90 ms, flip angle = 90°, isotropic voxel size = 1.8 mm³, FOV = 200 × 200 mm², acquisition matrix = 110 × 110, 80 contiguous axial slices, multiband acceleration factor = 4, GRAPPA acceleration factor = 2, bandwidth = 2,066 Hz/pixel. The multi-shell protocol comprised 214 diffusion directions distributed across multiple b-value shells (b = 500, 1,000, 2,000, 3,000, 4,000, and 6,000 s/mm²) acquired with anterior-to-posterior phase encoding. Immediately after this acquisition, a short reverse phase-encoded, posterior-to-anterior sequence was collected to acquire b0 volumes for susceptibility-induced distortion correction. This acquisition was designed to meet the requirements of the SANDI model, which requires high b-value shells to resolve the soma compartment (Palombo et al. 2020).

Fluid attenuated acquisition recovery (FLAIR)

A 3D fluid-attenuated inversion recovery (FLAIR) image was acquired using a T2-SPACE sequence with the following parameters: TR = 5,500 ms, TE = 396 ms, TI = 1,700 ms, flip angle = 120°, isotropic voxel size = 0.6 mm³, FOV = 250 × 201 mm², base resolution = 224, echo train length = 237, 2 averages, fat saturation, and CAIPIRINHA parallel acceleration (total factor = 4).

Structural MRI preprocessing and segmentation

Anatomical images underwent preprocessing before FreeSurfer segmentation, including denoising and bias-field correction of the MP2RAGE UN11 image, skull stripping, vascular lesion segmentation and filling, and tissue segmentation (see Supplementary Materials). Lesion-filled, skull-stripped T1-weighted images were then processed with FreeSurfer v8.0.0 (Fischl 2012) using the automated *recon-all* pipeline.

Hippocampal volumes and subfields were segmented using FreeSurfer's Bayesian hippocampal subfield module (Iglesias et al. 2015). All reconstructions were visually inspected for the accuracy of tissue and surface boundaries. Fine-grained labels were aggregated into six major subdivisions: presubiculum, subiculum, fimbria, CA1, CA2/3, and CA4/dentate gyrus (DG). Each hippocampus was additionally divided along its longitudinal axis into head, body, and tail to examine anterior–posterior variation in the structural associations. Total brain volume (TBV) was calculated in FreeSurfer as the sum of cerebral and cerebellar grey and white matter, excluding ventricles, cerebrospinal fluid, and dura, following the definition used by Engvig et al. (2012).

DWI processing

Diffusion-weighted images underwent denoising (dwidenoise) and Gibbs-ringing correction (mrdegibbs; MRtrix3; RRID:SCR_006971), followed by susceptibility- and eddy-current distortion correction (FSL topup and Eddy) and N4 bias-field correction (Andersson et al. 2003; Andersson and Sotiropoulos 2016; Bautista et al. 2021; Cordero-Grande et al. 2019; Kellner et al. 2016; Smith et al. 2004; Tournier et al. 2019; Veraart, Fieremans, et al. 2016; Veraart, Novikov, et al. 2016).

The preprocessed data were fitted using the SANDI model (Palombo et al. 2020) implemented in the SANDI-MATLAB toolbox, using all six b-value shells ($b = 500\text{--}6,000\text{ s/mm}^2$). Analyses focused on the neurite signal fraction (f_{neurite}) and soma signal fraction (f_{soma}), reflecting the relative contributions of neurite and soma compartments, respectively. These measures were treated as indirect markers of tissue microstructure rather than direct histological measures.

Hippocampal ROIs derived from FreeSurfer segmentation were transformed to native diffusion space by rigidly registering the diffusion b0 image to the denoised MP2RAGE UNI image and applying the inverse transformation to the hippocampal labels using nearest-neighbor interpolation. ROIs were visually inspected before parameter extraction. Median SANDI values were extracted within each ROI to reduce sensitivity to boundary voxels. Given the lower spatial resolution of DWI relative to the anatomical images, microstructural analyses were restricted to anterior (head) and posterior (body and tail) hippocampal subdivisions, matching the longitudinal-axis segmentation used for volumetric analyses (see Supplemental Materials for the complete technical description).

Statistical analysis

All statistical analyses were conducted in R version 4.6.0 (R Core Team 2026). Follow-up subregional analyses were considered exploratory and were not corrected for multiple comparisons.

Covariate Rationale and cognitive predictors.

Total brain volume (TBV) was included as a covariate in volumetric models to account for individual differences in head size. Education was included in all neural models as a theory-driven proxy for cognitive reserve. SANDI signal fractions were not adjusted for TBV because they are normalized microstructural measures. The Geriatric Depression Scale (GDS) was included as an a priori covariate in

the cognitive predictor analysis. Associations between sociodemographic, cognitive, and affective measures and MaxLevel were examined using Spearman correlations. Variables significantly associated with MaxLevel were entered into a multiple regression model together with GDS. Predictors identified in this model were subsequently included as covariates in the neural analyses.

Associations between sociodemographic, cognitive, and affective measures and MaxLevel were examined using Spearman correlations. Variables significantly associated with MaxLevel were entered into a multiple regression model together with GDS. Predictors identified in this model were subsequently included as covariates in the neural analyses.

Object-Position Combination task analyses.

Sensitivity indices (d') were calculated separately for each condition and time point using the log-linear correction for extreme hit and false-alarm rates (Hautus 1995). Repeated-measures ANOVAs tested the effects of condition and time on d' and log-transformed reaction times. A global d' index was calculated by averaging across the three conditions at each time point to provide a more reliable measure of overall associative memory performance. Spearman correlations were used to examine associations between MaxLevel and binding performance at each time point, and between left hippocampal volume and binding performance while controlling for TBV. See Object-Position Combination task sensitivity index correction section in the Supplementary Materials for a detailed description of the sensitivity index correction's formula.

Volumetric and microstructural models.

Multiple linear regression models tested whether baseline hippocampal volume predicted MaxLevel, with education and TBV included as covariates. Models were estimated separately for the left and right whole hippocampus. Significant whole-hippocampal models were followed by exploratory analyses of hippocampal subfields and longitudinal-axis subdivisions. For SANDI measures, separate regression models tested whether f_{soma} or f_{neurite} predicted MaxLevel, with education as a covariate, separately for each hemisphere. Significant whole-hippocampal models were followed by analyses of anterior (head) and posterior (body and tail) subdivisions.

Cross validation and direct comparison of associations

The predictive performance of significant neural models was evaluated using leave-one-out cross-validation (LOOCV). Model performance was quantified using cross-validated R^2 , RMSE, MAE, and the Pearson correlation between observed and predicted MaxLevel. Statistical significance of predictive performance was assessed using 5,000 permutations (see Supplementary Materials for details).

Selected associations were formally compared using tests for dependent partial correlations sharing a common variable (Williams 1959). Differences were expressed as Δr with 95% confidence intervals obtained from 5,000 bootstrap resamples and were additionally evaluated using 5,000 permutations. Two pre-specified confirmatory contrasts were tested: left versus right whole-hippocampal volume and

f_{soma} versus f_{neurite} , with family-wise error controlled using the Holm procedure. All other subregional analyses were considered exploratory.

Results

Participants completed a mean of 26.70 training days ($SD = 2.91$). MaxLevel ranged from 8 to 29 ($M = 17.56$, $SD = 6.39$), reflecting the substantial individual variability in training responsiveness. Figure 1 shows individual training trajectories. Sociodemographic characteristics and baseline neuropsychological measures are reported in Supplementary Table S1.

Cognitive and affective predictors

Spearman correlations showed that only years of education ($\rho = .61$, $p < .001$) and Dot matrix accuracy ($\rho = .39$, $p = .03$) were significantly associated with MaxLevel (Supplementary Table S2). When both predictors and GDS (retained as an a priori relevant covariate) were entered into a multiple linear regression entering both predictors together with GDS, the model explained a 29% of the variation in MaxLevel [adjusted $R^2 = .21$; $F(3, 26) = 3.49$, $p = .03$]. Education was the only significant independent predictor [$\beta = .49$, $t(26) = 2.83$, $p = .009$]; neither Dot matrix accuracy nor GDS approached significance (all $ps > .40$). Education was therefore included as a covariate in subsequent neural models.

Object-Position Combination Task Results

The Object-Position Combination task was used as the nearest transfer measure. A 3 (Condition: Object, Location, Combination) $\times$ 2 (Time: pre, post) repeated-measures ANOVA on d' scores revealed significant main effects of Condition [$F(2, 58) = 10.11$, $p < .001$, $\eta^2 = .26$] and Time [$F(1, 29) = 32.82$, $p < .001$, $\eta^2 = .53$], with performance improving from pre to post training (see Table 1). The Condition $\times$ Time interaction was not significant [$F(2, 58) = 1.11$, $p = .33$, $\eta^2 = .007$], indicating a comparable improvement across conditions.

