

ORIGINAL ARTICLE

Hippocampal and Parahippocampal Gray Matter Structural Integrity Assessed by Multimodal Imaging Is Associated with Episodic Memory in Old Age

Ylva Köhncke¹, Sandra Düzel¹, Myriam C. Sander¹,
Ulman Lindenberger^{1,2}, Simone Kühn^{3,4} and
Andreas M. Brandmaier^{1,2}

¹Center for Lifespan Psychology, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany, ²Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany, ³Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany and ⁴Lise Meitner Group for Environmental Neuroscience, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany

Address correspondence to Ylva Köhncke, Center for Lifespan Psychology, Max Planck Institute for Human Development, Lentzeallee 94, D-14195 Berlin, Germany. Tel: +49 30 82406 683; Fax: +49 30 82406 571; Email: koehncke@mpib-berlin.mpg.de.

Simone Kühn and Andreas M. Brandmaier take joint responsibility for senior authorship.

Abstract

Maintained structural integrity of hippocampal and cortical gray matter may explain why some older adults show rather preserved episodic memory. However, viable measurement models for estimating individual differences in gray matter structural integrity are lacking; instead, findings rely on fallible single indicators of integrity. Here, we introduce multitrait-multimethod methodology to capture individual differences in gray matter integrity, based on multimodal structural imaging in a large sample of 1522 healthy adults aged 60–88 years from the Berlin Aging Study II, including 333 participants who underwent magnetic resonance imaging. Structural integrity factors expressed the common variance of voxel-based morphometry, mean diffusivity, and magnetization transfer ratio for each of four regions of interest: hippocampus, parahippocampal gyrus, prefrontal cortex, and precuneus. Except for precuneus, the integrity factors correlated with episodic memory. Associations with hippocampal and parahippocampal integrity persisted after controlling for age, sex, and education. Our results support the proposition that episodic memory ability in old age benefits from maintained structural integrity of hippocampus and parahippocampal gyrus. Exploratory follow-up analyses on sex differences showed that this effect is restricted to men. Multimodal factors of structural brain integrity might help to improve our biological understanding of human memory aging.

Key words: episodic memory, healthy aging, multitrait-multimethod model, structural equation modeling

Introduction

Performance in episodic memory tasks typically declines after the age of 60 years (Schaie et al. 1998; Rönnlund et al. 2005), but there are pronounced age-related individual differences in levels and changes of performance (de Frias et al. 2007; Josefsson et al. 2012), with some older individuals displaying little or no performance decline. The “brain maintenance” hypothesis suggests that an older person’s level of behavioral performance reflects the degree to which this person’s brain has maintained its integrity across a variety of levels, including structure, function, and neurochemistry (Nyberg et al. 2012; Lindenberger 2014; Cabeza et al. 2018; Nyberg and Pudas 2019; Nyberg and Lindenberger 2020). Modern neuroimaging techniques allow us to better describe and understand various characteristics of brain tissue through the application of different imaging modalities such as structural and functional magnetic resonance imaging (MRI), diffusion tensor imaging (DTI), or positron-emission tomography (PET). Yet, it is unclear whether these measures converge on constructs that reflect the “integrity” of a given brain region.

Here, we combine multimodal imaging with multitrait-multimethod (MTMM) modeling (Campbell and Fiske 1959; Eid et al. 2008) to represent the gray matter structural integrity of different regions of the human brain, and to investigate their associations with episodic memory. We selected a number of regions of interest (ROIs) that are part of the episodic memory network (for a review, see Dickerson and Eichenbaum 2009; Benoit and Schacter 2015). Specifically, we included hippocampus, parahippocampal gyrus, precuneus, dorsolateral prefrontal cortex, and medio-orbitofrontal cortex as ROIs. The hippocampus plays a key role in episodic memory (Eichenbaum 2017). Hippocampal volume is, on average, smaller in healthy older than in healthy younger adults, and shrinks with time in normal aging (Raz 2005; Raz et al. 2005; Fjell et al. 2009; Walhovd et al. 2011; Jäncke et al. 2020). Smaller hippocampal volume is related to poorer episodic memory performance in cross-sectional studies in old age (for reviews, see Squire 1992 or Kaup et al. 2011; Ward et al. 2015). Longitudinal studies show that less decline over time in hippocampal volume in older adults is related to less decline in episodic memory performance (Persson et al. 2012; Gorbach et al. 2017). Similar results were found at the functional level, such that smaller decrements in activation were associated with better preservation of memory performance (Persson et al. 2012).

Regarding the role of hippocampal microstructural integrity in age-related cognitive decline, there is some indication that higher mean diffusivity (MD) in hippocampus, indicating a less dense tissue structure, is related to poorer episodic memory performance in older adults (Carlesimo et al. 2010). In studies using magnetization transfer (MT) imaging, it could be shown that a higher MT ratio, indicating denser microstructure, is related to lower MD (Düzel et al. 2010), faster processing speed, and higher fluid intelligence (Aribisala et al. 2014), but not better memory (Düzel et al. 2010, 2008; Aribisala et al. 2014). Still, taken together, these findings suggest that the macro- and micro-structural integrity of hippocampus might be critical for preserving its functionality for episodic memory in older age.

Hippocampus does not operate in isolation (Rugg et al. 2008; Dickerson and Eichenbaum 2009; Rugg and Vilberg 2013; Eichenbaum 2017). The parahippocampal gyrus is, together with the entorhinal cortex, a major input source for hippocampus, and critically supports episodic memory (Persson et al. 2012). Parahippocampal volume is smaller in older than in younger

adults (Henson et al. 2016; Gorbach et al. 2017; Foster et al. 2019), and longitudinal decline in parahippocampal gyrus’ volume is related to decline in episodic memory performance (Gorbach et al. 2017), but cross-sectional associations between parahippocampal volume and episodic memory performance are not necessarily observed (Henson et al. 2016; Foster et al. 2019). MD in parahippocampal gyrus is higher in older than in younger adults (Grydeland et al. 2013).

Further cortical regions are known to interact with hippocampus in support of episodic memory, including prefrontal areas, retrosplenial/posterior cingulate cortex, and lateral parietal cortices (e.g., angular gyrus; Rugg and Vilberg 2013). Here, we restrict the analysis to prefrontal cortex and precuneus (as it includes the retrosplenial region), given that some of the structural properties that we examine here have been reported to be associated with memory performance in relation to these regions. To begin with, prefrontal cortex is involved in memory retrieval processes (Eichenbaum 2017). Prefrontal cortex volume is shrinking with advancing age (Raz et al. 2005), its MD is higher in older than in younger individuals (Grydeland et al. 2013), and larger prefrontal gray matter volumes are related to better associative (Becker et al. 2015) as well as item episodic memory performance (Persson et al. 2017). Finally, precuneus contributes to memory retrieval (Cavanna and Trimble 2006), precuneus volume is related to autobiographic memory (Fretton et al. 2014), and MD in precuneus shows age-specific associations with cognitive performance (Grydeland et al. 2013).

