Cerebral Cortex, April 2020;30: 2465–2477

doi: 10.1093/cercor/bhz252
Advance Access Publication Date: 4 December 2019
Original Article

ORIGINAL ARTICLE

Hippocampal Subfields and Limbic White Matter Jointly Predict Learning Rate in Older Adults

Andrew R. Bender[1,2], Andreas M. Brandmaier[2,3], Sandra Düzel[2], Attila Keresztes[2,4,5], Ofer Pasternak[6], Ulman Lindenberger[2,3,7] and Simone Kühn[2,8]

[1]Departments of Epidemiology and Biostatistics, Neurology and Ophthalmology, College of Human Medicine, Michigan State University, East Lansing, MI 48824, USA, [2]Center for Lifespan Psychology, Max Planck Institute for Human Development, D-14195 Berlin, Germany, [3]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, D-14195 Berlin, Germany and London, UK WC1B 5EH, [4]Research Centre for Natural Sciences, Hungarian Academy of Sciences, H-1117 Budapest, Hungary, [5]Faculty of Education and Psychology, Eötvös Loránd University, H-1053 Budapest, Hungary, [6]Departments of Psychiatry and Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA, [7]European University Institute, I-50014. San Domenico di Fiesole, Italy and [8]Department of Psychiatry and Psychotherapy, University Clinic Hamburg-Eppendorf, 20246 Hamburg, Germany

Address correspondence to Andrew R. Bender, Department of Epidemiology and Biostatistics, Michigan State University, 909 Wilson Rd., Rm B601, East Lansing, MI 48824, USA. Email: arbender@msu.edu

Abstract

Age-related memory impairments have been linked to differences in structural brain parameters, including cerebral white matter (WM) microstructure and hippocampal (HC) volume, but their combined influences are rarely investigated. In a population-based sample of 337 older participants aged 61–82 years (Mage = 69.66, SDage = 3.92 years), we modeled the independent and joint effects of limbic WM microstructure and HC subfield volumes on verbal learning. Participants completed a verbal learning task of recall over five repeated trials and underwent magnetic resonance imaging (MRI), including structural and diffusion scans. We segmented three HC subregions on high-resolution MRI data and sampled mean fractional anisotropy (FA) from bilateral limbic WM tracts identified via deterministic fiber tractography. Using structural equation modeling, we evaluated the associations between learning rate and latent factors representing FA sampled from limbic WM tracts, and HC subfield volumes, and their latent interaction. Results showed limbic WM and the interaction of HC and WM—but not HC volume alone—predicted verbal learning rates. Model decomposition revealed HC volume is only positively associated with learning rate in individuals with higher WM anisotropy. We conclude that the structural characteristics of limbic WM regions and HC volume jointly contribute to verbal learning in older adults.

Key words: aging, hippocampus, memory, verbal learning, white matter

Introduction

Age-related deficits in verbal learning and memory have a long history of study in psychological and cognitive sciences (Kausler 1994; Korchin and Basowitz 1957). Individual differences in learning and memory in older adults are linked with differences in both regional gray matter volumes in the medial temporal lobe (Petersen et al. 2000) and microstructural measures of limbic white matter (WM) pathways (see Madden et al. 2012 for a review). However, few studies to date have investigated

the influences of both classes of neuroanatomical correlates of verbal learning, using two magnetic resonance imaging (MRI) modalities, in a population-based cohort of older adults. Arguably, modeling learning and memory as a function of a larger, integrated neural system affords a more balanced perspective over traditional univariate modeling of individual neural structures (Aggleton 2014).

Verbal learning is commonly tested via serial presentation of lexical stimuli, followed by tests of free or cued recall and recognition. For instance, multiple neuropsychological instruments assessing verbal learning repeatedly present the same stimulus list over multiple successive learning trials; following each presentation, participants freely recall as many items as possible (Baldo et al. 2002; Schmidt 1996). Although the sum of recalled items is commonly used as a measure of aggregate performance, the slope of change in memory performance across learning trials can serve as an estimate of the rate of learning (Jones et al. 2005). Although many studies have explored the neural correlates of age-associated decrements in delayed mnemonic retrieval, fewer have investigated the structural brain correlates associated with learning rate (Gifford et al. 2015). These initial studies suggest that individual differences in learning "rate" are associated with hippocampal (HC) volumes in normal aging and mild cognitive impairment (Bonner-Jackson et al. 2015; Gifford et al. 2015), and may afford a more sensitive behavioral correlate of brain organization. Although prior work has linked larger HC subregional volumes in Cornu Ammonis (CA) and dentate gyrus (DG) to better verbal learning, associative memory, and higher longitudinal retest improvements (Bender et al. 2013; Mueller et al. 2011; Shing et al. 2011), this has not been investigated as a correlate of learning rate.

In addition to HC, other established neuroanatomical correlates of age-associated memory decline include structurally linked afferent and efferent limbic WM pathways measured via diffusion MRI (dMRI; Bender et al. 2016; Bennett et al. 2014; Charlton et al. 2013; Fletcher et al. 2013; Henson et al. 2016; Metzler-Baddeley et al. 2011a; Sasson et al. 2013; Sepulcre et al. 2008; Stoub et al. 2006; see Preston and Eichenbaum 2013; Shing et al. 2008 for reviews). In particular, these extant reports implicate cingulum bundle, including dorsal and parahippocampal segments, fornix, and uncinate fasciculus (UF) as primary WM correlates of age-related memory declines. Furthermore, associations between total HC volume and memory in older adults are inconsistent (Van Petten 2004) and suggest that other, less elucidated factors may modify this relationship. One possibility is that its structure includes multiple functionally and cytoarchitectonically distinct subregions, which show differential associations with aging and with mnemonic processes (Braak et al. 1996; Duvernoy 2005; Insausti et al. 1998; Kiernan 2012; Wilson et al. 2006). This perspective suggests that the relationship between total HC volume and memory may be attenuated by the lack of functional and structural specificity (Van Petten 2004). Alternatively, associations between HC volume and memory may be modified by related factors, such as the extent of WM connectivity (Foster et al. 2019; Metzler-Baddeley et al. 2019). However, the statistical interaction between HC volumes and limbic WM microstructure has not been tested previously.

Investigating the links of different HC subfields and limbic WM fiber tracts to episodic memory requires a sufficiently large sample and the incorporation of multiple MRI measurement approaches. The Berlin Aging Study-II (BASE-II; Bertram et al. 2014; Gerstorf et al. 2016) includes over 1500 healthy older adults; of these, a subset of whom underwent MRI neuroimaging. The sample is sufficiently large to permit the use of structural equation modeling (SEM), which allows the confirmatory construction of memory performance and brain parameters as latent factors, moving from an observed level to a more valid construct level. These latent factors then serve as a basis for exploring brain—behavior associations as between—construct correlations, independent of measurement error. Thus, following the plea of Brandmaier et al. (2013), we use SEM as a statistical tool that combines the benefits of both confirmatory and explanatory modes of scientific inquiry.

Specifically, we were interested in modeling the rate of learning across five free recall trials of a test of verbal learning (Helmstaedter and Durwen 1989; Schmidt 1996) to test the notion that the rate of learning is a more sensitive correlate of brain structure and organization than the intercept. By evaluating the combined contributions from both limbic WM and HC subfield volumes under the SEM framework, our goal was to simultaneously model multiple distinct, but interdependent structural neural correlates of learning. Our initial approach was more exploratory by freely estimating all associations between individual HC and WM factors and verbal learning slope and intercept. We expected a similar pattern of anatomical associations with learning as previously reported: higher fractional anisotropy (FA) in cingulum and fornix tracts and larger HC volumes, particularly of CA and DG (Bender et al. 2016; Bennett et al. 2014; Madden et al. 2012). Herein, we also tested the general hypothesis that aggregate, latent measures of HC volume and limbic WH-FA in older adults are both associated with verbal learning rate. Furthermore, we hypothesized that individual differences in rates of learning would be associated with individual differences in the interaction between HC volume and FA in limbic WM.

Materials and Methods

Participants

Study data were drawn from the first wave of the BASE-II cohort (Bertram et al. 2014; Gerstorf et al. 2016), a population-based study of older and younger adults living in Berlin, Germany. The baseline cohort included 1532 older adults from 60 to 88 years of age. None of the participants took medication that might affect memory function, and none had neurological disorders, psychiatric disorders, or a history of head injuries. All participants reported normal or corrected to normal vision and were right-handed. All participants were invited to two cognitive sessions with an exact interval of 7 days and at the same time of day to avoid circadian confounding effects on session differences in performance. Participants were tested in groups of 4–6 individuals. The ethics section of the German Psychological Society approved the study (SK 012013_6). All participants had provided informed consent in accord with the Declaration of Helsinki.