Table 1
Descriptive statistics for the Object-Position Combination task by condition and time point. Values are M (SD).

		Hit rate	FA	d'
Location	Pre	.92(.10)	.15(.05)	2.15(.47)
	Post	.95(.10)	.02(.04)	2.87(.44)
Object	Pre	.62(.38)	.01(.04)	1.86(1.15)
	Post	.80(.31)	.008(.03)	2.44(1.00)
Combination	Pre	.95(.08)	.42(.27)	1.61(.84)
	Post	.96(.10)	.29(.25)	2.02(.86)
FA = false alarm rate; values are based on log-linear correction (Hautus 1995).				

We next examined whether individual differences in training success were associated with binding performance. MaxLevel was not associated with pre-training binding performance ($\rho = .22, p = .242$), whereas it was significantly correlated with post-training performance ($\rho = .42, p = .020$). This dissociation is consistent with transfer of training-related gains to an independent measure of binding.

Macrostructural predictors

Two multiple linear regression models examined whether baseline hippocampal volumes predicted MaxLevel, controlling for total brain volume (TBV) and education (Table 2, Fig. 2).

The left hippocampus model explained 52% of the variance in MaxLevel [adjusted $R^2 = .46$; $F(3, 26) = 9.27, p < .001$], with left hippocampal volume emerging as a significant positive predictor [$\beta = .66, t(26) = 3.31, p = .003$]. Neither TBV nor education reached significance in this model (both $ps > .06$). The right hippocampal model also explained a 34% of the variance adjusted $R^2 = .26$; $F(3, 26) = 4.38, p = .013$, but education was the only significant predictor of training performance [$\beta = .48, t(26) = 2.90, p = .008$]. Neither right hippocampal volume nor TBV (both $ps > .36$) significantly predicted MaxLevel.

The association with MaxLevel was significantly stronger for left than right hippocampal volume. Comparing the dependent partial correlations, controlling for education and TBV, yielded $\Delta r = .37, 95\% \text{ CI } [.13, .59]$, Williams's $t(25) = 3.79, p < .001$; the difference was confirmed by permutation testing ($p = .001$) and remained significant after Holm correction across the two confirmatory contrasts ($p = .002$). Restricting subsequent analyses to the left hemisphere therefore reflects a difference in association strength rather than statistical significance alone.

To examine whether left hippocampal volume was related to baseline associative memory ability, we correlated it with pre- and post-training binding performance on the Object-Position Combination task, controlling for TBV. Left hippocampal volume was not associated with pre-training binding performance

($p = .23$, $p = .241$), whereas the association with post-training performance was numerically stronger but remained non-significant ($p = .34$, $p = .070$).

Exploratory follow-up models examined left hippocampal anatomical subfields (presubiculum, subiculum, fimbria, CA1, CA2/3, CA4/DG), each controlling for education and TBV (Table 2; Figure S1 in the Supplementary Materials), to determine whether the effect was driven by specific subregions or distributed across the hippocampal formation. All exploratory models were significant overall [unadjusted R^2 s = .37–.43, adjusted R^2 s = .29–.37; $F(3, 26) \geq 4.97$, $ps \leq .007$], with education emerging as a significant predictor of MaxLevel in each model [β s = .35–.49, $t(26) > 2.17$, $ps \leq .04$]. At the subfield level, larger CA2/3 volume ($\beta = .45$, $p = .03$) and CA4/DG volume ($\beta = .44$, $p = .04$) were positively associated with higher training performance. In contrast, the models for CA1, subiculum, presubiculum and fimbria did not reveal significant associations with MaxLevel [β s = $-.23$ to $.43$, $ps \geq .06$].

Exploratory models for left hippocampal head, body, and tail volumes revealed significant associations for the tail and body but not the head (Table 2; Figure S2 in the Supplementary Materials). All models were significant overall [unadjusted R^2 s = .40–.52, adjusted R^2 s = .33–.47; $F(3, 26) \geq 5.76$, $ps \leq .004$]. Education was a significant predictor in each model [β s = .36–.41, $t(26) > 2.27$, $ps \leq .031$], whereas TBV was not significant in any model [all $ps \geq .45$]. The hippocampal tail [$\beta = .49$, $t(26) = 3.34$, $p = .003$] and body [$\beta = .44$, $t(26) = 2.31$, $p = .03$] were significant predictors, whereas the head was not [$\beta = .42$, $t(26) = 1.92$, $p = .07$] (see Supplementary Table S3 for full output models). This pattern of results, significant in posterior subregions but not in the head would be consistent with the fine-grained spatial and sequential processing demands of the MoL, as outlined in the introduction.

Table 2
Multiple regression models predicting MaxLevel from hippocampal volume, adjusted for TBV and education.

Model	Adjusted R^2	Volume		Education	
		β	p	β	p
Whole hippocampus					
Left hemisphere	.46	.66	.003	.30	.06
Right hemisphere	.26	.26	.36	.48	.008**
Anatomical subfields (left)					
Subiculum	.31	.30	.11	.45	.008**
Presubiculum	.29	.26	.16	.46	.007**
Fimbria	.29	– .23	.16	.49	.004**
CA1	.33	.43	.06	.38	.03*
CA2/3	.37	.45	.03	.35	.04*
CA4-DG	.35	.44	.04	.38	.03*
Antero-posterior axis (left)					
Head	.33	.42	.07	.41	.02*
Body	.37	.44	.03	.37	.03*
Tail	.47	.49	.003	.36	.02*
All models include TBV, education, and the corresponding hippocampal volume. β = standardized regression coefficient; adjusted R^2 = adjusted R^2 of the full model. * $p < .05$; ** $p < .01$.					

SANDI microstructural predictors

We next examined whether hippocampal microstructural properties assessed using SANDI provided convergent predictive information, and whether associations with MaxLevel differed between soma- and neurite-related tissue compartments. Results are summarized in Table 3, with full model output provided in Supplementary Table S4.

f_{soma} results

Neither left nor right hippocampal f_{soma} predicted MaxLevel. For the left hippocampus, the model explained 33% of the variance [adjusted $R^2 = .28$; $F(2, 27) = 6.56$, $p = .005$]. However, only education [$\beta = .59$, $t(27) = 3.57$, $p = .001$] was a significant predictor whereas left hippocampal f_{soma} was not [$\beta = -.25$, $t(27) = -1.53$, $p = .14$]. Results were similar for the right hippocampus. The model explained 29% of the

total variance of the training performance [unadjusted $R^2 = .29$, adjusted $R^2 = .24$; $F(2, 27) = 5.45$, $p = .01$], with education again the sole significant predictor [$\beta = .54$, $t(27) = 3.28$, $p = .003$] and right f_{soma} non-significant [$\beta = -.14$, $t(27) = -0.86$, $p = .40$].

f_{neurite} results

In contrast, left hippocampal f_{neurite} was a significant predictor of MaxLevel. The model explained 41% of the variance [adjusted $R^2 = .37$; $F(2, 27) = 9.49$, $p < .001$], with both education ($\beta = .50$, $t(27) = 3.36$, $p = .002$) and left f_{neurite} ($\beta = .38$, $t(27) = 2.58$, $p = .02$) as significant predictors (Fig. 3). For the right hippocampus, the model explained 27% of the variance (unadjusted $R^2 = .27$, adjusted $R^2 = .22$; $F(2, 27) = 5.11$, $p = .01$), but only education was a significant predictor ($\beta = .51$, $t(27) = 3.08$, $p = .005$); right f_{neurite} did not predict MaxLevel ($\beta = .07$, $t(27) = 0.48$, $p = .63$). This left-lateralized pattern mirrors the volumetric findings.

For the left hippocampus, f_{neurite} and f_{soma} related to MaxLevel in opposite directions. In particular, f_{neurite} has a positive partial correlation ($r = .44$) and a negative partial correlation for f_{soma} ($r = -.28$). The two associations differed reliably, $\Delta r = .73$, 95% CI [.19, 1.20], Williams's $t(26) = 2.39$, $p = .025$, with agreement from the permutation test ($p = .022$) and after Holm correction ($p = .025$). The predictive signal is carried by the neurite-related fraction and does not extend to the soma-related fraction.