“Brain maintenance” should be reflected by preserved tissue integrity on many biological levels, which most likely interact with one another as they change in the course of healthy aging. In structural imaging, brain maintenance should be reflected by relatively sparse and little microstructural damage as well as a relative lack of macrostructural atrophy. Maintenance can thus be assessed by several imaging modalities carried out at the same time in the same subjects. Such a multimodal imaging approach might provide a more comprehensive account of interindividual variability in brain maintenance (Nyberg et al. 2012) than an approach that considers each measure separately. In the present study, we combine different structural characteristics of gray matter tissue into region-specific constructs of structural integrity. Our approach takes advantage of commonalities across measures from different modalities while removing the modality-specific measurement error; the rationale being that all selected measures reflect some aspect of structural integrity, so that their commonality should be a more robust index of structural integrity than any one measure alone.

Our approach differs from approaches adopted in other multimodal imaging studies. There are many good reasons to invest more effort into the use of multimodal brain imaging in aging research (Nevalainen et al. 2015; Fjell and Walhovd 2016). Many of the existing multimodal imaging studies aim at maximizing predictive accuracy by combining information from multiple modalities (Ritchie et al. 2015; Ward et al. 2015; Hedden et al. 2016; Liem et al. 2017). These approaches capitalize on the unique information that each modality adds to predicting cognitive performance (Ritchie et al. 2015; Ward et al. 2015; Hedden et al. 2016) or brain age (Liem et al. 2017), so they benefit from the fact that each modality measures different aspects of integrity. In contrast, in the present study, we were not primarily interested in the unique contribution of each measure to predicting an outcome, but instead in the variance that is shared across modalities. The presumed primary advantage of multimodality in our approach is that the resulting latent factor might yield a

more reliable and valid estimate of the target concept, which is the purported tissue property of “structural integrity.” A latent factor of regional gray matter integrity expresses what the indicators have in common, is free of imaging modality-specific variance, and free of residual variance (measurement error). In comparison to currently available indicators, we postulate that a factor of this sort is more likely to do justice to the level of generality and abstraction that the term integrity suggests.

In this study, we aimed to quantify and jointly model different structural properties of gray matter by using three common imaging techniques, each being differentially sensitive to macro- and micro-structural properties of brain tissue (Bartrés-Faz and Arenaza-Urquijo 2011), namely gray matter volume, MT ratio, and MD from DTI. In the next paragraph, we describe which structural properties of gray matter are captured by the imaging modalities we selected for the current study.

Structural imaging provides static anatomical information derived from MR signal properties. T1-weighted, 3D, high-resolution images are commonly used to estimate the volume of brain ROIs to study interindividual differences in volume and volume changes over time. When using the voxel-based morphometry (VBM) method (Ashburner and Friston 2000; Good et al. 2001), signal intensity in every voxel is used to gauge regional variations in structural properties of the tissue and provides voxel-wise estimations of the local volume of specific tissue compartments (gray matter, white matter, or cerebrospinal fluid).

Microstructural properties of gray matter regions, and age-related differences therein, can be probed by MT imaging (for a review, see Seiler et al. 2014). MT imaging capitalizes on the transfer of energy and related magnetization exchange between mobile water protons and protons that are immobilized by macromolecules (Wolff and Balaban 1989). MT ratio values are calculated as the ratio between values measured with a MT pulse and values without MT pulse. MT can detect subtle microstructural abnormalities due to age-related or pathological changes otherwise not detectable with standard MRI (Seiler et al. 2014). MT ratio values depend on content and concentration of macromolecules bound to water molecules in relation to free water molecules. Lower MT ratio values can result from an increase in the mobile proton pool, occurring as a result of inflammation and edema, or a decrease in the semisolid proton pool, associated with cell damage, axonal loss, and demyelination (Seiler et al. 2014).

DTI can detect subtle changes in cellular microstructure by measuring patterns of water diffusion that likewise cannot be quantified using more traditional structural MRI sequences. MD is a DTI metric that measures the rate of water diffusion in all directions within an image voxel (Pierpaoli and Basser 1996) and is commonly used as an index of white matter microstructural integrity. MD can be used to characterize one form of structural integrity under the assumption that region-specific diffusion is based on (1) less diffusion across cell membranes in denser structures or structures with a main direction as seen in white matter tracts (Sundgren et al. 2004; Jespersen et al. 2007), (2) more diffusion within less dense brain structures or structures with no principal direction as seen in gray matter (Sundgren et al. 2004). Although more often used to characterize white matter, MD is also informative of gray matter microstructural properties and age-related differences in it (Abe et al. 2008; Grydeland et al. 2013), with lower MD indicating a denser structure, most probably indicating more cell membranes and intracortical myelin (Grydeland et al. 2013).

In the present study, we combined macro- and micro-structural imaging modalities as indicators of gray matter integrity in a multimodal approach. Thus, we set out to validate a structural equation model representing the commonalities of specific tissue characteristics resulting from different imaging modalities. To establish the plausibility of integrity factors, we examined whether the empirically observed covariance structure shows substantial commonalities among the various gray-matter indices.

We used cross-sectional data from the older participants of the Berlin Aging Study II (BASE-II; Bertram et al. 2013), which amount to a fairly large sample of 1532 healthy adults aged 60–88 years, with structural brain imaging measures of VBM, MT, and MD taken from a subsample of 333 participants who underwent MR imaging. Our goal of this analysis approach was two-fold. First, we sought to demonstrate the benefits of a multivariate latent variable modeling approach to representing regional structural integrity while doing justice to the complexity of the underlying measurement problem. Second, we wished to demonstrate that such an approach can be put to use to identify the associations between structural properties of gray matter regions belonging to the episodic memory network and episodic memory ability in old age.

In a first set of analyses, we established region-specific latent brain integrity factors. To this end, we specified confirmatory factor models within each of the brain regions by defining a latent brain integrity factor representing the variance that is shared across the three imaging modalities. This latent factor should capture the statistical communality of different physical properties of gray matter tissue. This parallels psychometric approaches targeting a nonobservable (or latent) psychological construct by measuring a range of indicators and interpreting their common variance as representative of the target construct. By using different indicators, we triangulate our target construct, integrity, which is sensible even if our indicators should only have limited overlap in variance (Little et al. 1999).