After completing the comprehensive cognitive assessment in BASE-II, MR-eligible participants were invited to take part in one MRI session within a mean time interval of 3.2 months after the cognitive testing. The subsample consisted of 345 older adults aged 61–82 years (mean age 70.1 years, SD = 3.9 years, 39% female). We excluded six participants following technical errors in cognitive test administration, and we excluded two additional participants with scores below 25 on the Mini-Mental Status Examination (MMSE; Folstein et al. 1975). Most participants' MMSE scores were well above this cut-off (mean = 28.61, SD = 1.15). BASE-II participants in the MRI cohort did not differ from those who did not undergo MRI scanning in terms of educational attainment, cognitive performance, or MMSE scores (for

all, $t < 1.0$), although the MRI cohort was significantly younger than the non-MRI cohort ($t = 2.577$, $P < 0.05$) by approximately 6 months. The final sample retained for analysis included 337 older adults (mean age = 69.66, SD = 3.92 years). Sample demographics showed a greater proportion of men (61.7%) than women (38.3%) and the mean level of years of education nearing 1 year of university (mean = 14.07, SD = 2.90 years).

Magnetic Resonance Imaging

Image Acquisition

All MRI data were acquired on a 3 T Siemens MagnetomTim Trio scanner. For most cases, a 32-channel head coil was used, although in two cases a 12-channel coil was used as the 32-channel coil provided an uncomfortable fit. MRI data acquisition included a T_1-weighted magnetization-prepared rapid gradient echo (MPRAGE) sequence, acquired in the sagittal plane with a single average, repetition time (TR) = 2500 ms, echo time (TE) = 4.77 ms, with an isotropic voxel size of $1.0 \times 1.0 \times 1.0$ mm, using a 3D distortion correction filter and prescan normalization with FOV = 256, and generalized autocalibrating partially parallel acquisitions (GRAPPA) acceleration factor = 2. Acquisition also included a single T_2-weighted, turbo spin echo high-resolution sequence in a coronal direction, oriented perpendicularly to the long axis of the left HC, with voxel size = $0.4 \times 0.4 \times 2.0$ mm^3, 30 slices. TR = 8150 ms, TE = 50 ms, and flip angle = 120°, positioned to cover the entire extent of the HC. A single-shot, echo-planar imaging, diffusion weighted sequence was also acquired in transverse plane with TR = 11 000 ms, TE = 98 ms, in 60 gradient directions, diffusion weighting of $b = 1000$ s/mm^2, seven volumes collected without diffusion weighting ($b = 0$), and GRAPPA acceleration factor = 2 with an isotropic voxel of 1.70 mm^3.

Diffusion MRI Processing

All diffusion-weighted images (DWIs) underwent an initial quality control (QC) process using DTIPrep v. 1.2.4 (Oguz et al. 2014), software to eliminate noisy gradient directions and correct for motion and eddy currents.

Free Water Estimation and Removal

The influence of partial volume artifacts from cerebral spinal fluid (CSF) is an established limitation of the single-tensor diffusion tensor imaging (DTI) model that is often used in the context of tractography (Concha et al. 2005; Metzler-Baddeley et al. 2011b; Pasternak et al. 2009), particularly in regions directly adjacent to ventricular CSF, such as the fornix (Jones and Cercignani 2010; Metzler-Baddeley et al. 2011a). To address this limitation, tensor data were corrected for CSF contamination on a voxel-by-voxel basis, using the free water elimination MATLAB code (Pasternak et al. 2009), resulting with free water corrected diffusion tensors. The free water corrected tensors were then decomposed using FSL to produce FA image maps.

Sample Template Creation

We used DTI-TK (Zhang et al. 2006), software to align all participants' data into a common template. The complete procedures for intersubject registration are detailed in Supplementary Material.

Region of Interest Creation

Individual regions of interest (ROIs) reflecting seed regions, regions of inclusion, and regions of exclusion were drawn on the template-space FA image, colored by orientation image output as an option by DTI-TK in ITK-SNAP (www.itksnap.org; Yushkevich et al. 2006). All template-space ROIs were nonlinearly deprojected to native space, where they were inspected for errors by one of the authors (A.R.B.). Following published deterministic tractography approaches for these regions (Bennett et al. 2014; Malykhin et al. 2008; Metzler-Baddeley et al. 2011a), we created ROIs for tractography of four, bilateral limbic WM: dorsal cingulum bundle (CBD), parahippocampalcingulum bundle (CBH), posterior fornix, and UF. The procedures for tractography mask creation, placement, and spatial transformation from standard to native space are detailed in Supplementary Material.

Constrained Spherical Deconvolution Tractography

To enhance the anatomical validity and minimize the potentially confounding influences of crossing fiber populations, we performed diffusion MR tractography using constrained spherical deconvolution (CSD), a method that fits a fiber-orientation distribution (FOD) to each voxel and performs tractography based on peaks in the FOD (Tournier et al. 2007; Tournier et al. 2004). CSD tractographyis considered superior to other commonly used approaches for delineating WM tracts of interest, such as WM skeletonization, due to greater anatomical precision for any given tract (Metzler-Baddeley et al. 2011a). MRtrix3 (Tournier et al. 2007; Tournier et al. 2004) software was used for CSD-based deterministic tractography on the DWI data following QC, but before any free water correction. Response function estimation used the method previously described (Tournier et al. 2013) with maximum spherical harmonic degree = 4. Following default procedures (Beginner DWI Tutorial 2017), two separate FOD images were produced using the estimated response function, one using a whole brain mask, and the other using the thresholded FA mask (here, FA prior to free water elimination was used). The whole brain mask was used only for tractography of the fornix, as using the thresholded FA masked FOD data did not permit sufficient information for reliable tractography of fornix, whereas the use of the whole brain mask for other regions produced excessive spurious and anatomically implausible streamlines (i.e., crossing sulci). For fornix, a secondary mask was also applied, in which the operator (A.R.B.) drew an inclusionary ROI to cover the fornix, but excludes the other regions of the ventricles, the thalamus, and choroid plexus. Additional exclusionary masks were liberally applied outside these regions to limit any spurious streamlines. All masks were deprojected for individual, native space tractography. The thresholded, FA-masked FOD data were used for CSD tractography of the other regions: CBD, CBH, and UF. All streamline outputs were inspected by the same person (A.R.B.) using the MRtrix image viewer to ensure the complete inclusion and to check for spurious streamlines (see Supplementary Fig. SM1 for an example of sampled streamlines). Additional information on streamline inspection is available in Supplementary Material. Following streamline generation and inspection, we used streamlines to sample median free water-corrected FA values and then calculated the mean value across all streamlines in each tract of interest.

HC Subfield Morphometry

HC subfield regions included were based on methods from prior work (Bender et al. 2013; Daugherty et al. 2016; Mueller et al. 2011; Mueller et al. 2007; Shing et al. 2011), and included separate regions for SUB and the aggregations of CA1 and 2 (CA1/2) and an aggregation of CA3 and DG (CA3/DG).

Optimized Automated Segmentation

We used the Automated Segmentations of HC Subfields (ASHS; Yushkevich et al. 2015; Yushkevich et al. 2010) software with a customized atlas for HC subfield morphometry (Bender et al. 2018). Because it uses multi-atlas label fusion methods, ASHS may be more sensitive to individual differences in HC subfield morphology, than single-atlas approaches, such as Freesurfer. The customized atlas was built using a modified version of the manual demarcation and tracing rules described previously (Daugherty et al. 2016; Keresztes et al. 2017; see Supplementary Material for more information).

ICV Correction

We sampled the intracranial vault (ICV) as described previously (Bender et al. 2013; Keihaninejad et al. 2010), using the brain extraction tool (Smith 2002) in FSL 5.0 (Jenkinson et al. 2005, 2012) on the MPRAGE images (further detail is provided in Supplementary Material).

Cognitive Testing

The Verbal Learning and Memory Test (VLMT; Helmstaedter and Durwen 1989) is a German version of the Rey Auditory Verbal Learning Test (Schmidt 1996). Participants heard a list of 15 nouns, serially presented via headphones. Presentation of the list was followed by a recall phase in which a computer screen prompted the participants to type as many words as they could remember from the list. This was repeated over five learning trials, each with the same word list. The same list of German nouns was used for all participants. The present analyses are based on the five verbal learning trials and do not include data from the delayed recall and recognition tasks that are also a part of the VLMT.

Data Conditioning

ICV values were divided by 1000 to align the scales of HC subfields and ICV and to increase numerical stability in the parameter estimation. We corrected each of the subregional HC volumes for ICV using the analysis of covariance approach (Bender et al. 2013; Jack et al. 1989; Raz et al. 2005). All analyses reported below used the adjusted HC subfield volumes. In addition, all FA values were centered at their respective sample means and HC subfield volumes were standardized to z-scores.

Data Analysis

Overview

Data modeling and analysis was performed in Mplus 7 (Muthén and Muthén 2012). We used latent factor analysis to explore associations among verbal learning, HC subfield volumes, and WM-FA within an overall multivariate model. Model specification proceeded in an iterative fashion, as we sought to establish the validity of each modeled domain (i.e., WM-FA, HC subfields, verbal learning). The first step involved specifying, testing, and refining individual measurement models for latent variables (factors) in each domain, including: 1) a latent growth model (LGM) across learning trials (specified below) yielded factors for intercept and slope; 2) three separate HC subfield volume factors; and 3) four latent factors representing mean FA of four WM tracts of interest (see Supplementary Table SM1 for factor loadings). Following this, we combined these individual measurement models with a structural model to test associations across domains between all latent factors (Fig. 1). Next, we specified an additional model to test the fit of separate second-order factors representing the HC and WM factors (Fig. 2A). Second-order factors are constructed from other estimated factors, rather than observed indicators. Thus, this approach permits the estimation of overall HC and WM factors based on their constituent factors by subregion or tract, which should provide more reliable brain factor estimates. We then tested regression paths from both second-order factors for HC and WM to the learning slope factor. Last, we modeled the latent interaction (Fürst and Ghisletta 2009; Maslowsky et al. 2015; Little et al. 2006) between the second-order factors for HC and WM to test the hypothesis that the association between HC volume and learning is modified by WM microstructure.