Given the left-hemisphere association with f_{neurite} , follow-up models examined anterior and posterior subdivisions, controlling for education (Figure S3). Anterior f_{neurite} explained 35% of the variance in the training performance [adjusted $R^2 = .30$; $F(2, 27) = 7.12$, $p = .003$], but only education [$\beta = .50$, $t(27) = 3.25$, $p = .003$] was a significant predictor of MaxLevel, whereas anterior f_{neurite} was not [$\beta = .28$, $t(27) = 1.78$, $p = .08$]. This is consistent with the volumetric pattern, where posterior subregions were significant and the head was not.

In contrast, posterior f_{neurite} was a significant predictor of MaxLevel. The model explained 39% of the variance in MaxLevel [adjusted $R^2 = .34$; $F(2, 27) = 8.44$, $p = .001$]. Both education ($\beta = .49$, $t(27) = 3.24$, $p = .003$) and posterior f_{neurite} ($\beta = .34$, $t(27) = 2.26$, $p = .03$) appeared as significant predictors of MaxLevel. Taken together, the macrostructural and microstructural analyses showed the same left-hemisphere pattern: left HC volume and left HC f_{neurite} each predicted MaxLevel, whereas right-hemisphere and f_{soma} were not related to training performance.

Table 3

Multiple regression models predicting MaxLevel from SANDI measures, adjusted for education.

Model	Adj- R^2	SANDI metric		Education	
		β	p	β	p
f_{soma}					
Left hemisphere	.23	-.25	.14	.59	.001**
Right hemisphere	.24	-.14	.40	.54	.003**
$f_{neurite}$					
Left hemisphere	.37	.38	.02	.50	.002**
Right hemisphere	.22	.08	.63	.51	.005**
Antero-posterior axis, $f_{neurite}$ (left)					
Anterior	.30	.28	.09	.51	.003**
Posterior	.34	.34	.03	.49	.003**
All models include education and the corresponding SANDI metric. β = standardized regression coefficient; adjusted R^2 = adjusted R^2 of the full model. * $p < .05$; ** $p < .01$.					

LOOCV results

Given the limited sample size, significant hippocampal models were evaluated using LOOCV to assess internal generalizability (Table 4; Fig. 4). Cross-validation was restricted to models in which the neural predictor was significant in the primary regression, consistent with the hierarchical follow-up logic.

The left hippocampal volume model (TBV + Education + Left HC) explained 35% of variance in held-out participants ($R^2CV = .35$, $r = .61$, $p < .001$), compared with 46% in the full sample. Among the subfield models, CA2/3 ($R^2CV = .25$, $r = .53$, $p = .002$), and CA4-DG ($R^2CV = .21$, $r = .50$, $p = .002$) showed significant cross-validation performance. Along the longitudinal axis, the hippocampal tail ($R^2CV = .36$, $r = .62$, $p < .001$) and the body ($R^2CV = .25$, $r = .53$, $p = .002$) showed significant generalization effects. LOOCV plots for anatomical subregions and longitudinal-axis subdivisions are provided in Supplementary Figures S4, S5. For the SANDI microstructural models, left hippocampal $f_{neurite}$ generalized significantly to held-out participants ($R^2CV = .23$, $r = .52$, $p = .002$), with the posterior subdivision also showing a significant generalization ($R^2CV = .21$, $r = .46$, $p = .003$, Figure S6).

Permutation-based p -values remained significant across models, and the cross-validated pattern was consistent with the primary analyses, supporting the robustness of the left-hemisphere and posterior effects and the association with $f_{neurite}$ rather than f_{soma} .

Table 4
Leave-one-out cross-validation (LOOCV) of left hippocampal models predicting MaxLevel.

Model	<i>N</i>	R^2_{cv}	RMSE _{cv}	MAE _{cv}	<i>r</i>	<i>p</i>
Left HC	30	.35	5.07	4.35	.61	< .001***
Anatomical subregions						
CA2/3	30	.25	5.44	4.45	.53	.002***
CA4-DG	30	.21	5.59	4.56	.50	.002***
Antero-posterior axis						
Body	30	.25	5.46	4.65	.53	.002***
Tail	30	.36	5.01	4.18	.62	< .001***
SANDI microstructural models						
$f_{neurite}$ Left HC	30	.23	5.50	4.34	.52	.002***
$f_{neurite}$ (posterior)	30	.21	5.60	4.61	.50	.003***
R ² CV = cross-validated R ² ; RMSECV = root mean squared error; MAECV = mean absolute error; <i>r</i> = Pearson correlation between observed and LOOCV-predicted MaxLevel. <i>p</i> = one-tailed permutation <i>p</i> -value (5,000 permutations, continuity-corrected). No correction for multiple comparisons was applied to exploratory subregion and SANDI analyses. * <i>p</i> < .05; ** <i>p</i> < .01; *** <i>p</i> < .001.						

Discussion

The present study shows that hippocampal integrity is associated with individual differences in responsiveness to MoL training in healthy older adults. Left hippocampal volume predicted MaxLevel after controlling for education and TBV, replicating previous evidence that pre-training hippocampal structure is associated with memory-training gains (Engvig et al. 2012; Bråthen et al. 2018). Importantly, our direct comparison of the two hemispheres showed that the association was significantly stronger in the left hippocampus. This extends previous findings by demonstrating that the relationship is not simply attributable to a general association between hippocampal size and training performance, it is preferentially expressed in the left hemisphere.

The findings also refine previous subfield-level evidence. Engvig et al. (2012) reported associations between training gains and left hippocampal, CA2/3, and CA4/dentate gyrus volumes, although their analyses relied on an earlier segmentation protocol. Here, the association with training responsiveness was again observed in CA2/3 and CA4/dentate gyrus using contemporary segmentation methods. The regional specificity of the volumetric effect is consistent with hippocampal circuitry supporting pattern separation and detailed spatial associations (Hainmueller and Bartos 2020; Liu et al. 2022), processes central to MoL performance, which requires discriminating similar item–location associations, retrieving

them from partial spatial cues, and preserving their sequence (Rolls 2018). Together with the whole-hippocampus finding, this convergence suggests that the structural features associated with MoL responsiveness extend across hippocampal levels of organization, from the entire formation to specific subregions.

Beyond subfield specificity, the present findings provide evidence that hippocampal contributions to training responsiveness vary along the longitudinal axis. Significant associations were observed in the hippocampal body and tail, with the volumetric effect concentrated in posterior regions. This pattern is consistent with the functional organization of the hippocampus, in which anterior and posterior regions differ in connectivity and representational granularity. Whereas anterior hippocampal regions have been associated with broader, more gist-like representations, posterior regions are more strongly involved in fine-grained spatial and contextual information (Schlichting et al. 2015; Evensmoen et al. 2013; Grady 2020; Poppenk et al. 2013; Strange et al. 2014; Zeidman and Maguire 2016).

This organization is particularly relevant to MoL training, which requires participants to construct and repeatedly retrieve a detailed mental route while binding individual items to specific locations and maintaining their ordered sequence. The preferential involvement of posterior hippocampal regions may therefore reflect the representational demands of the strategy itself. The convergence between the posterior pattern observed across longitudinal segments and the involvement of CA2/3 and CA4/dentate gyrus provides a more anatomically specific characterization of the hippocampal substrate of training responsiveness than whole-volume measures alone. Nevertheless, these subfield and longitudinal-axis analyses were exploratory and uncorrected for multiple comparisons. They should therefore be interpreted as convergent anatomical evidence.

The SANDI analyses provide a complementary tissue-level perspective. Only left hippocampal f_{neurite} was associated with MaxLevel, and this effect again showed a posterior predominance. To our knowledge, this is the first study to examine SANDI-derived hippocampal measures as predictors of mnemonic training benefit. The finding extends previous work showing that diffusion-based measures can capture individual differences in memory training and training-related change (de Lange et al. 2016, 2017), by applying a model designed to characterize microstructural properties within hippocampal grey matter. The stronger association with f_{neurite} than f_{soma} may suggest that variation in neurite-related tissue properties is more closely related to training responsiveness than variation in the soma-related compartment. However, SANDI-derived f_{neurite} is not specific to a single cellular process and may reflect contributions from both dendritic and axonal compartments. Moreover, because the measures were acquired before training, the present results cannot determine whether they represent biological mechanisms supporting training-induced plasticity or pre-existing neural reserve that enables some individuals to benefit more from training. Thus, the findings identify a microstructural correlate of responsiveness rather than a specific cellular mechanism of cognitive plasticity.

The convergence between macrostructural and microstructural findings is nevertheless informative. Both measures implicated the left hippocampus, both showed a posterior predominance, and both

generalized to held-out participants under leave-one-out cross-validation. Although external validation in larger independent samples remains necessary, this convergence suggests that individual differences in training responsiveness are reflected at multiple levels of hippocampal organization.