In a second set of analyses, we combined the latent brain integrity model with a latent episodic memory factor to investigate the associations of brain integrity and episodic memory performance. We hypothesize that gray matter integrity in regions of the episodic memory network is related to episodic memory performance.

Materials and Methods

Participants and Study Design

Healthy older participants were recruited from BASE-II, a multi-institutional and multidisciplinary study assessing variables from a wide range of domains for each participant (Bertram et al. 2013). Participants completed a comprehensive cognitive examination (see Düzel et al. 2016, for further details). A subsample of eligible participants was then invited to take part in a separate MRI session within a couple of weeks (mean time interval 3.8 months, $SD=4.4$) after completing cognitive testing. None of the participants took any medication that is known to potentially affect memory function or had a history of head injuries, medical (e.g., heart attack), neurological (e.g., epilepsy), or psychiatric disorders (e.g., depression). Additionally, all participants had completed at least 8 years of education. BASE-II includes a larger sample of persons above age 60 and a smaller sample of participants in early adulthood. Here,

Table 1 Descriptive statistics of the background variables and episodic memory task scores both in the full sample and the magnetic resonance (MR) sample, and selectivity (in SD) of the MR sample

Variable	N	Mean		SD		Skewness		Kurtosis		Selectivity ^a
		Total sample	MR sample							
Age	1511	70.60	70.08	3.84	3.83	0.31	0.22	3.59	2.99	0.13
Sex (% f)	1522	50.72	38.44							
Education	1339	14.16	14.08	2.89	2.90	0.14	0.13	1.62	1.69	0.03
VLMT	1497	8.50	8.46	2.68	2.52	-0.43	-0.35	2.49	2.66	0.02
FP	1495	0.27	0.27	0.21	0.20	-0.02	0.11	2.89	2.70	-0.04
SE	1495	0.28	0.29	0.14	0.14	-0.02	-0.09	2.93	2.91	-0.10
OL	1505	13.26	13.56	4.00	4.03	0.05	0.10	3.12	3.52	-0.08

Notes: VLMT, verbal learning and memory test—sum of remembered words after five learning trials with the same 15 words; FP, Face-profession task—hits minus false alarms; SE, scene-encoding task—hits minus false alarms; OL, object location task—sum of correct placements across two trials (each consisting of 12 items).

^aSelectivity = $(\text{mean}_{(\text{total sample})} - \text{mean}_{(\text{MR sample})}) / \text{SD}_{(\text{total sample})}$.

Table 2 Descriptive statistics of the MR variables, in the MR sample

Variable	N	Mean	SD	Skewness	Kurtosis
VBM hippocampus	330	0.5218	0.0458	-0.3126	3.8741
VBM parahippocampal gyrus	330	0.4316	0.0317	-0.6077	4.4632
VBM precuneus	330	0.3969	0.0304	-0.037	2.9566
VBM medio-orbitofrontal cortex	330	0.4156	0.0307	-0.0886	2.7687
VBM dorsolateral prefrontal cortex	330	0.4147	0.0297	-0.0652	2.9019
MT hippocampus	197	332.58	31.15	-1.041	5.5332
MT parahippocampal gyrus	197	363.62	14.92	-1.0487	8.5256
MT precuneus	197	335.81	18.81	-0.0541	3.0495
MT medio-orbitofrontal cortex	197	354.50	16.79	0.2575	4.321
MT dorsolateral prefrontal cortex	197	300.51	21.40	0.6943	3.8887
MD hippocampus	274	0.0014	0.000135	0.6734	3.4187
MD parahippocampal gyrus	274	0.0013	0.000111	0.481	3.3893
MD precuneus	274	0.0012	0.000112	0.4331	3.4798
MD medio-orbitofrontal cortex	274	0.0012	0.000102	0.3673	3.276
MD dorsolateral prefrontal cortex	274	0.0012	0.000081	0.4319	3.4547

Notes: N, sample size in the MR subsample; SD, standard deviation; VBM, voxel-based morphometry, MT, magnetization transfer ratio; MD, mean diffusivity (in gray matter).

we selected only data from participants above age 60, of which 1532 had completed cognitive testing, of which 342 had additionally taken part in MR imaging. We had to exclude nine cases with erroneous data from the cognition sample, two of which were in the MR sample as well. We then excluded multivariate outliers with highly unlikely combinations of values ($P < 0.0001$ of robust Mahalanobis distances; detected using R-package faoutlier, version 0.7.2, Chalmers and Flora 2015, method “mve,” in complete cases only). We detected multivariate outliers separately for the 4 episodic memory tests in the total sample ($n = 1523$; one outlier found) and for the 12 MR variables in the MR sample ($n = 340$; seven outliers found). Hence, the effective sample with cognitive data consisted of 1522 older adults (Table 1), the effective sample with MR-data consisted of 333 older adults (Table 2). Thus, of the total sample, for 78.12% we had no MR data. The ethics committee of the Max Planck Institute for Human Development had approved the cognitive testing and the ethics committee of the Deutsche Gesellschaft für Psychologie (DPGs) had approved the imaging study. Participants received monetary compensation for their participation in the cognitive and imaging sessions and

provided informed consent in accordance to the Declaration of Helsinki.

MRI Acquisition

Images were acquired on a Siemens Tim Trio 3T scanner (Erlangen, Germany) using a 32-channel head coil. The T1 images were obtained using a 3D T1-weighted magnetization prepared gradient-echo (MPRAGE) sequence based on the ADNI protocol (www.adni-info.org; repetition time (TR) = 2500 ms; echo time (TE) = 4.77 ms; TI = 1100 ms, acquisition matrix = $256 \times 256 \times 176$, flip angle = 7° ; $1 \times 1 \times 1 \text{ mm}^3$ voxel size). Diffusion-weighted images were obtained with a single-shot diffusion-weighted spin-echo-refocused echo-planar imaging sequence (FOV $218 \times 218 \text{ mm}$; 128×128 matrix interpolated to 256×256 ; TE = 98 ms; TR = 11 000 ms; 73 slices; slice thickness 1.7 mm; b -value 1000 s/mm^2 ; 60 directions). MT ratio images consisting of two volumes were acquired with identical settings (transversal, 256×256 pixels, TE = 5.5 ms, TR = 28 ms, 48 slices, voxel size $1 \times 1 \times 3 \text{ mm}^3$). The first image (MT image) was acquired with a magnetic saturation pulse (1200 Hz off-resonance, 16 ms) and

the second (noMT image) without a magnetic saturation pulse resulting in a proton-density-like image.