All models used full information maximum likelihood to account for missing data without requiring pairwise deletion. Goodness of model fit was assessed by multiple indices, including chi-square (χ^2), chi-square value divided by degrees of freedom (χ^2/df), comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). We evaluated models according to commonly accepted goodness-of-fit thresholds, that is, nonsignificant chi-square values, CFI values >0.95, RMSEA values reliably <0.05, and SRMR values <0.05 indicate good model fit (Bentler 1990; Hooper et al. 2008).

Verbal Learning Latent Growth Model

We modeled verbal learning by fitting a latent free basis model (McArdle 1986), a type of LGM (Duncan et al. 2013; Meredith and Tisak 1990) to the five learning trials. LGM is commonly used to assess latent change over successive occasions, and generally includes separate factors for the intercept and for change trajectories. LGM commonly specifies fixed factor loadings in incrementing value following a specific change function (i.e., linear, quadratic, etc.) across sequential indicators. In contrast, the latent free basis model fixes the factor loadings of the first and last trials at 1 and 0, respectively; the other factor loadings are freely estimated. Descriptions of the full modeling procedures for the LGM are reported in Supplementary Material. Briefly, however, it should be noted that initial attempts to fit the LGM with only intercept and linear slopes were poor fit for the data (Supplementary Table SM3). However, latent basis models may be the best approaches for estimating relative change over trials when linear, quadratic, cubic, or other functional patterns do not provide a good fit for the data (Berlin et al. 2013).

WM Tract and HC SubregionalFactor Models

To model each brain imaging parameter, we used a confirmatory factor analysis approach (see Supplementary Material for a complete description). That is, we fit individual, single-factor models for each of the three HC subfield volumes and for each of the four WM tracts. For all single-factor models, left and right hemisphere brain parameter measures (i.e., FA or mm^3 brain volume) served as dual indicators (Bender and Raz 2015; Raz et al. 2005). In all dual-indicator models of brain imaging parameters, the factor's variance parameter was fixed to 1 and the factor loadings for both indicators were freely estimated, as were the residual variances. This approach of separate factors per region was preferred over combining all indicators loading into a single factor (see Supplemental Materials for a complete description).

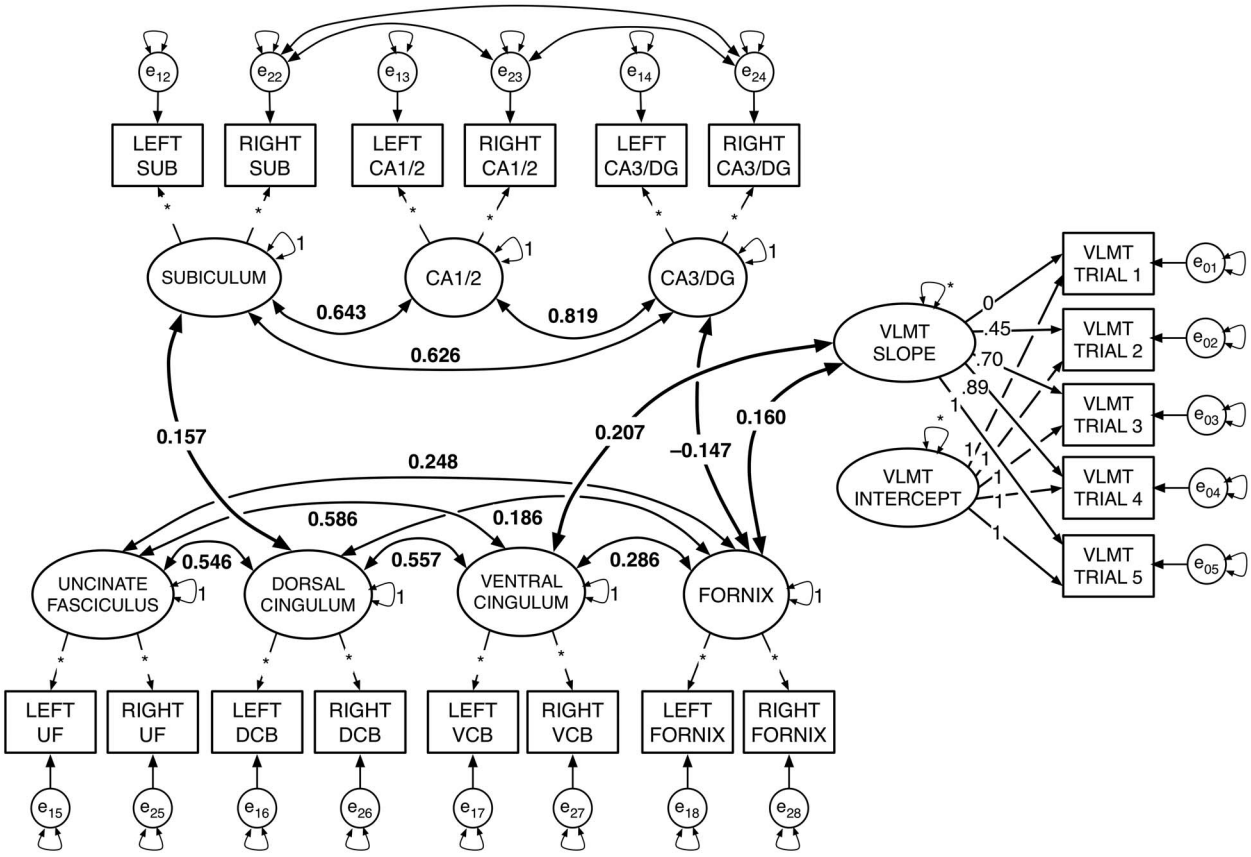

Figure 1. Diagram of the "combined model," with only significant correlations (i.e., $P < 0.05$) shown. The estimated model included the fully saturated latent correlation matrix. Larger ellipses with small double-headed arrows represent latent factors with variance either fixed at 1 or freely estimated (*). Small single-headed arrows between factors and their respective observed indicators reflect factor loadings, with the value of the factor loading prespecified (i.e., based on estimated loadings from earlier modeling steps) or freely estimated (*). The rectangles reflect the observed indicators for each measurement, and the small circles with double-headed arrows reflect their residuals and residual variance; all residual error variance parameters were freely estimated. In the LGM portion on the right side of the figure, factor loadings for the LGM slope factor were originally estimated using a latent basis free model. The intercept (i.e., all factor loadings fixed to 1), and slope with factor loadings represent individual differences in growth across the verbal learning trials. (See Supplementary Table SM3 for a comparison for factor loadings and fit to a model with linear slope). Larger curved bidirectional arrows represent significant covariances between factors ($P < 0.05$), and covariance path values reflect standardized parameters. The figure shows only the significant associations within each brain domain (i.e., among HC subfield factors, and among factors for WM tract FA), and between WM and HC factors. In addition, the bold covariance double-headed arrows show significant associations between brain factors and the slope factor for the LGM on learning trials.

Combined Model

Next, we specified a combined model that included: 1) the verbal learning intercept and slope factors; 2) the three factors of HC subfields volume; and 3) four factors of FA in WM tracts. Then, we estimated a fully crossed latent covariance matrix for each combined model. We used bootstrapped resampling with 1000 draws to generate confidence intervals around the combined model parameters and to test significance of the combined model parameters.

Second-Order Factor Model

In the following steps, we respecified the model to estimate two second-order factors: one representing all four WM factors and another one representing the four HC subfield factors (Fig. 2A). This model was initially estimated with covariances freely estimated among factors. Following the observation of good fit, we then respecified the model to include directional regression paths from the factors representing both WM and HC to the LGM slope and intercept factors. Although such regression procedures imply causal relationships, it is worth noting that the data were cross-sectional and thus cannot inform regarding order or directionality of age-related changes (Lindenberger et al. 2011).