An important question is whether the neural predictors identified here relate specifically to training responsiveness or instead reflect baseline memory ability. The absence of a clear association between pre-training binding performance and hippocampal volume, together with the robust association between hippocampal integrity and MaxLevel, supports the interpretation that hippocampal measures capture a component of trainability rather than simply initial associative-memory performance. This distinction is important because cognitive training studies often focus on mean improvement while overlooking substantial individual variability in response. In fact, the improvement observed in the untrained Object-Position Combination task provides further evidence for near transfer to an associative-memory measure. Such transfer is notable because mnemonic training effects often remain relatively close to the trained paradigm, with near transfer generally easier to demonstrate than far transfer (Schubert et al. 2014; Sala et al. 2019). The present findings therefore suggest that greater responsiveness to MoL training may be accompanied by improvements in a related but untrained memory process. Whether the hippocampal predictors identified here also explain transfer to more distant cognitive domains remains unknown.

More broadly, these findings contribute to the study of individual differences in cognitive ageing by identifying a measurable neurobiological dimension associated with responsiveness to intervention. The results may ultimately inform approaches aimed at understanding why some older adults derive greater benefit from cognitive training than others and could provide a basis for future stratification or personalization strategies. However, these implications remain conceptual. The present sample was modest, the design was observational with respect to the neural predictors, and the predictive models have not yet been externally validated. Larger independent samples will therefore be needed to establish the robustness and generalizability of these associations before hippocampal measures can support individual-level prediction or personalized intervention decisions.

Overall, the findings indicate that left hippocampal integrity, particularly in posterior regions, is associated with individual differences in MoL training responsiveness. The convergence of volumetric and neurite-related microstructural measures suggests that this relationship is expressed across multiple levels of hippocampal organization. By linking these neural characteristics to variation in training benefit rather than baseline memory performance alone, the study provides a more specific account of the neurobiological basis of cognitive trainability in healthy ageing.

Declarations

CRedit authorship contribution statement

R.L.: Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Writing – original draft. R.P.: Conceptualization, Data curation, Formal analysis, Investigation, Writing – original draft. D.D.C.: Conceptualization, Investigation, Data curation, Formal analysis, Visualization, Writing – original draft. G.M.: Data curation, Formal analysis. M.G.P.: Funding acquisition, Project administration, Resources, Writing – review & editing. B.F.: Conceptualization, Writing – review & editing. S.D.P.: Conceptualization, Writing – review & editing. G.C.: Conceptualization, Funding acquisition, Writing – review & editing. M.C.: Supervision, Writing – review & editing. R.G.W.: Supervision, Writing – review & editing. C.S.: Conceptualization, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Funding

This work and the development of the Method of Loci training platform were supported by a grant from the European Union-NextGenerationEU–Italian Ministry of University and Research (MUR), National Plan for Recovery and Resilience (PNRR) and Projects of National Relevance (PRIN), Project Title: “Training Memory for Successful Aging: social robots, neural predictors and mechanisms of plasticity (REFUEL)”, Project Code: P2022CWYTJ, CUP: D53D23020660001. Founding call No.1409 of 14.09.2022. PI: Carlo Sestieri. Mission 4, Component 2, Investment 1.1, and was conducted under the framework of the Departments of Excellence 2023-2027 initiative of the Italian Ministry of Education, University and Research for the Department of Neuroscience, Imaging and Clinical Sciences (DNISC) of the University of Chieti-Pescara.

Declaration of competing interests

R.P. is a doctoral candidate whose PhD fellowship is co-funded by the University "G. d'Annunzio" of Chieti–Pescara and by Maggioli S.p.A. The Method of Loci training platform used in this study was developed by Maggioli S.p.A. under a paid service agreement, at the request of and to specifications set by the research team, specifically for this study; it is not a publicly available commercial product. The other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data and scripts that support the findings of this study are available on the project repository on Github (<https://github.com/RaquelLezCan/Hippocampus-predicts-training-gains/tree/main>).

Declaration of generative AI use

During the preparation of this work, authors used Claude (Anthropic 2026) to improve the spelling and grammar of the text and to depurate statistical analyses scripts. After using this tool, the first author reviewed and edited the content as needed and takes full responsibility for the content of the present article.

References

1. Amaral DG, Witter MP (1989) The three-dimensional organization of the hippocampal formation: a review of anatomical data. *Neuroscience* 31:571–591. [https://doi.org/10.1016/0306-4522\(89\)90424-7](https://doi.org/10.1016/0306-4522(89)90424-7)
2. Andersson JLR, Sotiropoulos SN (2016) An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage* 125:1063–1078. <https://doi.org/10.1016/j.neuroimage.2015.10.019>
3. Andersson JLR, Skare S, Ashburner J (2003) How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage* 20:870–888. [https://doi.org/10.1016/S1053-8119\(03\)00336-7](https://doi.org/10.1016/S1053-8119(03)00336-7)
4. Anthropic (2026) Claude (Opus 4.8) [Large language model]. <https://claude.ai>
5. Apolone G, Mosconi P (1998) The Italian SF-36 Health Survey: translation, validation and norming. *J Clin Epidemiol* 51:1025–1036. [https://doi.org/10.1016/S0895-4356\(98\)00094-8](https://doi.org/10.1016/S0895-4356(98)00094-8)
6. Battaglini M, Jenkinson M, De Stefano N (2012) Evaluating and reducing the impact of white matter lesions on brain volume measurements. *Hum Brain Mapp* 33:2061–2071. <https://doi.org/10.1002/hbm.21344>
7. Bautista T, O'Muircheartaigh J, Hajnal JV, Tournier JD (2021) Removal of Gibbs ringing artefacts for 3D acquisitions using subvoxel shifts. *Proc Int Soc Magn Reson Med* 29:1482
8. Baykara E, Kühn S, Kessler J, Rapp MA, Heinzel S (2021) MRI predictors of cognitive training outcomes. *J Cogn Enhanc* 5:180–191. <https://doi.org/10.1007/s41465-020-00198-w>
9. Bosch M, Hayashi Y (2012) Structural plasticity of dendritic spines. *Curr Opin Neurobiol* 22:383–388. <https://doi.org/10.1016/j.conb.2011.09.002>
10. Bourne JN, Harris KM (2008) Balancing structure and function at hippocampal dendritic spines. *Annu Rev Neurosci* 31:47–67. <https://doi.org/10.1146/annurev.neuro.31.060407.125646>
11. Brodeur MB, Dionne-Dostie E, Montreuil T, Lepage M (2010) The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS ONE* 5:e10773. <https://doi.org/10.1371/journal.pone.0010773>
12. Brodeur MB, Guérard K, Bouras M (2014) Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. *PLoS ONE* 9:e106953. <https://doi.org/10.1371/journal.pone.0106953>
13. Bråthen ACS, Sneve MH, Amlien IK, Grydeland H, Fjell AM, Walhovd KB (2018) Multimodal cortical and hippocampal prediction of episodic-memory plasticity in young and older adults. *Hum Brain Mapp* 39:4329–4342. <https://doi.org/10.1002/hbm.24287>