MR Preprocessing

Voxel-Based Morphometry

Structural data were processed with Computational Anatomy Toolbox 12 (CAT12, Structural Brain Mapping group, Jena University Hospital), a toolbox that is implemented in Statistical Parametric Mapping (SPM12, Institute of Neurology) for VBM analysis of imaging data. We applied the CAT12 default cross-sectional preprocessing stream, which implements correction of the T1-weighted images for bias-field inhomogeneities, segmentation into gray matter, white matter and CSF, and spatial normalization using the Diffeomorphic Anatomical Registration Through Exponentiated Lie algebra (DARTEL) algorithm. Modulation with Jacobian determinants was applied in order to preserve the volume of a particular tissue within a voxel leading to a measure of volume of gray matter. Gray matter images were used for the current set of analyses and smoothed with a Gaussian kernel of 8 mm (full width at half maximum).

Magnetization Transfer Imaging

The MT ratio (MTR) maps for each subject were calculated on a voxel-by-voxel basis according to the formula $MTR = (noMT - MT) / noMT$. The data were then normalized into MNI space.

Diffusion Tensor Imaging

Diffusion-weighted images were preprocessed using the FSL software package (Smith et al. 2004; Jenkinson et al. 2012), version 5.0. This included corrections of potential head movement and inspection of image quality. The first non diffusion-weighted image of each individual image set was used as a brain mask. The difference in alignment between this initial image and recurrent ones in the sequence was estimated using FMRIB's Linear Image Registration Tool (FLIRT; Jenkinson et al. 2002) and then corrected for by means of re-alignment. The resulting data were then processed via FSL's *dtfit* to fit a diffusion tensor model at each voxel and obtain the MD values. The MNI based maps were produced using the standard TBSS pipeline (Smith et al. 2006).

ROI Extraction and Adjustment for Differences in Intracranial Volume

Based on prior studies on associations between regional gray matter structure and episodic memory (Squire 1992; Pantel et al. 2003; Cavanna and Trimble 2006; Raz and Rodrigue 2006; Becker et al. 2015; Gorbach et al. 2017) as well as functional correlates of episodic memory (Persson et al. 2012; Benoit and Schacter 2015), we extracted mean values of CAT12/VBM, MD, and MT ratio bilaterally from the ROIs hippocampus, medio-orbitofrontal cortex, dorsolateral prefrontal cortex, parahippocampal gyrus, and precuneus, as defined by the automated anatomical labelling (AAL) atlas (Tzourio-Mazoyer et al. 2002). ROI masks were fitted in MNI space after normalization to a standard template.

We used intracranial volume (ICV) to adjust the VBM values for each ROI via the analysis of covariance formula (Raz et al. 2005): $\text{adjusted volume} = \text{raw volume} - b * (\text{ICV} - \text{mean ICV})$, where b is the slope of regression of the appropriate ROI volume on ICV.

Episodic Memory Assessment

All BASE-II participants were invited to 2 cognitive test sessions with an exact interval of 7 days and at the same time of day to avoid circadian confounding effects on performance. Each session lasted about 3.5 h. Participants were tested in groups of 4–6 individuals. Each group was instructed via a standardized session manual. Each task started with a practice trial to ensure that every participant understood the task. Depending on the task, responses were given via button boxes, the computer mouse, or a keyboard.

The cognitive battery of BASE-II covers key cognitive abilities measured by 21 tasks, 4 of which assess aspects of episodic memory and were thus selected for the present study: 1) "Verbal Learning and Memory," assessing free recall of auditorily presented words after each of five learning trials each consisting of the same 15 words (score: sum of remembered words across the responses to five identical learning trials). Participants typed the words they recalled on the keyboard, one by one visible on the screen; (1) "Face-Profession," testing associative recognition memory 5 min after incidental encoding of 45 face-profession pairs. Participants were instructed to judge whether the face matches the profession. During recall, they were presented with 27 old, 9 new, and 18 rearranged pairs and were asked to provide old-new judgments (score: hits minus false alarms); (3) "Scene Encoding," measuring recognition memory of 88 incidentally encoded scenes (task: indoor-outdoor judgment) after a delay of 2.5 h (score: hits minus false alarms); (4) "Object Location," assessing free recall of 12 deliberately encoded object locations in a 6 by 6 locations grid (score: sum of correct placements across two trials). The tasks are described in detail in Düzel et al. (2016) and in the [Supplementary Materials](#).

Statistical Analyses

We used structural equation modeling to investigate the relationship between episodic memory performance and structural gray matter integrity in a multivariate approach, for two reasons. First, it enables us to capture variance shared across three different structural brain-imaging modalities in a latent factor of structural integrity for any given brain region. This makes sense from a theoretical perspective, as we aim to define a statistically plausible index of region-specific gray matter structural integrity. While the three imaging modalities are designed to assess different characteristics of gray matter structure, their shared variance can be interpreted as indicating a common cause for relatively good or relatively poor gray matter structural integrity. By separating the shared variance (i.e., what is common across measures) from the unique, modality-specific variance (i.e., what is specific to the measurement instrument), we hope to acquire a more reliable and valid estimate of integrity. In addition, we defined latent method factors for each modality (VBM, MTR, and MD) that capture common variance within the modality across all regions (capturing what is common to the measurement instrument only but not to the common factor). As a consequence, the residual variance estimates in our model represent variance that is neither shared by all ROIs in a given modality nor shared by all modalities in a given ROI (e.g., measurement error).

Second, we used structural equation modeling to examine whether the region-specific factors of gray matter structural integrity were associated with episodic memory performance.

In this context, a particular virtue of structural equation modeling is that we can model gray matter integrity for each of the ROIs as well as episodic memory performance as latent factors.

We specified and estimated structural equation models in Onyx (Oertzen et al. 2014), version 1.0-1029, and lavaan (Rosseel 2012), version 0.6-5, a SEM package in R (R Core Team 2019), version 3.6.2 (2019-12-12). To account for missing data, we used full information maximum likelihood estimation. Given that large differences in measurement scales, like those in our data, typically pose problems for numerical optimization algorithms, all observed variables were rescaled. We chose a scale with a mean of 5 and a standard deviation of 2. To evaluate model fit, we used the root-mean-square error of approximation (RMSEA), the comparative fit index (CFI), and the standardized root mean residual (SRMR). We interpret an RMSEA < 0.08, a CFI > 0.90, and SRMR < 0.8 as acceptable model fit (Schermelleh-Engel et al. 2003). To assess statistical significance of individual parameter estimates within a model, we used the likelihood ratio test. That is, we compared the model with the parameter of interest freely estimated to a nested model with this parameter fixed to zero, and compared whether the χ^2 difference between the models indicated a significant difference in fit (Kline 2015). For loadings and variance parameters, we used the Z-value (or Wald statistic; parameter estimate divided by its standard error) from the lavaan output.