Latent Moderation Models

Following successful convergence and estimation of the second-order factor model, we followed published suggestions for modeling the latent interaction (Fürst and Ghisletta 2009; Maslowsky et al. 2015; Little et al. 2006) between the factors for the HC and WM second-order factors. We specified the regression paths from the two latent factors representing the brain (i.e., WM and HC), and from their latent interaction to verbal learning to test if the effect of HC on learning varies across levels of WM, and vice versa. We then compared model fit between the models with and without the latent interaction using log-likelihood ratio tests. In addition, we estimated difference in R^2 and variance accounted for in learning rate with and without estimating the interaction between WM and HC. Next, we applied the Johnson-Neyman (1936; Preacher et al. 2006) technique for plotting the effects of each factor in the interaction, WM and HC, on the

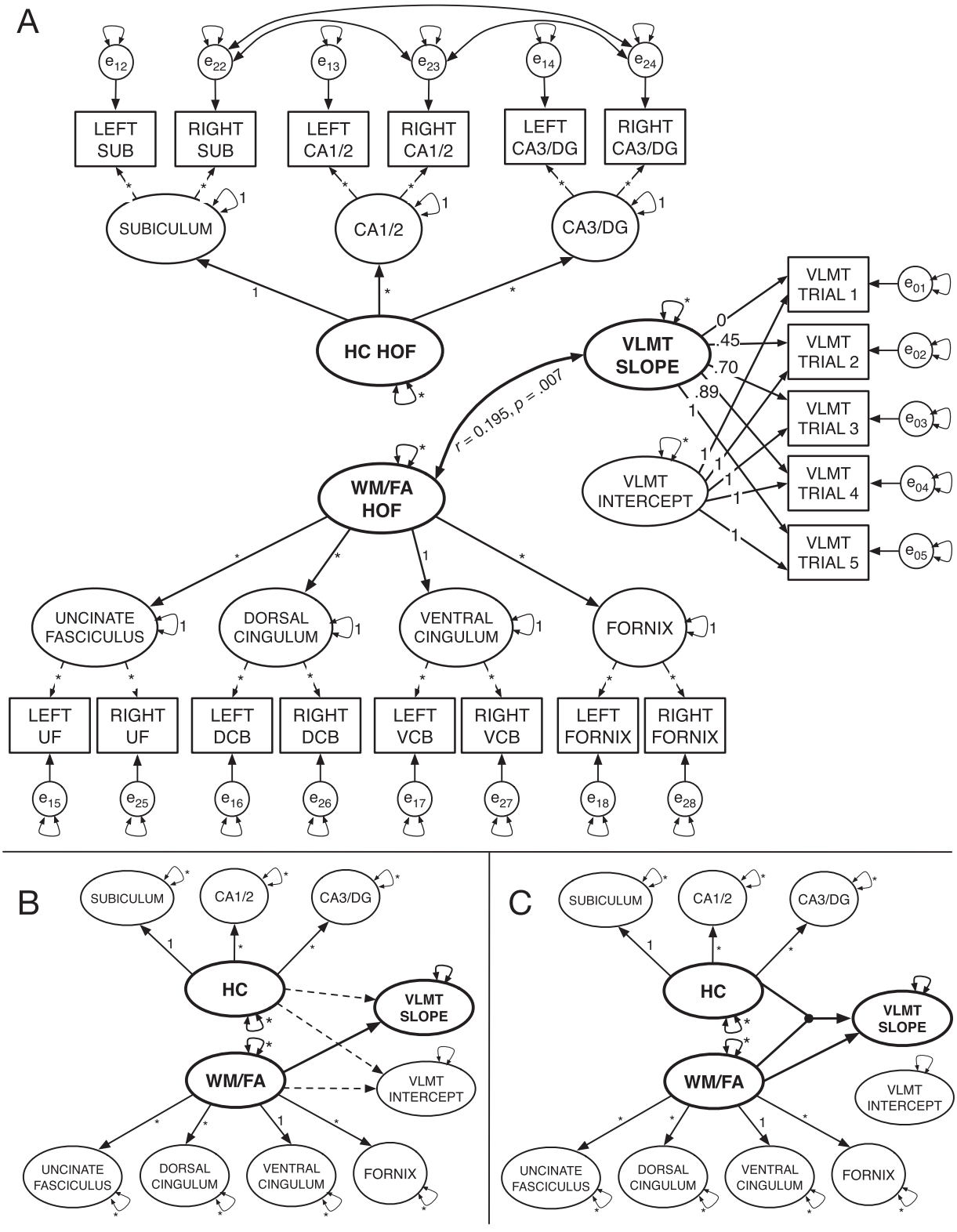

Figure 2. Data modeling steps for second-order models and latent moderated structural equation models. (A) Initial specification of the second-order latent factor model. No covariances or regression paths between factors are illustrated. (B) Reduced illustration of the specified model (indicators and error variances not shown). Initial specification without latent interaction included regression paths from the HC and WM/FA second-order factors to the slope and intercept factors. Dashed lines indicate nonsignificant regression paths, and solid lines reflect significant paths. (C) The latent moderation model showing significant paths from HC and WM/FA factors to the slope factor. The dot symbolizes the latent interaction between the HC and WM/FA factors.

learning rate factor, at different levels of the other. That is, we plotted the effects of WM on learning rate at different levels of HC volume, and vice versa. Last, we tested simple slopes of each WM or HC predictor on learning slope for each level of the other.

Covariate Models

To determine if our findings were influenced by relevant demographic variables, we re-evaluated the combined, second-order factor and latent interaction models with the inclusion of the covariates age in years, sex, and number of years of formal education. Years of age and education were centered at their respective sample means.

Results

Associations between WM, HC, and Learning Parameters

Initial models for WM and HC showed that individual factor models by subregions or WM tracts fit better than single factors models (Supplementary Table SM2). Similarly, the latent basis free LGM fit better than modeling learning as a linear slope. Furthermore, combining the verbal learning LGM with the individual factors for the seven factors representing HC subfield volumes and limbic WM tracts also resulted in excellent fit. This combined model estimated associations between the seven factors representing the structural brain parameters—free water corrected FA in four limbic WM fiber tracts and ICV-corrected volumes in three HC subregions—and verbal learning (Fig. 1).

In the combined model, we found that higher learning rate was significantly associated with higher FA in CBH (standard estimate = 0.207, $P = 0.002$) and fornix (standard estimate = 0.160, $P = 0.025$). No additional significant associations were observed between HC subfield factors and verbal learning. Of note, the intercept factor was not significantly associated with any brain factor or with the slope factor.

Second-Order Factor Model

Based on the combined model, we also specified a model in which the four WM factors UF, CBD, CBH, and fornix, load onto a second-order factor representing WM, and the three HC subfield factors SUB, CA1/2, and CA3/DG load onto the HC volume second-order factor (Fig. 2A). Following the observation of a nonsignificant relationship between brain factors and the intercept of the LGM, we respecified the model to only estimate the direct paths between the HC and WM second-order factors and the learning slope factor; this model specification proved a good fit for the data (Supplementary Table SM2). In addition, the R^2values output by Mplus showed the second-order factors accounted for a large and significant proportion of the variance in their constituent factors (Supplementary Table SM4). However, the only significant covariance between factors was a positive association between the WM factor representing combined FA and the learning slope parameter ($r = 0.195$, $P = 0.007$). Notably, HC was associated with neither the WM factor nor the learning curve factor. In addition, the LGM intercept factor was also unrelated to other latent factors.

Latent Moderation Model

Next, using the latent moderated SEM (LMS) approach as implemented in Mplus, we estimated: 1) the latent interaction between the HC and WM second-order factors and 2) the direct regression paths from the HC and WM second-order factors and the latent interaction parameter to the intercept and slope on the LGM (Fürst and Ghisletta 2009; Maslowsky et al. 2015; Little et al. 2006). The use of the maximum likelihood estimation with robust standard errors estimator necessary for LMS in Mplus does not provide standard fit indices for model comparison. Thus, we used log-likelihood ratio tests to compare fit between the models with and without the latent interaction. Modeling the latent moderation effect resulted in a significantly better fit (Supplementary Table SM6). In addition, we estimated difference in R^2 and variance accounted for in learning rate with and without estimating the interaction between WM and HC and calculated the differences in explained variance using the formula provided by Maslowsky et al. (2015; Table 1). The model with the latent interaction explained an additional 3.6% of variance in verbal learning over models without estimating parameter.

Next, we followed the Johnson-Neyman (1936) technique for plotting the effects of each factor in the interaction, WM, and HC, on the learning rate factor, at different levels of the other. That is, we plotted the effects of WM on learning rate at different levels of HC volume, and vice versa. Last, we extended this approach to test simple slopes of each WM or HC predictor on learning slope for each level of the other predictor (Clavel 2015; Fig. 3).The effects of both HC and WM factors on rate of verbal learning are only apparent at higher levels of the other. In other words, the positive effect of HC volume on the learning slope factor is only apparent at values of the WM factor above the sample mean. Moreover, the effect of WM on the verbal learning slope factor is only significant in individuals with HC factor scores above the sample mean (Supplementary Fig. SM3). We confirmed this by respecifying the model to include additional constraints to test the simple slope of HC on verbal learning separately for WM factor standardized values of ±0.5 (Clavel 2015). Model results showed that whereas the low slope of HC on the learning factor at −0.5 on the WM factor was nonsignificant (estimate = 0.050, $P = 0.841$; 95% CI = −0.436 to 0.535), the high slope was significant (estimate = 0.536, $P = 0.015$; CI = 0.105–0.967). The positive relationship between HC volume and verbal learning rate was only apparent among those with higher FA in limbic WM.

To further probe the moderation effect, we saved the standardized factor scores from the model using the factor score regression method. Subsequently, we subdivided the sample distributions for the standardized WM and HC factors into tertiles and used these to examine the results of Johnson-Neyman plots. Our objective for these follow-up analyses was to identify whether different WM × HC patterns in this population-based sample might further qualify differences in learning. Examining bootstrapped (1000 draws), zero-order correlations between learning slope and HC volume by three different levels of WM (Fig.4) showed that individuals in the lowest tertile of FA in limbic WM exhibit negative associations between HC volume and learning slope ($r = -0.221$, $P = 0.016$; 95% CI = −0.381 to −0.036), which differed from those both in the middle WM factor tertile ($r = 0.261$, $P = 0.005$; 95% CI = 0.084–0.428) and in the third WM factor tertile representing highest FA ($r = 0.392$, $P < 0.001$; 95% CI = 0.265–0.497).