14. Carlesimo GA, Caltagirone C, Gainotti G, Fadda L, Gallassi R, Lorusso S, Marfia GA (1996) The Mental Deterioration Battery: normative data, diagnostic reliability and qualitative analyses of cognitive impairment. *Eur Neurol* 36:378–384. <https://doi.org/10.1159/000117297>
15. Cordero-Grande L, Christiaens D, Hutter J, Price AN, Hajnal JV (2019) Complex diffusion-weighted image estimation via matrix recovery under general noise models. *NeuroImage* 200:391–404. <https://doi.org/10.1016/j.neuroimage.2019.06.039>
16. Daugherty AM, Carr V, Canada KL, Ekstrom AD, Kanel D, La Joie R, Berron D, Cecchetti G, De Flores R, Ding SL, Insausti R, Kasa LW, Malykhin NV, Olsen RK, Palombo M, Postel C, Biegler KA, Yushkevich PA (2026) Harmonized protocol for subfield segmentation in the hippocampal body on high-resolution in vivo MRI from the Hippocampal Subfields Group (HSG). *Hippocampus* 36:e70073. <https://doi.org/10.1002/hipo.70073>
17. de Lange AMG, Bråthen ACS, Grydeland H, Sexton C, Johansen-Berg H, Andersson JLR, Fjell AM, Walhovd KB (2016) White matter integrity as a marker for cognitive plasticity in aging. *Neurobiol Aging* 47:74–82. <https://doi.org/10.1016/j.neurobiolaging.2016.07.007>
18. de Lange AMG, Bråthen ACS, Rohani DA, Grydeland H, Fjell AM, Walhovd KB (2017) The effects of memory training on behavioral and microstructural plasticity in young and older adults. *Hum Brain Mapp* 38:5666–5680. <https://doi.org/10.1002/hbm.23756>
19. Dimsdale-Zucker HR, Ritchey M, Ekstrom AD, Yonelinas AP, Ranganath C (2018) CA1 and CA3 differentially support spontaneous retrieval of episodic contexts within human hippocampal subfields. *Nat Commun* 9:294. <https://doi.org/10.1038/s41467-017-02752-1>
20. Ding SL, Van Hoesen GW (2015) Organization and detailed parcellation of human hippocampal head and body regions based on a combined analysis of cyto- and chemoarchitecture. *J Comp Neurol* 523:2233–2253. <https://doi.org/10.1002/cne.23786>
21. Droby A, Thaler A, Giladi N, Hutchison RM, Mirelman A, Ben Bashat D, Karni A (2021) Whole brain and deep gray matter structure segmentation: quantitative comparison between MPRAGE and MP2RAGE sequences. *PLoS ONE* 16:e0254597. <https://doi.org/10.1371/journal.pone.0254597>
22. Eldridge LL, Engel SA, Zeineh MM, Bookheimer SY, Knowlton BJ (2005) A dissociation of encoding and retrieval processes in the human hippocampus. *J Neurosci* 25:3280–3286. <https://doi.org/10.1523/JNEUROSCI.3420-04.2005>
23. Elkana O, Beheshti I, Alzheimer's Disease Neuroimaging Initiative (2025) The protective role of education in white matter lesions and cognitive decline. *Brain Cogn* 184:106304. <https://doi.org/10.1016/j.bandc.2025.106304>
24. Engvig A, Fjell AM, Westlye LT, Skaane NV, Sundseth Ø, Walhovd KB (2012) Hippocampal subfield volumes correlate with memory training benefit in subjective memory impairment. *NeuroImage* 61:188–194. <https://doi.org/10.1016/j.neuroimage.2012.02.072>
25. Evensmoen HR, Lehn H, Xu J, Witter MP, Nadel L, Håberg AK (2013) The anterior hippocampus supports a coarse, global environmental representation and the posterior hippocampus supports

- fine-grained, local environmental representations. *J Cogn Neurosci* 25:1908–1925.
https://doi.org/10.1162/jocn_a_00436
26. Fischl B (2012) *FreeSurfer Neuroimage* 62:774–781.
<https://doi.org/10.1016/j.neuroimage.2012.01.021>
27. Galeoto G, Sansoni J, Scuccimarri M, Bruni V, De Santis R, Colucci M, Tofani M (2018) A psychometric properties evaluation of the Italian version of the Geriatric Depression Scale. *Depress Res Treat* 2018:1797536. <https://doi.org/10.1155/2018/1797536>
28. Gallen CL, D'Esposito M (2019) Brain modularity: a biomarker of intervention-related plasticity. *Trends Cogn Sci* 23:293–304. <https://doi.org/10.1016/j.tics.2019.01.014>
29. Gazes Y, Li P, Sun E, Razlighi Q, Tsapanou A (2019) Age specificity in fornix-to-hippocampus association. *Brain Imaging Behav* 13:1133–1142. <https://doi.org/10.1007/s11682-018-9958-1>
30. Gelman A, Stern H (2006) The difference between significant and not significant is not itself statistically significant. *Am Stat* 60:328–331. <https://doi.org/10.1198/000313006X152649>
31. Gertje EC, Pluta J, Das S, Mancuso L, Kliot D, Yushkevich P, Wolk DA (2016) Clinical application of automatic segmentation of medial temporal lobe subregions in prodromal and dementia-level Alzheimer's disease. *J Alzheimers Dis* 53:1421–1432. <https://doi.org/10.3233/JAD-160014>
32. Giuliano A, Donatelli G, Cosottini M, Tosetti M, Retico A, Fantacci ME (2017) Hippocampal subfields at ultra high field MRI: an overview of segmentation and measurement methods. *Hippocampus* 27:929–948. <https://doi.org/10.1002/hipo.22717>
33. Grady CL (2020) Meta-analytic and functional connectivity evidence from functional magnetic resonance imaging for an anterior to posterior gradient of function along the hippocampal axis. *Hippocampus* 30:456–471. <https://doi.org/10.1002/hipo.23164>
34. Gross AL, Parisi JM, Spira AP, Kueider AM, Ko JY, Saczynski JS, Rebok GW (2012) Memory training interventions for older adults: a meta-analysis. *Aging Ment Health* 16:722–734.
<https://doi.org/10.1080/13607863.2012.667783>
35. Guo H, Ali T, Que J, Zhou Y, Bai Y (2023) Dendritic spine dynamics in associative memory: a comprehensive review. *FASEB J* 37:e22812. <https://doi.org/10.1096/fj.202202166R>
36. Guye S, De Simoni C, von Bastian CC (2017) Do individual differences predict change in cognitive training performance? A latent growth curve modeling approach. *J Cogn Enhanc* 1:412–431.
<https://doi.org/10.1007/s41465-017-0049-9>
37. Hainmueller T, Bartos M (2020) Dentate gyrus circuits for encoding, retrieval and discrimination of episodic memories. *Nat Rev Neurosci* 21:153–168. <https://doi.org/10.1038/s41583-019-0260-z>
38. Hautus MJ (1995) Corrections for extreme proportions and their biasing effects on estimated values of d'. *Behav Res Methods Instrum Comput* 27:46–51. <https://doi.org/10.3758/BF03203619>
39. Iglesias JE, Augustinack JC, Nguyen K, Player CM, Player A, Wright M, Fischl B (2015) A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: application to adaptive segmentation of in vivo MRI. *NeuroImage* 115:117–137.
<https://doi.org/10.1016/j.neuroimage.2015.04.042>

40. Isensee F, Schell M, Tursunova I, Brugnara G, Bonekamp D, Neuberger U, Maier-Hein K (2019) Automated brain extraction of multi-sequence MRI using artificial neural networks. *Hum Brain Mapp* 40:4952–4964. <https://doi.org/10.1002/hbm.24750>
41. Jaeggi SM, Buschkuhl M, Shah P, Jonides J (2014) The role of individual differences in cognitive training and transfer. *Mem Cognit* 42:464–480. <https://doi.org/10.3758/s13421-013-0364-z>
42. Jung CKE, Herms J (2014) Structural dynamics of dendritic spines are influenced by an environmental enrichment: an in vivo imaging study. *Cereb Cortex* 24:377–384. <https://doi.org/10.1093/cercor/bhs317>
43. Kellner E, Dhital B, Kiselev VG, Reiser M (2016) Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magn Reson Med* 76:1574–1581. <https://doi.org/10.1002/mrm.26054>
44. Knierim JJ (2015) The hippocampus. *Curr Biol* 25:R1183–R1187. <https://doi.org/10.1016/j.cub.2015.10.049>
45. Liu J, Zhang H, Yu T, Ren L, Ni D, Yang Q, Wang L (2022) Hippocampal representations of event structure and temporal context during episodic temporal order memory. *Cereb Cortex* 32:2588–2601. <https://doi.org/10.1093/cercor/bhab304>
46. Maggioli SA (2024) Software di valutazione neuropsicologica. Maggioli Editore
47. Magni E, Binetti G, Bianchetti A, Rozzini R, Trabucchi M (1996) Mini-Mental State Examination: a normative study in Italian elderly population. *Eur J Neurol* 3:198–202. <https://doi.org/10.1111/j.1468-1331.1996.tb00423.x>
48. Marques JP, Kober T, van der Krueger G, Gruetter R (2010) MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *NeuroImage* 49:1271–1281. <https://doi.org/10.1016/j.neuroimage.2009.10.002>
49. MathWorks (2023) MATLAB (version R2023b) [Computer software]. The MathWorks Inc., Natick
50. Mendonça AR, Loureiro LM, Nórte CE, Landeira-Fernandez J (2022) Episodic memory training in elderly: a systematic review. *Front Psychol* 13:947519. <https://doi.org/10.3389/fpsyg.2022.947519>
51. Meng X, D'Arcy C (2012) Education and dementia in the context of the cognitive reserve hypothesis: a systematic review with meta-analyses and qualitative analyses. *PLoS ONE* 7:e38268. <https://doi.org/10.1371/journal.pone.0038268>
52. Mitchell KJ, Johnson MK, Raye CL, Mather M, D'Esposito M (2000) Aging and reflective processes of working memory: binding and test load deficits. *Psychol Aging* 15:527–541. <https://doi.org/10.1037/0882-7974.15.3.527>
53. Miyake A, Friedman NP, Rettinger DA, Shah P, Hegarty M (2001) How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *J Exp Psychol Gen* 130:621–640. <https://doi.org/10.1037/0096-3445.130.4.621>
54. Monaco M, Costa A, Caltagirone C, Carlesimo GA (2013) Forward and backward span for verbal and visuo-spatial data: standardization and normative data from an Italian adult population. *Neurol Sci* 34:1749–1754. <https://doi.org/10.1007/s10072-012-1130-x>