In a first set of analyses, we specified separate CFAs to validate each latent integrity factor model for the selected ROIs (dorsolateral prefrontal cortex, medio-orbitofrontal cortex, hippocampus, parahippocampal gyrus, precuneus) defined by the indicators representing the three imaging modalities. In a next set of analyses, we combined the individual, fully saturated models to one structural equation model, in which all latent integrity factors were allowed to covary with one another. The intercorrelation of the medio-orbitofrontal cortex factor and the dorsolateral prefrontal cortex factor was too high for the two factors to be meaningfully modeled as separate latent variables (estimated $r > 1$), hence we decided to specify a single latent prefrontal cortex factor with six indicators, two per modality, from medio-orbitofrontal cortex and dorsolateral prefrontal cortex. Next, we added modality-specific latent factors (methods factors), which were defined to be orthogonal to the ROI factors, so that they represent the modality-specific share of the variance in the measures after the ROI-specific variance is accounted for. The methods factors were allowed to correlate with one another, and to be measured by all indicators that were derived from the respective modality, with loadings freely estimated (see Fig. 1). This type of model is known as a MTMM model (Campbell and Fiske 1959; Eid et al. 2008). It is the appropriate measurement model if multiple characteristics (usually traits, but here, ROIs) are each measured by several distinct measures (usually raters, but here, imaging modalities), yielding a latent integrity factor for each ROI and a latent method factor for each imaging modality. The latent scales of both the ROI factors and method factors were identified by fixing the loading of a reference indicator to one, which was the respective VBM indicator for the ROI factors and the respective measure of precuneus for the method factors. Within this model, we then correlated the ROI factors and the method factors with age to investigate age differences in variables of interest. We expected age differences for all ROI integrity factors (Raz et al. 2005; Bjell et al. 2009; Grydeland et al. 2013; Seiler et al. 2014). We

did not formulate any specific hypotheses for age differences in the methods factors, because method-specific variance was not in the focus of the investigation. Age differences in any method factor would indicate that this method is especially age-sensitive, over and above the age-related variance it shares with the other methods and that is captured in the ROI-wise integrity factors. Another purpose of investigating age differences was to be able to statistically control for age differences in both gray matter and episodic memory, potentially underlying observable associations between the two.

In a second set of analyses, we investigated associations between gray matter structure and episodic memory. An episodic memory latent factor based on these tasks and data was reported before (Düzel et al. 2016; Kühn et al. 2017). Performance on these four tasks is well captured by a latent factor of episodic memory ability as indicated by good fit both in the total sample, CFI > 0.999; RMSEA < 0.001; SRMR = 0.004, and in the MR sample, CFI > 0.999; RMSEA = 0.009; SRMR = 0.018. We set up the latent associations between gray matter integrity factors and episodic memory in two ways that are statistically equivalent but highlight different aspects of the multivariate associations: once we report covariances as estimates of the first-order correlations among ROIs and episodic memory (correlational model) and once we report multiple regression coefficients to assess the unique associations of each ROI factor with episodic memory while controlling for the other factors' association with it (regression model). With the correlational model, we sought to assess ROI-memory associations independent of the other ROIs in the model. As complementary piece of information, with the multiple regression model, we were able to assess how much variance in episodic memory each ROI factor accounts for, over and above the other ROI factors. In addition, we estimated the total variance in episodic memory that all ROI factors together accounted for (see MIMIC model in Kievit et al. 2012, p. 93).

In a next step, we entered age, and education (years) as well as sex into the correlational model to statistically control for the extent to which potential associations between episodic memory and gray matter integrity are being caused by these covariates. With respect to age, we expected that older participants tended to show lower gray matter integrity and lower episodic memory ability, so that not controlling for age in this age-heterogeneous sample would most likely yield a strong association between gray matter and episodic memory that is at least partly driven by those age differences. Education was expected to be related to episodic memory performance such that persons with more years of education tend to score better in episodic memory tests (Stern 2009). Education may affect performance as a consequence from education-related early training of memory, but also serves as a proxy for socioeconomic status, differences in which are not of interest in this study. Sex differences in episodic memory performance were to be expected (for a review, see Asperholm et al. 2019), and possibly even in gray matter structure (for a review, see Ruigrok et al. 2014). These then could induce an association if not adjusted for.

As an additional ad-hoc exploratory analysis, we investigated differences between men and women in the factor models and in the associations between gray matter integrity and episodic memory (Supplementary Material).

Models that entailed only brain data or brain data and covariates were fitted to data from the MR sample ($n = 333$), and models that entailed episodic memory data were fitted to data from

the total sample, under the assumption that the MR data were missing at random (Rubin 1976; Schafer and Graham 2002). This assumption holds as long as missingness in the gray matter variables is either completely random or can be explained by variables in the model.

Results

Sample Descriptives

Table 1 displays the descriptive statistics of the selected covariates and episodic memory task performance variables in the full sample and the MR sample. None of the variables were heavily skewed, and kurtosis was high only in the verbal learning and memory test in the full sample, so we assumed that all variables follow the normal distribution to an acceptable extent. The MR sample did not differ much from the full BASE-II sample, with selectivity below 0.13 standard deviations in the continuous variables. The only considerable difference was found in the sex distribution, with ~38% females in the MR sample and 51% females in the full sample. In Table 2, we report the descriptive statistics of the gray matter structure variables in their original scale in the MR sample. We deemed skewness and kurtosis levels acceptable in these variables, too. Days between the cognitive assessment and the MR-session differed between individuals (absolute difference in days: mean = 113.8, SD = 126.9, min = 8, max = 774). For sex differences in all variables of interest, see the Supplementary Material. For pairwise correlations between all variables of interest, see Table 3.