Next, because the learning slope and intercept factors were not significantly related to the total sample, we inquired whether this association also might jointly depend on HC and WM characteristics. We evaluated differences in the correlation between the slope and intercept factors across the different levels of WM, separately for each tertile of the HC factor

Table 1 Differences in explained variance with and without latent interaction

Model	Original model			Latent moderation model				Difference
	β_{YX1}	β_{YX2}	R^2	β_{YX1}	β_{YX2}	β_{X1X2}	R^2	ΔR^2
No covariates	0.452	0.290	0.074	0.372	0.293	0.486	0.110	0.036
Age	0.426	0.030	0.038	0.358	0.048	0.479	0.077	0.040
Age, sex	0.416	0.019	0.036	0.297	−0.050	0.461	0.060	0.025
Age, sex, and education	0.391	0.024	0.032	0.227	−0.273	0.483	0.050	0.017

Note. R^2 and ΔR^2 (i.e., change in R^2) were calculated using the formula provided by Maslowsky et al. (2015). β_{YX1}: SEM model parameter for regression path to verbal learning slope latent factor from latent factor WM (X1). β_{YX2}: SEM model parameter for regression path to verbal learning slope factor from latent factor HC (X2). β_{X1X2}: Covariance between the factors for HC and WM. R^2 values reflect only variance in verbal learning slope factor explained by the two latent factors WM (X1) and HC (X2), and by their latent interaction. ΔR^2: difference in R^2 values with and without latent interaction. In covariate models, R^2 estimates reflect the inclusion of age, sex, and education.

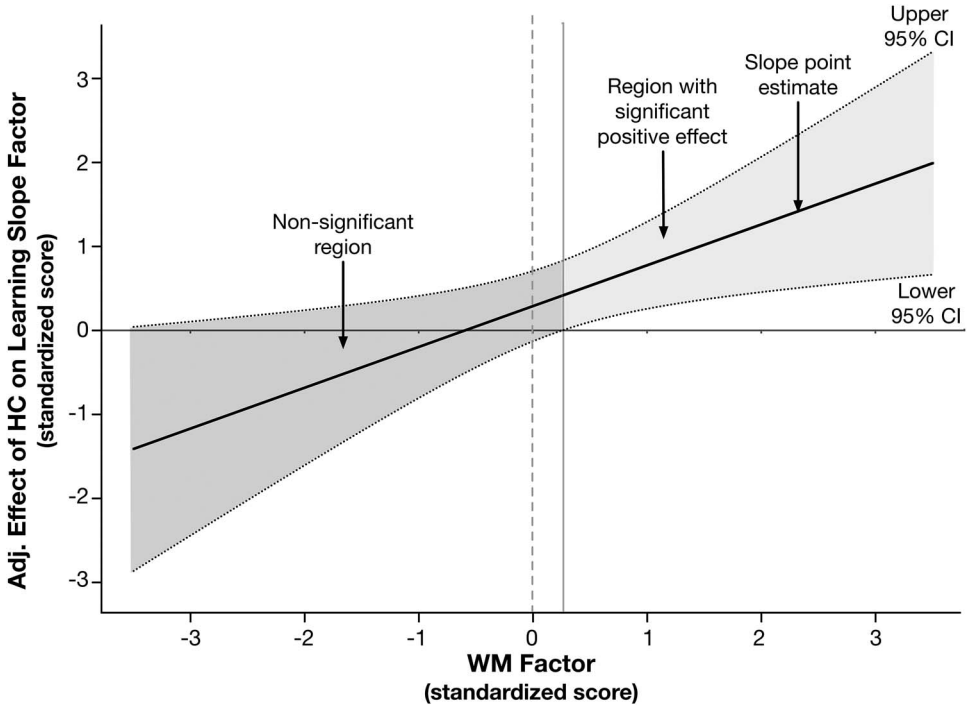

Figure 3. Johnson-Neyman plot illustrating the decomposed interaction to test the moderating effect of limbic WM FA on the effect of HC volume on the verbal learning slope factor. The *x*-axis represents the continuous moderator—here, the standardized WM factor score, and the y-axis represents the effect of the hippocampus (HC) latent factor in the latent interaction on the verbal learning slope parameter, adjusted for other model parameters. The solid regression line reflects the association between the adjusted effect of the HC factor on the learning slope factor, as a function of level in the WM factor. The dotted lines represent the upper and lower 95% confidence band around the regression slope. The solid horizontal line at $y = 0$, and the dotted vertical line at $x = 0$ are superimposed to assist with interpretation. Regions where the confidence bands overlap with $y = 0$ indicates the levels of the *x*-variable in which the effect represented by the regression slope are not significant; this is denoted by dark gray shading. The confidence bands overlap with zero until the WM factor score is slightly greater than 0.15, demonstrating that the adjusted effect of HC volume on learning is only apparent at nonnegative values of the WM factor (i.e., area with lighter gray shading).

distribution. Among participants in the tertile for largest HC, higher learning intercept was associated with less positive slope in those with the highest FA ($r = -0.248$, $P = 0.154$; 95% CI = −0.481 to −0.020); however, this relationship was positive in both the middle ($r = 0.522$, $P = 0.001$; 95% CI = 0.251–0.718) and lower WM tertiles ($r = 0.261$, $P = 0.131$; 95% CI = −0.001 to 0.485). Thus, whereas higher initial recall performance left less room for improvement across learning trials in those with the most robust brain parameters, intercept served as a positive correlate of learning in participants whose brain parameter estimates were near or below the sample mean.

Covariate Models

Although the LMS model without covariates (i.e., the 'No Covariates' model in Table 1) provided the primary findings of interest, we repeated the modeling process to assess how these effects are influenced by three relevant demographic covariates: years of age and educational attainment, both centered at their respective sample means, and participant sex. Initially, we tested the inclusion of covariates in the combined model by specifying paths from each covariate to each of the nine latent factors for HC subfields, WM tracts, and the intercept and slope, which proved an acceptable fit (Supplementary Table SM2). In

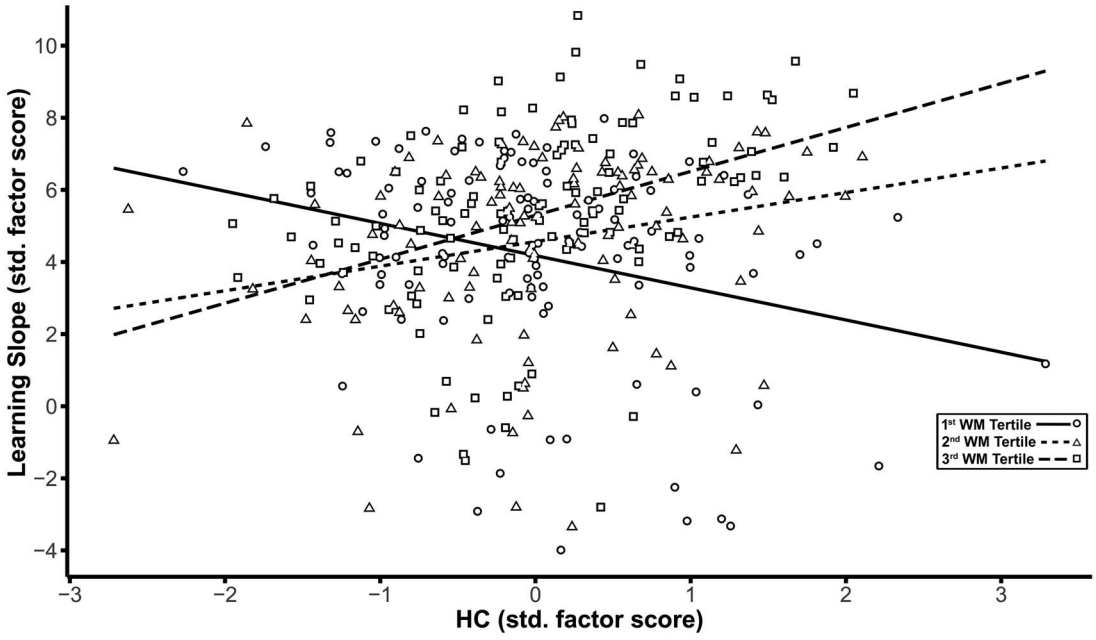

Figure 4. Decomposition of the effects of the latent interaction between hippocampus (HC) and WM on the latent factor representing the slope across learning trials based on WM tertiles. The scatter plot shows the HC factor score on the *x*-axis plotted against the learning slope factor on the *y*-axis, with linear smoothers fitted separately for each of the three tertiles of the WM distribution. Scales for both axes are depicted using standardized scores. Separate symbols and fitted regression lines represent each of the three tertiles of the WM distribution representing low, middle, and high FA values. Greater HC volume is associated with higher learning slope only in the middle (short-dashed regression line and triangle symbols) and highest tertiles (long-dashed regression line and square symbols) of the WM factor. For the lowest tertile of WM (solid regression line and circle symbols), higher HC volume is associated with lower learning rate.

addition, covariances between learning slope and WM brain parameters (i.e., FA in fornix and CBH) remained significant in the combined model, with the inclusion of covariates (for both, $P < 0.05$).