55. Nelson ME, Jester DJ, Petkus AJ, Andel R (2021) Cognitive reserve, Alzheimer's neuropathology, and risk of dementia: a systematic review and meta-analysis. *Neuropsychol Rev* 31:233–250. <https://doi.org/10.1007/s11065-021-09478-4>
56. Ngo CT, Weisberg SM, Newcombe NS, Olson IR (2016) The relation between navigation strategy and associative memory: an individual differences approach. *J Exp Psychol Learn Mem Cogn* 42:932–948. <https://doi.org/10.1037/xlm0000193>
57. Nieuwenhuis S, Forstmann BU, Wagenmakers EJ (2011) Erroneous analyses of interactions in neuroscience: a problem of significance. *Nat Neurosci* 14:1105–1107. <https://doi.org/10.1038/nn.2886>
58. Nyberg L, Walhovd KB, Fjell AM (2024) Neurobiological aging and memory: average patterns and individual differences. In: Kahana MJ, Wagner AD (eds) *The Oxford handbook of human memory*, vol 2. Oxford University Press, Oxford, pp 560–580. <https://doi.org/10.1093/oxfordhb/9780190917982.013.56>
59. Ondřej J (2025) The method of loci in the context of psychological research: a systematic review and meta-analysis. *Br J Psychol* 116:45–68. <https://doi.org/10.1111/bjop.12799>
60. Orsini A, Pezzuti L (2013) WAIS-IV: Wechsler Adult Intelligence Scale – Fourth Edition. Contributo alla taratura italiana. Giunti O.S. Organizzazioni Speciali, Florence
61. Palani D, Baginskaskas A, Raastad M (2010) Bursts and hyperexcitability in non-myelinated axons of the rat hippocampus. *Neuroscience* 167:304–313. <https://doi.org/10.1016/j.neuroscience.2010.03.021>
62. Palombo M, Ianus A, Guerreri M, Nunes D, Alexander DC, Shemesh N, Zhang H (2020) SANDI: a compartment-based model for non-invasive apparent soma and neurite imaging by diffusion MRI. *NeuroImage* 215:116835. <https://doi.org/10.1016/j.neuroimage.2020.116835>
63. Picconi L, Fairfield B, Sergi MR, Cataldi F, Padulo C, Brugnera A, Compare A (2023) Development and validation of a short form of the Geriatric Anxiety Scale (GAS-12) among Italian older adults. *Clin Gerontol* 46:512–524. <https://doi.org/10.1080/07317115.2022.2120445>
64. Poppenk J, Evensmoen HR, Moscovitch M, Nadel L (2013) Long-axis specialization of the human hippocampus. *Trends Cogn Sci* 17:230–240. <https://doi.org/10.1016/j.tics.2013.03.005>
65. Posit Team (2026) RStudio: integrated development environment for R [Computer software]. Posit Software, PBC, Boston. <https://posit.co>
66. R Core Team (2026) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org>
67. Read M, Umla-Runge K, Lawrence AD, Costigan AG, Hsieh L, Chamberland M, Graham KS (2023) A role for the fornix in temporal sequence memory. *Eur J Neurosci* 57:980–996. <https://doi.org/10.1111/ejn.15940>
68. Rolls ET (2018) The storage and recall of memories in the hippocampo-cortical system. *Cell Tissue Res* 373:57–78. <https://doi.org/10.1007/s00441-017-2744-3>
69. Sala G, Aksayli ND, Tatlidil KS, Tatsumi T, Gondo Y, Gobet F (2019) Near and far transfer in cognitive training: a second-order meta-analysis. *Collabra Psychol* 5:203.

<https://doi.org/10.1525/collabra.203>

70. Salthouse TA, Babcock RL (1991) Decomposing adult age differences in working memory. *Dev Psychol* 27:763–776. <https://doi.org/10.1037/0012-1649.27.5.763>
71. Schlichting ML, Mumford JA, Preston AR (2015) Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nat Commun* 6:8151. <https://doi.org/10.1038/ncomms9151>
72. Schubert T, Strobach T, Karbach J (2014) New directions in cognitive training: on methods, transfer, and application. *Psychol Res* 78:743–755. <https://doi.org/10.1007/s00426-014-0619-8>
73. Simfukwe C, An SSA, Youn YC (2025) Mechanisms and impact of cognitive reserve in normal aging and Alzheimer's disease. *Diagnostics* 15:3068. <https://doi.org/10.3390/diagnostics15233068>
74. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 23:S208–S219. <https://doi.org/10.1016/j.neuroimage.2004.07.051>
75. Stern Y, Arenaza-Urquijo EM, Bartrés-Faz D, Belleville S, Cantilon M, Chetelat G, Ewers M (2020) Whitepaper: defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimers Dement* 16:1305–1311. <https://doi.org/10.1016/j.jalz.2018.07.219>
76. Strange BA, Witter MP, Lein ES, Moser EI (2014) Functional organization of the hippocampal longitudinal axis. *Nat Rev Neurosci* 15:655–670. <https://doi.org/10.1038/nrn3785>
77. Tournier JD, Smith R, Raffelt D, Tabbara R, Dhollander T, Pietsch M, Connelly A (2019) MRtrix3: a fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage* 202:116137. <https://doi.org/10.1016/j.neuroimage.2019.116137>
78. Traut HJ, Guild RM, Munakata Y (2021) Why does cognitive training yield inconsistent benefits? A meta-analysis of individual differences in baseline cognitive abilities and training outcomes. *Front Psychol* 12:662139. <https://doi.org/10.3389/fpsyg.2021.662139>
79. Tucker-Drob EM, Salthouse TA (2011) Individual differences in cognitive aging. In: Chamorro-Premuzic T, von Stumm S, Furnham A (eds) *The Wiley-Blackwell handbook of individual differences*. Wiley-Blackwell, Chichester, pp 242–268. <https://doi.org/10.1002/9781444343120.ch9>
80. Tulving E, Thomson DM (1973) Encoding specificity and retrieval processes in episodic memory. *Psychol Rev* 80:352–373. <https://doi.org/10.1037/h0020071>
81. Tustison NJ, Cook PA, Holbrook AJ, Johnson HJ, Muschelli J, Devenyi GA, Yushkevich PA (2021) The ANTsX ecosystem for quantitative biological and medical imaging. *Sci Rep* 11:9068. <https://doi.org/10.1038/s41598-021-87564-6>
82. Twomey C, Kroneisen M (2021) The effectiveness of the loci method as a mnemonic device: meta-analysis. *Q J Exp Psychol* 74:882–899. <https://doi.org/10.1177/1747021821993457>
83. Veraart J, Fieremans E, Novikov DS (2016) Diffusion MRI noise mapping using random matrix theory. *Magn Reson Med* 76:1582–1593. <https://doi.org/10.1002/mrm.26059>

84. Veraart J, Novikov DS, Christiaens D, Ades-Aron B, Sijbers J, Fieremans E (2016) Denoising of diffusion MRI using random matrix theory. *NeuroImage* 142:394–406. <https://doi.org/10.1016/j.neuroimage.2016.08.016>
85. Whelan CD, Hibar DP, van Velzen LS, Zannas AS, Carrillo-Roa T, McMahon K, Thompson PM (2016) Heritability and reliability of automatically segmented human hippocampal formation subregions. *NeuroImage* 128:125–137. <https://doi.org/10.1016/j.neuroimage.2015.12.039>
86. Williams EJ (1959) The comparison of regression variables. *J R Stat Soc Ser B Methodol* 21:396–399. <https://doi.org/10.1111/j.2517-6161.1959.tb00346.x>
87. Wiltgen T, McGinnis J, Schlaeger S, Kofler F, Voon CC, Berthele A, Wiestler B (2024) LST-AI: a deep learning ensemble for accurate MS lesion segmentation. *Neuroimage Clin* 41:103611. <https://doi.org/10.1016/j.nicl.2024.103611>
88. Wisse LEM, Biessels GJ, Geerlings MI (2014) A critical appraisal of the hippocampal subfield segmentation package in FreeSurfer. *Front Aging Neurosci* 6:261. <https://doi.org/10.3389/fnagi.2014.00261>
89. Zammit AR, Ezzati A, Zimmerman ME, Lipton RB, Lipton ML, Katz MJ (2017) Roles of hippocampal subfields in verbal and visual episodic memory. *Behav Brain Res* 317:157–162. <https://doi.org/10.1016/j.bbr.2016.09.038>
90. Zehnder F, Martin M, Altgassen M, Clare L (2009) Memory training effects in old age as markers of plasticity: a meta-analysis. *Restor Neurol Neurosci* 27:497–515. <https://doi.org/10.3233/RNN-2009-0491>
91. Zeidman P, Maguire EA (2016) Anterior hippocampus: the anatomy of perception, imagination and episodic memory. *Nat Rev Neurosci* 17:173–182. <https://doi.org/10.1038/nrn.2015.24>
92. Zhang Y, Brady M, Smith S (2001) Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 20:45–57. <https://doi.org/10.1109/42.906424>