We succeeded in establishing a factor model with four ROI factors capturing shared variance across VBM, MT, and MD within each ROI (prefrontal cortex, hippocampus, parahippocampal gyrus, precuneus). We extracted ROI-wise data from five ROIs, but the two frontal ROIs medio-orbitofrontal cortex and dorsolateral-prefrontal cortex shared such a large amount of their variance (estimated correlation of $r > 1$) that we rather estimated only latent factor (prefrontal cortex) with all six indicators from the two regions, which is a more parsimonious representation of the common variance of these regions. The model included method-specific factors (VBM, MT, MD) that were orthogonal to the ROI factors and captured the shared variance of measures within imaging modalities and across ROIs (Fig. 1). We allowed for residual covariances between the VBM indicators of closely neighboring ROIs (namely, of medio-orbitofrontal and dorsolateral prefrontal cortex and of hippocampus and parahippocampal gyrus), as we expected dependencies between them to exceed the shared variance modeled in the latent VBM factor. Fixing these residual covariances to zero (as done with all other residual covariances among indicator variables) significantly decreased model fit. The proposed model with the residual covariances freely estimated fitted the data well, $\chi^2(df=64) = 152.876$; CFI = 0.965; RMSEA = 0.065; SRMR = 0.046. As the residual variance of the indicator variable MD medio-orbitofrontal cortex was estimated at a low negative value, we constrained that parameter to zero in all following models. This constraint did not result in worse model fit ($\Delta\chi^2(df=1) = 0.01$, $P = 0.91$). All observed variables loaded reliably on the postulated latent ROI factors except for MD of medio-orbitofrontal cortex on the prefrontal cortex factor (standardized loading = -0.02, $z = -0.16$, $P = 0.87$, all other standardized loadings > 0.27 , z 's > 2.13 , P 's < 0.034). All

indicators loaded reliably on the postulated method factors (absolute standardized loadings > 0.18 , abs. z 's > 3.12 , $P < 0.003$). We estimated covariances among the ROI integrity factors and among method factors, while method factors were defined as being orthogonal to ROI factors. For covariances among the latent factors see Table 4.

Association Between Latent Factors and Age

We included age in the model to estimate covariances between age and all latent factors. Model fit remained acceptable, $\chi^2(df=73) = 171.707$; CFI = 0.962; RMSEA = 0.064; SRMR = 0.045. Age was negatively associated with all ROI factors except precuneus, $r_{age, PFC} = -0.21$, $\Delta\chi^2(df=1) = 5.66$, $P = 0.017$; $r_{age, HC} = -0.22$, $\Delta\chi^2(df=1) = 13.2$, $P = 0.0003$; $r_{age, PHG} = -0.22$, $\Delta\chi^2(df=1) = 9.22$, $P = 0.0024$; $r_{age, PRE} = 0.01$, $\Delta\chi^2(df=1) = 0.026$; $P = 0.87$. In addition, age was associated with the method factors ($r_{age, V} = -0.16$, $\Delta\chi^2(df=1) = 4.35$, $P = 0.037$; $r_{age, MT} = -0.29$, $\Delta\chi^2(df=1) = 11.3$, $P = 0.0008$; $r_{age, MD} = 0.40$, $\Delta\chi^2(df=1) = 35.4$, $P < 0.0001$). We had also tested for measurement invariance across age groups after a median split, which suggested no age differences in the loadings on the ROI factors (Supplementary Table 9).

Associations Between Gray Matter Integrity ROIs and Episodic Memory in the Correlational Model

We added an episodic memory latent factor to the model, which was indexed by the four episodic memory task scores. We then estimated correlations between the episodic memory factor and the ROI integrity factors (correlational model) as well as the method factors. We fit this model to the data from all included participants with episodic memory data available ($n = 1522$), assuming missingness of MR data to be at random. The fit was acceptable, $\chi^2(df=120) = 203.27$; CFI = 0.97; RMSEA = 0.02; SRMR = 0.05 (for more fit indices, see Supplementary Table 1). The scores that served as indicators all loaded significantly on the latent episodic memory factor with moderate effect size (standardized loadings between 0.46 and 0.56, z 's > 10.41 , P 's < 0.001), indicating that they contributed similarly to the latent factor. Episodic memory was significantly associated with all ROIs except precuneus, $r_{EM, PFC} = 0.24$, $\Delta\chi^2(df=1) = 7.02$; $P = 0.008$; $r_{EM, HC} = 0.33$, $\Delta\chi^2(df=1) = 16$; $P < 0.0001$; $r_{EM, PHG} = 0.27$, $\Delta\chi^2(df=1) = 9.28$; $P = 0.0023$; $r_{EM, PRE} = 0.16$, $\Delta\chi^2(df=1) = 3.29$; $P = 0.07$. In contrast, episodic memory was not reliably related to the method factors, $r_{EM, V} = 0.07$, $\Delta\chi^2(df=1) = 0.44$, $P = 0.51$; $r_{EM, MT} = -0.02$, $\Delta\chi^2(df=1) = 0.03$; $P = 0.86$; $r_{EM, MD} = -0.12$, $\Delta\chi^2(df=1) = 2.19$; $P = 0.41$ (for a list of all estimates, see Supplementary Table 1). Hence, a model with these associations fixed to zero (see Figure 1) did not fit worse than a model estimating them freely, $\Delta\chi^2(df=3) = 3.5$, $P = 0.32$, and was used in the following extensions of the model ($\chi^2(df=123) = 206.77$; CFI = 0.973; RMSEA = 0.021, SRMR = 0.05).

Unique Associations of ROIs with Episodic Memory in Regression Model

To assess how much variance in episodic memory is uniquely and jointly predicted by the latent ROI factors, we refit the previous model with directed paths from each ROI to latent episodic memory (regression model). This is a latent multiple regression model regressing episodic memory on each ROI integrity factor. Note that the directionality of effects (i.e., ROI

Table 3 Pairwise first-order correlations between all observed variables of interest