Next, we respecified the second-order models to include the paths from the three covariates to the latent factors for WM, HC, and learning slope. Model fit was acceptable across all second-order covariate models (Supplementary Table SM5). In the second-order model with all three covariates, the path from the WM factor to learning slope remained significant (estimate = 0.172, $P = 0.013$; 95% CI = 0.055–0.765). However, following estimation of the LMS covariate model to test the latent interaction between WM and HC factors on learning, the direct effect from WM to learning rate was no longer significant (estimate = 0.227, $P = 0.256$; 95% CI = −0.165 to 0.620). Of note, however, the path from the HC × WM latent interaction term to the learning rate factor remained significant (estimate = 0.483, $P = 0.013$; 95% CI = 0.102–0.864). The LMS covariate model also showed that older age was associated with lower learning slope (estimate = −0.530, $P = 0.011$; 95% CI = −0.937 to −0.124) and with smaller HC (estimate = −0.267, $P < 0.001$; 95% CI = −0.399 to −0.135). Model results also revealed a significant effect of sex on the learning intercept factor (estimate = 0.592, $P = 0.011$; 95% CI = 0.137–1.046) showing superior recall performance by women over men.

Next, we evaluated the differences in levels of covariates across the tertiles of the WM and HC factor distributions. Separately, for each of the three HC factor tertiles, we evaluated one-way ANOVAs with age as the dependent variable and WM tertile as the independent variable. The model for the lowest tertile of the HC factor yielded an effect of WM tertile on age, $F(1,110) = 4.837$, $P = 0.030$, which was rendered nonsignificant following Bonferroni correction for multiple comparisons across the three models. We repeated this process to evaluate differences in educational attainment. The ANOVA revealed a significant effect of WM tertile on education only for the second tertile of HC, $F(1,110) = 9.564$, $P = 0.003$. *Post hoc* Student's *t*-tests showed that among the participants in the middle tertile of the HC factor, those with lowest WM factor scores had fewer self-reported years of formal education than those with the highest limbic FA: $t(72) = -3.060$, $P = 0.003$.

Discussion

We used latent factor modeling to evaluate the relationships between multiple limbic structures and learning in a large, population-based cohort study of older adults. The present study yielded several notable results concerning associations between limbic WM microstructure, HC subfield volumes, and verbal learning. First, a latent factor formed from FA in limbic WM regions and uncinate was consistently associated with faster rate of learning. Moreover, the latent factor representing volume of the hippocampus was not significantly related to learning rate in the total sample. However, evaluating the latent interaction between HC and WM factors revealed an important moderation effect: HC volume was only positively related to learning rate in older adults with more coherent diffusion in limbic WM, possibly reflecting more intact WM microstructure. In contrast, larger HC volume was associated with lower learning rate for individuals with lower WM anisotropy. This has substantial implications for the use of HC volume as a biomarker of brain and cognitive aging.

Van Petten's (2004) meta-analysis of the relationship between HC volume and memory notes substantial heterogeneity in this

association among older adults. The present findings offer one possible explanation for some of this variation. Indeed, Van Petten notes that data from one study of population neuroimaging supported a weak but significant association between total HC volume, and immediate and delayed verbal recall (Hackert et al. 2002), with an age-residualized effect of $r = 0.12$. However, that review found that smaller and more selectively sampled study cohorts were often more likely to report the positive association between HC volume and memory in older adults. In comparison, the present population-based cohort study of aging more closely resembles the Rotterdam Study, in which the age-residualized effects of HC volume on memory were rather modest. Thus, population neuroimaging studies that include less selectively screened samples should evaluate the associations between HC volume and memory as conditioned on differences in WM microstructure.

Learning rate has been previously associated with total HC volume in older adults with and without memory impairments (Bonner-Jackson et al. 2015). However, this is the first study to link differences in limbic WM with learning, modeled as a growth function. HC afferent and efferent pathways via the fornix and cingulum play a crucial role in mnemonic encoding and recall (Aggleton and Brown 1999), underlining the need to examine structural connections beyond the HC (Aggleton 2014). Prior reports evaluating the combined associations of limbic WM diffusion parameters and HC volumes on episodic memory show mixed effects. Whereas higher FA in CBH and fornix has been linked with better episodic memory, the relationships of total HC volume are inconsistent (Ezzati et al. 2015; Metzler-Baddeley et al. 2011a). We found that higher FA in the ventral (i.e., parahippocampal) portion of the cingulum bundle and the fornix was consistently associated with higher learning rate. Moreover, the association between HC volume and learning rate was positive in individuals with higher FA in limbic WM; however, this relationship was negative in those with low limbic FA.

These results support the notion that the intercept and slope of learning may reflect different demographic factors, such as age, sex, and education, as well as differences in other cognitive abilities, including verbal knowledge, processing speed, and cognitive status (Jones et al. 2005). One possibility is that the different patterns of WM and HC reflect different genetic and life course influences. We found that higher educational attainment was associated with more coherent limbic WM microstructure in those in the middle tertile of HC. However, whether this might also serve as an indicator of risk for subsequent decline will require further analysis with longitudinal data. Future studies might also benefit from applying nonparametric approaches to identify nonlinear moderation patterns, such as local SEM (Hildebrandt et al. 2009; Hülür et al. 2011) or SEM trees (Brandmaier et al. 2013). For example, local SEM could be used to move a window over all participants sorted by WM factor values and then plot estimated model parameters over WM factor values.

The present findings also highlight the utility of SEM latent factor approaches for modeling relationships between multiple neural correlates and cognitive measures as latent factors, free from inherent measurement error. This also permits simultaneous estimation of associations between related factors while precluding the need to correct for multiple comparisons. To the best of our knowledge, this is the first time that the hippocampus has been modeled in this fashion—as a second-order latent factor formed by individual subregional factors. Such an approach may provide a more reliable volumetric estimate of HC structural integrity, particularly in comparison to age-biased estimates of single volumetric indicators from automated segmentation procedures (Wenger et al. 2014).

Furthermore, specifying the latent interaction between the HC and WM latent factors resulted in a better model fit and explained more variance in learning rate. There are a limited number of valid statistical approaches for demonstrating such differential patterns of relationships in cross-sectional data. Mediational approaches are sometimes used to model more complex relationships between brain regions, age, and cognition (Foster et al. 2019; Metzler-Baddeley et al. 2019; Salthouse 2011). Despite violating essential assumptions of temporal ordering necessary to test causal relationships, this nevertheless points to an important modeling need—showing that associations between two variables vary across levels of a third. Moderation approaches are more appropriate for these types of cross-sectional data, and as we show, it can illuminate new patterns of brain-cognition relations in the population.

Limitations and Directions for Future Work

The results of the present study need to be interpreted in light of its limitations. First, the present data are cross-sectional and hence cannot reveal the order and directionality of age-related changes (Lindenberger et al. 2011). Second, we chose to include participants with MMSE scores of 25 and 26, raising the possibility that a small number of participants may have been in the process of developing dementia. There were also several notable technical limitations. First, higher b-values and multishell dMRI data can improve the resolution of crossing fibers and we recommend their use in future studies. Second, we used the aggregate values of WM parameters across tracts of interest, which does not permit more specific anatomical localization of possible effects in cerebral WM. Future studies should try to discern whether specific tract segments are differentially associated with learning and memory (Colby et al. 2012). As the number of brain variables of interest grows (e.g., many ROIs, or even voxel-level analyses), one may consider statistical approaches that appropriately deal with situations with the large number of predictors and relatively small sample sizes, such as regularization (Jacobucci et al. 2019). Also, the 2-mm slice thickness associated with the high-resolution structural imaging sequence for HC subfield volumetry used in this study may have come with cost of inducing greater partial volume artifacts. In addition, HC subfield measurement was limited to the body. Although some published methods permit the segmentation of the head and tail, this may simply introduce further methodological heterogeneity (Yushkevich et al. 2015). Work currently in progress should help to extend the valid segmentation of HC subfields to head and tail of the HC using a harmonized protocol (Wisse et al. 2017).

Last, there are also assumptions and limitations associated with specifying interactions in latent space (Moosbrugger et al. 1997). One concern is that established estimation methods impose potentially problematic assumptions regarding the orthogonality of error structures (Little et al. 2006). However, most published work with such SEM approaches for testing latent factor moderation specify exogenous latent factors based on unreliable observed variables. Here, we used a second-order factor, and although this may be a viable method for circumventing such concerns, such an approach has not been compared before with other latent moderation approaches. Thus, further work is needed to establish better practices for estimating interactions between continuous factors. Moreover, future studies

should also compare changes in WM and HC measures as the correlates of longitudinal changes in learning (Bender and Raz 2015; Bender et al. 2016). It is unclear why the paths from both WM and HC factors to the learning slope were attenuated following the inclusion of the age covariate, but that their interaction was not. Further work is needed to investigate the possibly differentially age-related mechanisms that underlie HC and WM and their interaction.