Figures

Training change trajectories

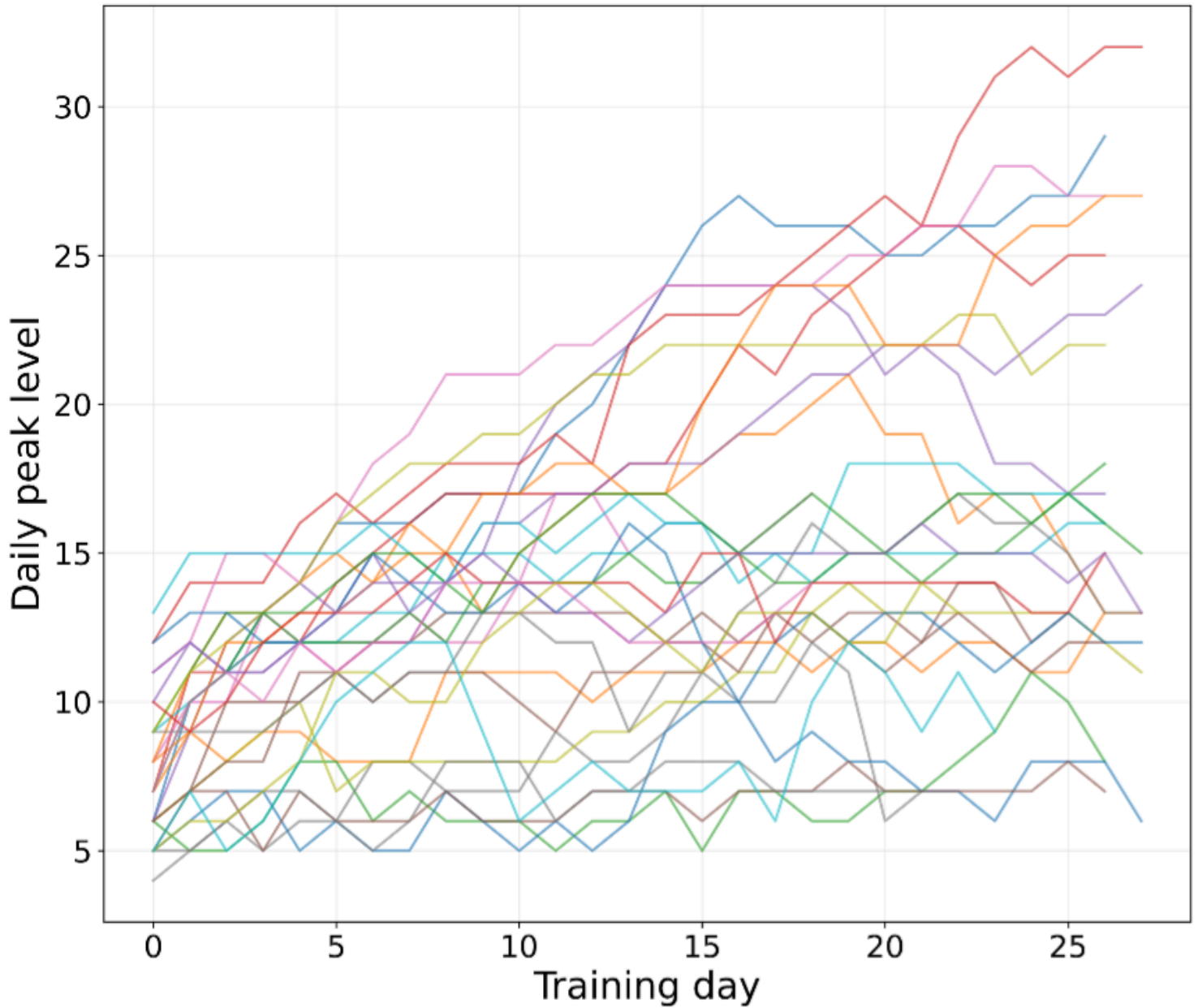


Figure 1

Individual training trajectories across days. Each line represents one participant's daily peak level.

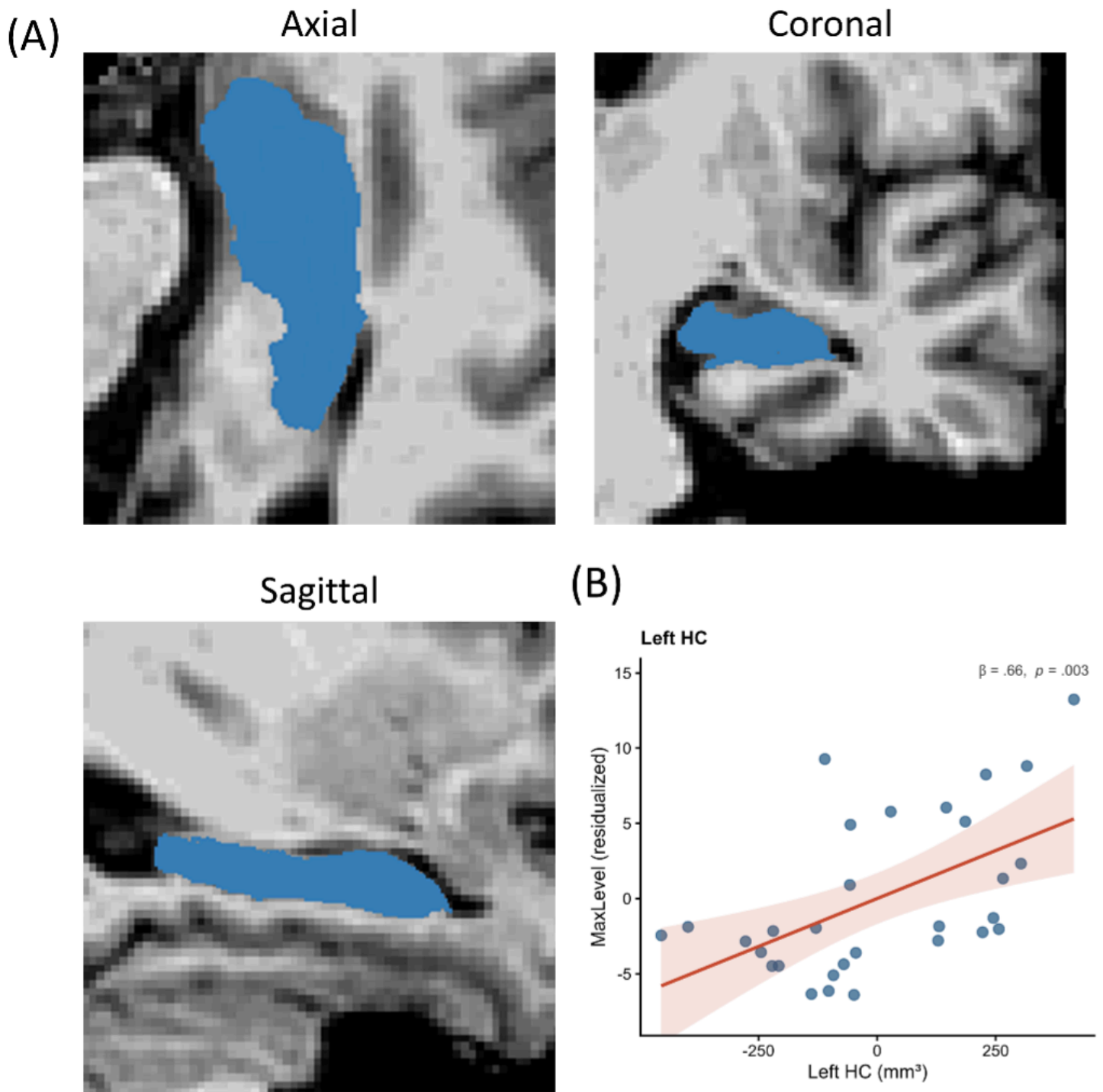
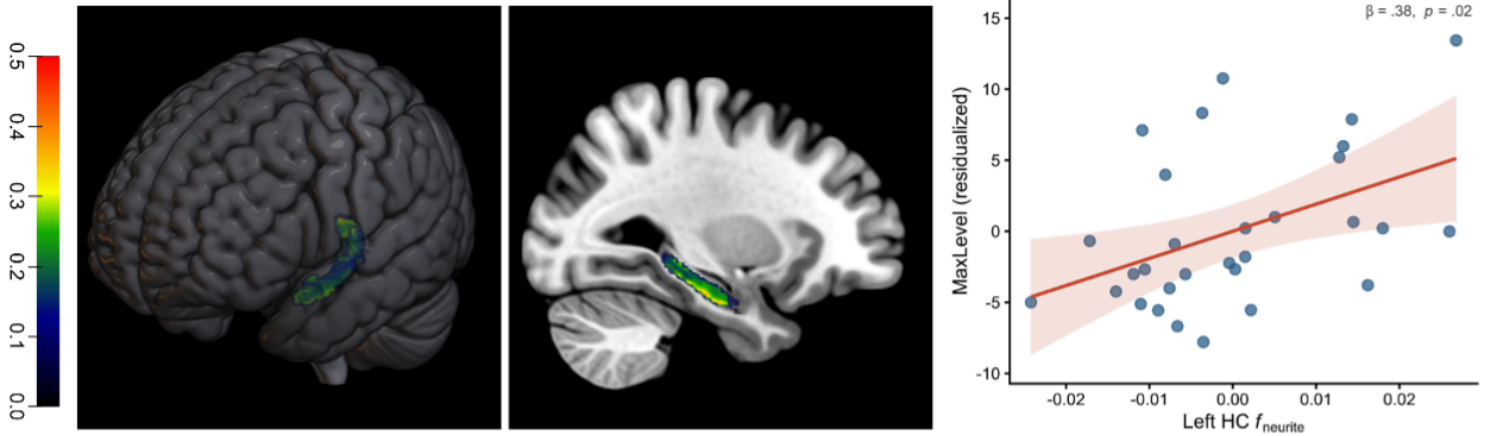