		Memory					VBM				MT				MD			Covariates				
		VLMT	FP	SE	OL	HC	PHG	PRE	OFC	DLPFC	HC	PHG	PRE	OFC	DLPFC	HC	PHG	PRE	OFC	DLPFC	Age	Edu
Memory	FP	0.25																				
	SE	0.26	0.22																			
	OL	0.30	0.25	0.29																		
VBM	HC	0.12	0.18	0.11	0.12																	
	PHG	0.08	0.09	0.01	-0.06	0.61																
	PRE	0.04	0.07	0.09	0.04	0.20	0.31															
	OFC	0.12	0.03	0.10	0.07	0.29	0.39	0.45														
MT	DLPFC	0.08	0.05	0.13	0.07	0.23	0.35	0.40	0.78													
	HC	0.16	0.23	0.21	0.05	0.57	0.34	0.20	0.32	0.36												
	PHG	0.17	0.17	0.13	-0.04	0.37	0.30	0.13	0.24	0.29	0.69											
	PRE	0.06	0.07	0.08	0.01	0.14	0.07	0.39	0.09	0.13	0.19	0.33										
MD	OFC	0.06	0.10	0.07	0.01	0.21	0.07	0.12	0.20	0.25	0.46	0.45	0.29									
	DLPFC	-0.03	0.08	0.02	-0.09	0.14	0.04	0.15	0.18	0.25	0.32	0.35	0.53	0.57								
	HC	-0.05	-0.16	-0.26	-0.11	-0.63	-0.35	-0.20	-0.25	-0.31	-0.86	-0.46	-0.13	-0.33	-0.26							
	PHG	-0.14	-0.13	-0.26	-0.09	-0.53	-0.44	-0.36	-0.36	-0.40	-0.69	-0.63	-0.39	-0.35	-0.32	0.64						
Covariates	PRE	-0.09	-0.12	-0.16	-0.07	-0.29	-0.26	-0.53	-0.26	-0.29	-0.27	-0.39	-0.76	-0.30	-0.37	0.29	0.58					
	OFC	-0.08	-0.14	-0.14	-0.04	-0.23	-0.25	-0.36	-0.30	-0.30	-0.25	-0.27	-0.46	-0.46	-0.61	0.24	0.52	0.66				
	DLPFC	-0.13	-0.15	-0.20	-0.03	-0.37	-0.34	-0.37	-0.39	-0.40	-0.47	-0.50	-0.40	-0.65	-0.54	0.48	0.65	0.55	0.71			
	Age	-0.13	-0.16	-0.16	-0.05	-0.30	-0.23	-0.11	-0.10	-0.11	-0.27	-0.28	-0.21	-0.30	-0.24	0.33	0.38	0.30	0.41	0.46		
Covariates	Edu	0.18	0.16	0.09	0.15	-0.03	-0.02	-0.01	-0.05	-0.08	-0.02	-0.05	-0.11	0.07	-0.04	0.06	0.11	0.12	-0.01	0.01	-0.06	
	Sex	0.09	0.03	0.04	0.04	0.21	0.09	0.21	0.32	0.35	0.37	0.28	0.26	0.26	0.16	-0.26	-0.40	-0.28	-0.18	-0.30	0.02	-0.15

Notes: VLMT, verbal learning and memory test; FP, face-profession task; SE, scene encoding; OL, object location task; VBM, volume-based morphometry; MT, magnetization transfer ratio; MD, mean diffusivity; HC, hippocampus; PHG, parahippocampal gyrus; PRE, precuneus; MOFC, medio-orbitofrontal cortex; DLPFC, dorsolateral prefrontal cortex; edu, years of education. Based on pairwise complete data.

integrity factor predicting memory) is merely hypothesized and cannot be tested with the data at hand. Importantly, instead of interpreting first-order correlations, we now examined the total effect and the unique effects of each ROI integrity factor on episodic memory. All ROIs together explained 12.5% of the variance in episodic memory ($R^2 = 0.125$). None of the ROIs showed a significant unique effect. However, the unique effect of hippocampus had the largest effect size; it was greater than those of parahippocampal gyrus, precuneus, and prefrontal cortex by a factor of 2.5, 3.2, and 4.8, respectively (std. $\beta_{EM, HC} = 0.38$, $\Delta\chi^2_{(df=1)} = 3.12$; $P = 0.08$; std. $\beta_{EM, PHG} = -0.15$, $\Delta\chi^2_{(df=1)} = 0.28$; std. $\beta_{EM, PRE} = 0.12$, $\Delta\chi^2_{(df=1)} = 1.05$; $P = 0.31$; $P = 0.59$; std. $\beta_{EM, PFC} = 0.08$, $\Delta\chi^2_{(df=1)} = 0.33$; $P = 0.57$).

Adjusting for Covariate Effects

In follow-up analyses, we entered age as a covariate into the correlational model. As before, in the brain-only model from the first set of analyses, age was negatively associated with all ROI factors except precuneus and with all method factors. Here, we regressed the ROI integrity factors, the method factors, and the episodic memory factor on age and examined the residual covariances between episodic memory and the ROI integrity factors. Age was negatively associated with episodic memory (Table 5). Moreover, the associations between episodic memory and the ROI factors were attenuated after controlling for age, leaving only hippocampus and parahippocampal gyrus being significantly associated with episodic memory, $r_{EM, PFC} = 0.16$, $\Delta\chi^2_{(df=1)} = 2.98$; $P = 0.08$; $r_{EM, HC} = 0.27$, $\Delta\chi^2_{(df=1)} = 10.22$; $P = 0.0014$; $r_{EM, PHG} = 0.21$, $\Delta\chi^2_{(df=1)} = 5.2$; $P = 0.023$; $r_{EM, PRE} = 0.13$, $\Delta\chi^2_{(df=1)} = 1.95$; $P = 0.16$. Thus, the gray matter integrity factors of hippocampus and

parahippocampal gyrus shared a significant amount of variance with episodic memory that was not collinear with age, further suggesting that the structural integrity of these two regions might be critical for episodic memory.

We then added years of education into the model as a covariate by regressing all ROI factors and episodic memory as well as the method factors on age and education, which had very little effect on estimates of associations and did not change the pattern of significance, $r_{EM, PFC} = 0.16$, $\Delta\chi^2_{(df=1)} = 3.04$; $P = 0.08$; $r_{EM, HC} = 0.27$, $\Delta\chi^2_{(df=1)} = 10.43$; $P = 0.0012$; $r_{EM, PHG} = 0.23$, $\Delta\chi^2_{(df=1)} = 5.92$; $P = 0.015$; $r_{EM, PRE} = 0.15$, $\Delta\chi^2_{(df=1)} = 2.6$; $P = 0.11$. For a test of measurement invariance across education, see Table S10. No differences in loadings nor residual variances were found. When we added sex as a covariate, all associations were somewhat further attenuated, but hippocampus and parahippocampal gyrus were still significantly associated with episodic memory, $r_{EM, PFC} = 0.05$, $\Delta\chi^2_{(df=1)} = 0.29$; $P = 0.59$; $r_{EM, HC} = 0.19$, $\Delta\chi^2_{(df=1)} = 4.66$; $P = 0.031$; $r_{EM, PHG} = 0.13$, $\Delta\chi^2_{(df=1)} = 4.08$; $P = 0.043$; $r_{EM, PRE} = 0.07$, $\Delta\chi^2_{(df=1)} = 1.76$; $P = 0.19$ (for a full list of parameter estimates and for additional fit indices, see Supplementary Table 2).