Conclusion

In the present study, we delineated multimodal neural correlates of verbal learning in older adults, including specific limbic WM fiber tracts and HC subregions. We show that HC volumetric associations with verbal learning are dependent on the levels of FA in limbic WM fiber tracts. Given that the present sample was unimpaired and did not widely differ in age, we consider this result as encouraging (cf. Salthouse, 2011) while recognizing that it needs to be replicated and extended in future cross-sectional and longitudinal investigations. These findings also suggest future studies should account for differences in WM microstructure when considering total HC volume as a correlate of learning and memory in older adults.

Supplementary Material

Supplementary material is available at *Cerebral Cortex* online.

Funding

Strategic Innovation fund of the Max Planck Society (grant 67–11HIPPOC to U.L.); National Institutes of Health grants (grants P41EB015902, R01AG042512 to O.P.);German Science Foundation (grants DFG KU 3322/1-1, SFB 936/C7 to S.K.). This project is also part of the BMBF funded EnergI consortium (01GQ1421B).

Notes

We thank Michael Krause for technical assistance in this study, Dhaval Gandhi for assistance with manuscript preparation and Ahnalee Brinks for suggestions regarding data visualization.
Conflict of Interest: None declared.

References

Aggleton JP. 2014. Looking beyond the hippocampus: old and new neurological targets for understanding memory disorders. *Proc Biol Sci*. 281:20140565.

Aggleton JP, Brown MW. 1999. Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behav Brain Sci*. 22:425–444.

Baldo JV, Delis D, Kramer J, Shimamura AP. 2002. Memory performance on the California verbal learning test-II: findings from patients with focal frontal lesions. *JINS*. 8:539–546.

Beginner DWI Tutorial. 2017. Retrieved fromhttp://mrtrix.readthedocs.io/en/latest/getting_started/beginner_dwi_tutorial.html

Bender AR, Daugherty AM, Raz N. 2013. Vascular risk moderates associations between hippocampal subfield volumes and memory. *J Cogn Neurosci*. 25:1851–1862.

Bender AR, Keresztes A, Bodammer NC, Shing YL, Werkle-Bergner M, Daugherty AM, Yu Q, Kühn S, Lindenberger U, Raz N. 2018. Optimization and validation of automated hippocampal subfield segmentation across the lifespan. *Hum Brain Mapp*. 39:916–931.

Bender AR, Prindle JJ, Brandmaier AM, Raz N. 2016. White matter and memory in healthy adults: coupled changes over two years. *NeuroImage*. 131:193–204.

Bender AR, Raz N. 2015. Normal-appearing cerebral white matter in healthy adults: mean change over two years and individual differences in change. *Neurobiol Aging*. 36:1834–1848.

Bennett IJ, Huffman DJ, Stark CE. 2014. Limbic tract integrity contributes to pattern separation performance across the lifespan. *Cereb Cortex*. 25:2988–2999.

Bentler PM. 1990. Comparative fit indexes in structural models. *Psychol Bull*. 107:238.

Berlin KS, Parra GR, Williams NA. 2013. An introduction to latent variable mixture modeling (part 2): longitudinal latent class growth analysis and growth mixture models. *J Pediatr Psychol*.39(2):188–203.

Bertram L, Bockenhoff A, Demuth I, Düzel S, Eckardt R, Li SC, Lindenberger U, Pawelec G, Siedler T, Wagner GG *et al*. 2014. Cohort profile: the Berlin aging study II (BASE-II). *Int J Epidemiol*. 43:703–712.

Bonner-Jackson A, Mahmoud S, Miller J, Banks SJ. 2015. Verbal and non-verbal memory and hippocampal volumes in a memory clinic population. *Alzheimer Res Therapy*. 7:61.

Braak H, Braak E, Yilmazer D, Bohl J. 1996. Functional anatomy of human hippocampal formation and related structures. *J Child Neurol*. 11:265–275.

Brandmaier AM, von Oertzen T, McArdle JJ, Lindenberger U. 2013. Structural equation model trees. *Psychol Methods*. 18:71–86.

Charlton RA, Barrick TR, Markus HS, Morris RG. 2013. Verbal working and long-term episodic memory associations with white matter microstructure in normal aging investigated using tract-based spatial statistics. *Psychol Aging*. 28:768–777.

ClavelF. Advanced topics: plotting better interactions using the Johnson-Neyman technique in Mplus [internet]. 2015. Retrieved from: https://clavelresearch.wordpress.com/2015/03/23/advanced-topics-plotting-better-interactions-using-the-johnson-neyman-technique-in-mplus/ (last accessed 3 April 2019).

Colby JB, Soderberg L, Lebel C, Dinov ID, Thompson PM, Sowell ER. 2012. Along-tract statistics allow for enhanced tractography analysis. *NeuroImage*. 59:3227–3242.

Concha L, Gross DW, Beaulieu C. 2005. Diffusion tensor tractography of the limbic system. *AJNR Am J Neuroradiol*. 26:2267–2274.

Daugherty AM, Bender AR, Raz N, Ofen N. 2016. Age differences in hippocampal subfield volumes from childhood to late adulthood. *Hippocampus*. 26:220–228.

Duncan TE, Duncan SC, Strycker LA. 2013. *An introduction to latent variable growth curve modeling: concepts, issues, and application*. New York: Routledge Academic.

Duvernoy HM. 2005. *The human hippocampus: functional anatomy, vascularization, and serial sections with MRI*. 3rd ed. New York: Springer.

Ezzati A, Katz MJ, Lipton ML, Zimmerman ME, Lipton RB. 2015. Hippocampal volume and cingulum bundle fractional anisotropy are independently associated with verbal memory in older adults. *Brain Imaging Behav*. 10:652–659.

Fletcher E, Raman M, Huebner P, Liu A, Mungas D, Carmichael O, DeCarli C. 2013. Loss of fornix white matter volume as a predictor of cognitive impairment in cognitively normal elderly individuals. *JAMA Neurol*. 70:1389–1395.

Folstein MF, Folstein SE, McHugh PR. 1975. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psych Res*. 12:189–198.

Foster CM, Kennedy KM, Hoagey DA, Rodrigue KM. 2019. The role of hippocampal subfield volume and fornix microstructure in episodic memory across the lifespan. *Hippocampus*. 1–18.

Fürst G, Ghisletta P. 2009. Statistical interaction between two continuous (latent) variables. *In 11th Congress of the Swiss Psychological Society, August* (pp. 19–20).

Gerstorf D, Bertram L, Lindenberger U, Pawelec G, Demuth I, Steinhagen-Thiessen E, Wagner GG. 2016. The Berlin aging study II–an overview. *Gerontology*. 62:311–315.

Gifford KA, Phillips JS, Samuels LR, Lane EM, Bell SP, Liu D, Alzheimer's Disease NeuroimagingInitiative. 2015. Associations between verbal learning slope and neuroimaging markers across the cognitive aging spectrum. *J Int Neuropsychol Soc*. 21:455–467.

Hackert VH, den Heijer T, Oudkerk M, Koudstaal PJ, Hofman A, Breteler MM. Hippocampal head size associated with verbal memory performance in nondemented elderly. 2002. *NeuroImage*. 17:1365–1372.

Helmstaedter C, Durwen H. 1989. The verbal learning and retention test. A useful and differentiated tool in evaluating verbal memory performance. *Schweiz Arch Neurol Psychiatr*. 1985(141):21–30.

Henson RN, Campbell KL, Davis SW, Taylor JR, Emery T, Erzinclioglu S, Kievit RA. 2016. Multiple determinants of lifespan memory differences. *Sci Rep*. 6:32527.

Hildebrandt A, Wilhelm O, Robitzsch A. 2009. Complementary and competing factor analytic approaches for the investigation of measurement invariance. *Rev Psych*. 16:87–102.

Hooper D, Coughlan J, Mullen MR. 2008. Structural equation modelling: Guidelines for determining model fit. *Journal of Business Research Methods*. 6:53–60.

Hülür G, Wilhelm O, Robitzsch A. 2011. Intelligence differentiation in early childhood. *J Individ Differ*. 32:170–179.

Insausti R, Insausti AM, Sobreviela MT, Salinas A, Martinez-Penuela JM. 1998. Human medial temporal lobe in aging: anatomical basis of memory preservation. *Microsc Res Tech*. 43:8–15.

Jacobucci R, Brandmaier AM, Kievit RA. 2019. A practical guide to variable selection in structural equation modeling by using regularized multiple-indicators, multiple-causes models. *Adv Methods Pract Psychol Sci*. 2:55–76.

Jack CR Jr, Twomey CK, Zinsmeister AR, Sharbrough FW, Petersen RC, Cascino GD. 1989. Anterior temporal lobes and hippocampal formations: normative volumetric measurements from MR images in young adults. *Radiology*. 172: 549–554.

Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. 2012. Fsl. *NeuroImage*. 62:782–790.

JenkinsonM, PechaudM, SmithS. 2005. BET2: MR-based estimation of brain, skull and scalp surfaces. In:*Eleventh Annual Meeting of the Organization for Human Brain Mapping*;vol.17, p. 167.

Johnson PO, Neyman J. 1936. Tests of certain linear hypotheses and their application to some educational problems. *Statistical Research Memoirs*. 1:57–93.