Figure 2

Hippocampal volume and MaxLevel. (A) Partial regression plots showing the association between hippocampal volume and MaxLevel, controlling for education and TBV. (B) Representative left hippocampal segmentation in axial, coronal, and sagittal views.

(A) Left hippocampal f_{neurite}



(B) Left Hippocampal f_{soma}

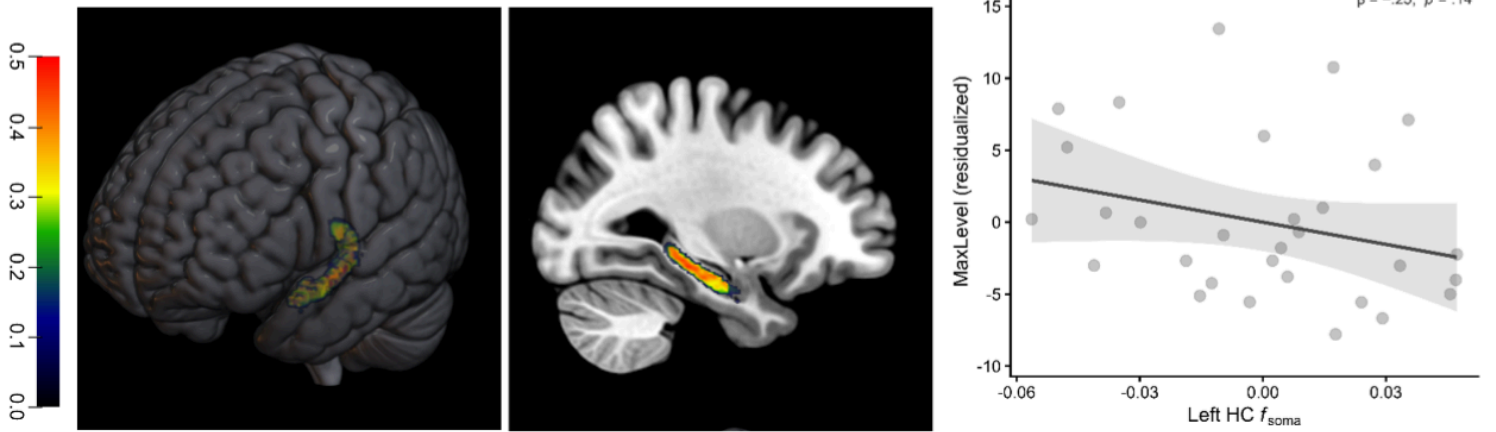


Figure 3

Microstructural metrics of the left hippocampus and their association with training performance. (A) Left hippocampal neurite fraction (f_{neurite}) and (B) soma fraction (f_{soma}) with surface and sagittal views showing their spatial distributions. Scatter plots show associations with residualized MaxLevel, controlling for education. f_{neurite} was positively associated with MaxLevel ($\beta = .38$, $p = .02$), whereas f_{soma} was not ($\beta = -.25$, $p = .14$). Solid lines indicate linear regression fits; shaded areas indicate 95% CIs.

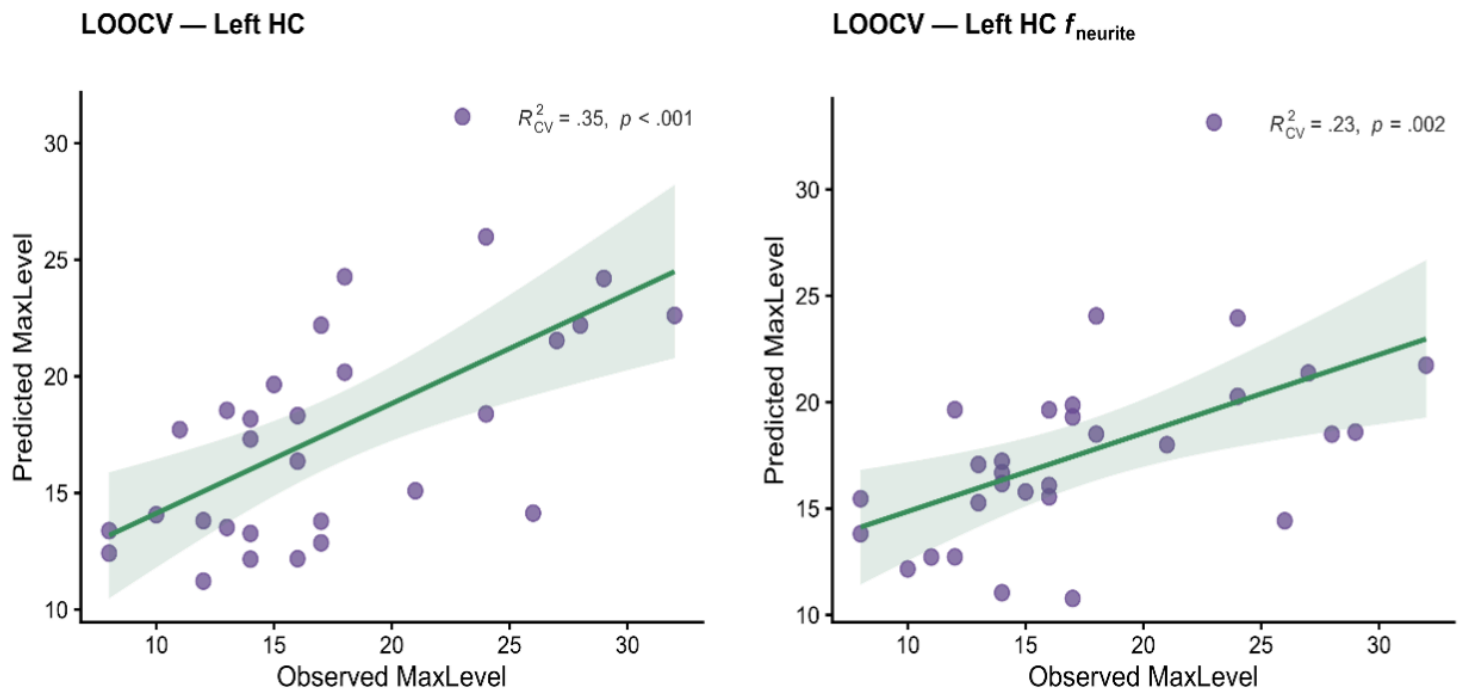


Figure 4

Leave-one-out cross-validation (LOOCV) of the main left hippocampal models predicting MaxLevel. (A) Volume model (TBV + education + left hippocampal volume). (B) SANDI model (education + left hippocampal $f_{neurite}$). Points represent observed and LOOCV-predicted MaxLevel scores; shaded areas indicate 95% CIs.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ESM1.docx](#)