Jones DK, Cercignani M. 2010. Twenty-five pitfalls in the analysis of diffusion MRI data. *NMR\ignorespacesBiomed.* 23: 803–820.

Jones RN, Rosenberg AL, Morris JN, Allaire JC, McCoy KJ, Marsiske M *et al*. 2005. A growth curve model of learning acquisition among cognitively normal older adults. *Exp Aging Res*. 31:291–312.

Kausler DH. 1994. *Learning and memory in normal aging*. San Diego (CA): Academic Press.

Keihaninejad S, Heckemann RA, Fagiolo G, Symms MR, Hajnal JV, Hammers A. 2010. A robust method to estimate the intracranial volume across MRI field strengths (1.5T & 3T). *NeuroImage*. 50:1427–1437.

Keresztes A, Bender AR, Bodammer NC, Lindenberger U, Shing YL, Werkle-Bergner M. 2017. Hippocampal maturity promotes memory distinctiveness in childhood and adolescence. *Proc Natl Acad Sci U S A*. 114:9212–9217.

Kiernan JA. 2012. Anatomy of the temporal lobe. *Epilepsy Res Treat*. 2012:176157.

Korchin SJ, Basowitz H. 1957. Age differences in verbal learning. *J Abnorm Psychol*. 54:64–69.

Lindenberger U, von Oertzen T, Ghisletta P, Hertzog C. 2011. Cross-sectional age variance extraction: what's change got to do with it? *Psychol Aging*. 26:34–47.

Little TD, Bovaird JA, Widaman KF. 2006. On the merits of orthogonalizing powered and product terms: implications for modeling interactions among latent variables. *Struct Equ Modeling*. 13:497–519.

Madden DJ, Bennett IJ, Burzynska A, Potter GG, Chen NK, Song AW. 2012. Diffusion tensor imaging of cerebral white matter integrity in cognitive aging. *Biochim Biophys Acta*. 1822:386–400.

Malykhin N, Concha L, Seres P, Beaulieu C, Coupland NJ. 2008. Diffusion tensor imaging tractography and reliability analysis for limbic and paralimbic white matter tracts. *Psychiatry Res*. 164:132–142.

Maslowsky J, Jager J, Hemken D. 2015. Estimating and interpreting latent variable interactions: A tutorial for applying the latent moderated structural equations method. *International Journal of Behavioral Development*. 39: 87–96.

McArdle JJ. 1986. Latent variable growth within behavior genetic models. *Behav Genet*. 16:163–200.

Meredith W, Tisak J. 1990. Latent curve analysis. *Psychometrika*. 55:107–122.

Metzler-Baddeley C, Jones DK, Belaroussi B, Aggleton JP, O'Sullivan MJ. 2011a. Frontotemporal connections in episodic memory and aging: a diffusion MRI tractography study. *J Neurosci*. 31:13236–13245.

Metzler-Baddeley C, Mole JP, Sims R, Fasano F, Evans J, Jones DK, Aggleton JP, Baddeley RJ. 2019. Fornix white matter glia damage causes hippocampal gray matter damage during age-dependent limbic decline. *Sci Rep*. 9: 1060–1074.

Metzler-Baddeley C, O'Sullivan MJ, Bells S, Pasternak O, Jones DK. 2012. How and how not to correct for CSF-contamination in diffusion MRI. *NeuroImage*. 59:1394–1403

Moosbrugger H, Schermelleh-Engel K, Klein A. 1997. Methodological problems of estimating latent interaction effects. *Methods Psychol Res Online*. 2:95–111.

Mueller SG, Chao LL, Berman B, Weiner MW. 2011. Evidence for functional specialization of hippocampal subfields detected by MR subfield volumetry on high resolution images at 4T. *NeuroImage*. 56:851–857.

Mueller SG, Stables L, Du AT, Schuff N, Truran D, Cashdollar N, Weiner MW. 2007. Measurement of hippocampal subfields and age-related changes with high resolution MRI at 4 T. *Neurobiol Aging*. 28:719–726.

Muthén LK, Muthén BO. 2012. *Mplus User's Guide*. 7th ed. Los Angeles (CA): Muthén & Muthén.

Oguz I, Farzinfar M, Matsui J, Budin F, Liu Z, Gerig G, Styner M. 2014. DTIPrep: quality control of diffusion-weighted images. *Front Neuroinform*. 8:4.

Pasternak O, Sochen N, Gur Y, Intrator N, Assaf Y. 2009. Free water elimination and mapping from diffusion MRI. *Magn Reson Med*. 62:717–730.

Petersen RC, Jack CR Jr, Xu YC, Waring SC, O'Brien PC, Smith GE, Ivnik RJ *et al*. 2000. Memory and MRI-based hippocampal volumes in aging and AD. *Neurology*. 54:581–587.

Preacher KJ, Curran PJ, Bauer DJ. 2006. Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *J Educ Behav Stat*. 31:437–448.

Preston AR, Eichenbaum H. 2013. Interplay of hippocampus and prefrontal cortex in memory. *Curr Biol*. 23:R764–R773.

Raz N, Lindenberger U, Rodrigue KM, Kennedy KM, Head D, Williamson A *et al*. 2005. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cereb Cortex*. 15:1676–1689.

Sasson E, Doniger GM, Pasternak O, Tarrasch R, Assaf Y. 2013. White matter correlates of cognitive domains in normal aging with diffusion tensor imaging. *Front Neurosci*. 7:32.

Salthouse TA. 2011. Neuroanatomical substrates of age-related cognitive decline. *Psychol Bull*. 137:753.

Schmidt M. 1996. *Rey auditory verbal learning test: a handbook*. Los Angeles (CA): Western Psychological Services.

Sepulcre J, Masdeu JC, Sastre-Garriga J, Goñi J, Vélez-de-Mendizábal N, Duque B, Pastor MA, Bejarano B, Villoslada P. 2008. Mapping the brain pathways of declarative verbal memory: Evidence from white matter lesions in the living human brain. *NeuroImage*. 42:1237–43.

Shing YL, Rodrigue KM, Kennedy KM, Fandakova Y, Bodammer N, Werkle-Bergner M, Lindenberger U, Raz N. 2011. Hippocampal subfield volumes: age, vascular risk, and correlation with associative memory. *Front Aging Neurosci*. 3:2.

Shing YL, Werkle-Bergner M, Li SC, Lindenberger U. 2008. Associative and strategic components of episodic memory: a lifespan dissociation. *J Exp Psychol Gen*. 137:495.

Smith SM. 2002. Fast robust automated brain extraction. *Hum Brain Mapp*. 17:143–155.

Stoub TR, deToledo-Morrell L, Stebbins GT, Leurgans S, Bennett DA, Shah RC. 2006. Hippocampal disconnection contributes to memory dysfunction in individuals at risk for Alzheimer's disease. *Proc Natl Acad Sci U S A*. 103:10041–10045.

Tournier JD, Calamante F, Gadian DG, Connelly A. 2004. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*. 23:1176–1185.

Tournier JD, Calamante F, Connelly A. 2013. Determination of the appropriate b-value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. *NMR Biomed*. 26:1775–1786.

Tournier JD, Calamante F, Connelly A. 2007. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*. 35:1459–1472.

Van Petten C. 2004. Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: review and meta-analysis. *Neuropsychologia*. 42:1394–1413.

Wenger E, Mårtensson J, Noack H, Bodammer NC, Kühn S, Schaefer S, Heinze HJ, Düzel E, Bäckman L, Lindenberger U *et al*. 2014. Comparing manual and automatic segmentation of hippocampal volumes: reliability and validity issues in younger and older brains. *Hum Brain Mapp*. 35: 4236–4248.

Wilson IA, Gallagher M, Eichenbaum H, Tanila H. 2006. Neurocognitive aging: prior memories hinder new hippocampal encoding. *Trends Neurosci*. 29:662–670.

Wisse LE, Daugherty AM, Olsen RK, Berron D, Carr VA, Stark CE, Amaral RS, Amunts K, Augustinack JC, Bender AR. 2017. A harmonized segmentation protocol for hippocampal and parahippocampalsubregions: why do we need one and what are the key goals? *Hippocampus*. 27:3–11.

Yushkevich PA, Amaral RSC, Augustinack JC, Bender AR, Bernstein JD, Boccardi M, Bocchetta M *et al*. 2015. Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampalsubregions in in vivo MRI: towards a harmonized segmentation protocol. *NeuroImage*. 111:526–541.

Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, Gerig G. 2006. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage*. 31:1116–1128.

Yushkevich PA, Pluta JB, Wang H, Xie L, Ding SL, Gertje EC, Mancuso L, Kliot D, Das SR, Wolk DA. 2015. Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. *Hum Brain Mapp*. 36:258–287.

Yushkevich PA, Wang H, Pluta J, Das SR, Craige C, Avants BB, Weiner MW, Mueller S. 2010. Nearly automatic segmentation of hippocampal subfields in in vivo focal T2-weighted MRI. *NeuroImage*. 53:1208–1224.

Zhang H, Yushkevich PA, Alexander DC, Gee JC. 2006. Deformable registration of diffusion tensor MR images with explicit orientation optimization. *Med Image Anal*. 10:764–785.