As an additional ad-hoc exploratory analysis, we investigated differences between men and women in the associations between gray matter integrity and episodic memory (Supplementary Figure 1) After testing for measurement invariance across sexes (Supplementary Tables 5 and 6), we ran the correlational model as a multigroup model, once with (Supplementary Table 7) and once without age and education as covariates (Supplementary Table 8). The associations between hippocampal and parahippocampal gray matter integrity were restricted to men (see Supplementary Material, Section 4, for details).

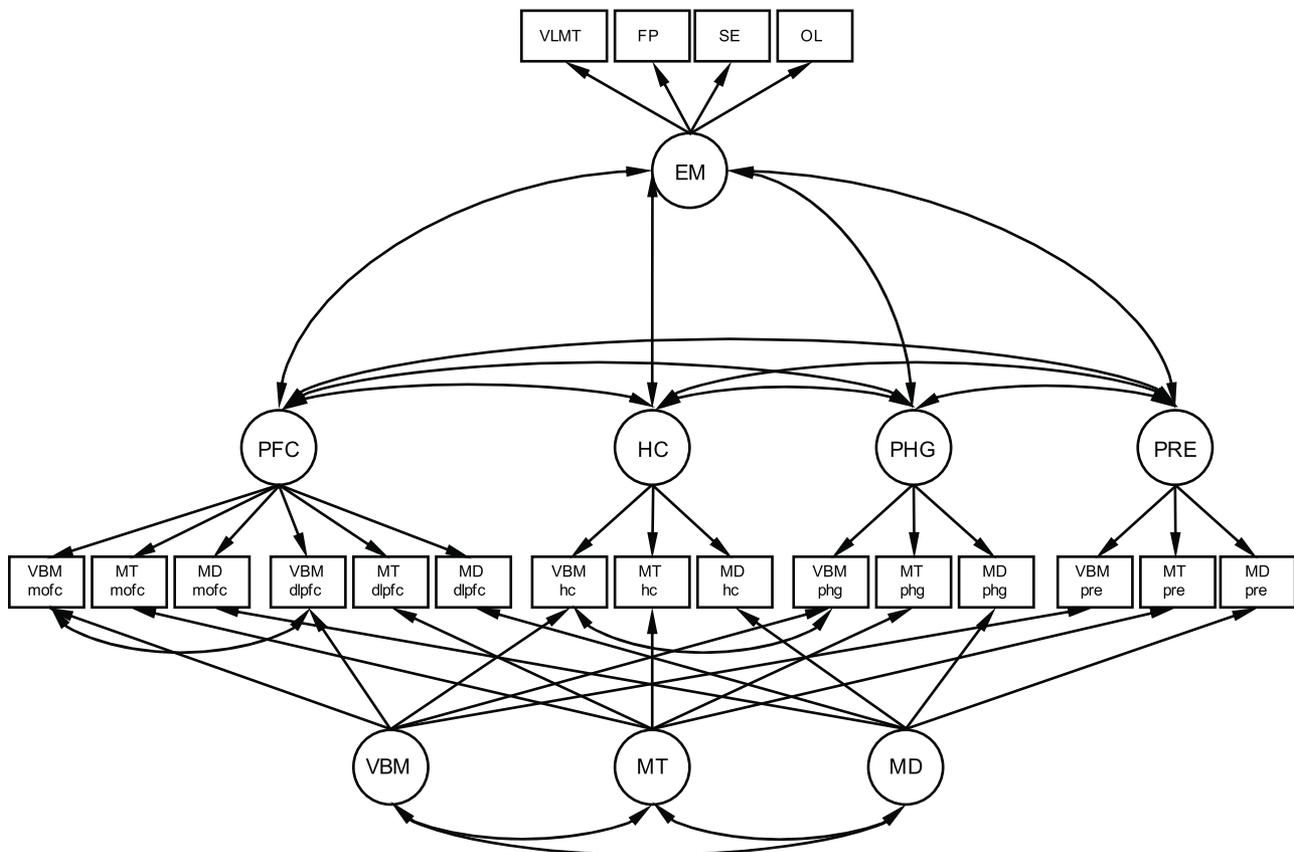


Figure 1. Correlational model. Gray matter integrity factor model separating common variance across structural imaging modalities for each of the ROIs (PFC, HC, PHG, PRE) from method-specific variance (VBM, MT, MD) and residual variance, and associating gray matter factors with episodic memory. Circles depict latent variables, squares depict observed variables. Double-headed arrows are covariances, single-headed arrows are loading parameters. Variances of all observed and all latent variables were estimated, but not visualized in this figure for the sake of clarity. Variable names beginning with VBM are VBM-derived gray matter probability measures, names with MT are magnetization transfer ratio measures, names with MD are mean diffusivity measures. Further, names indicate the regions: mofc, medio-orbitofrontal cortex; dlpfc, dorsolateral prefrontal cortex; hc, hippocampus; phg, parahippocampal gyrus. Upper part: Episodic memory factor, measured by four tasks. EM, episodic memory; VLMT, verbal learning and memory test; FP, face-profession task; SE, scene-encoding task; OL, object location task.

Table 4 Correlations among latent factors in the gray matter integrity model

	Prefrontal cortex	Hippocampus	Parahippocampal g.	Precuneus	Method VBM	Method MD	Method MT
Prefrontal cortex	1						
Hippocampus	0.59 ^a	1					
Parahippocampal g.	0.68 [*]	0.84 ^a	1				
Precuneus	0.18	0.13	0.37 [*]	1			
Method: VBM	0 ^b	0 ^b	0 ^b	0 ^b	1		
Method: MD	0 ^b	0 ^b	0 ^b	0 ^b	-0.53 [*]	1	
Method: MT	0 ^b	0 ^b	0 ^b	0 ^b	0.25 [*]	-0.77 [*]	1

Notes: ^aModel with this covariance restricted to 0 did not converge, so the Wald test was used instead for statistical inference

^bThe parameter was defined as zero in the model.

* $P < 0.05$; according to χ^2 difference test/likelihood ratio test with 1 df.

Discussion

In this study, we used cross-sectional data on multimodal structural imaging and episodic memory tasks from a large cohort to establish a structural equation model of regional gray matter structure integrity and its associations with episodic memory. We show that a MTMM latent factor representation of regional individual differences in gray matter structure enables

researchers to examine links of structural brain properties to behavior. Specifically, this representation allows researchers to separate three sources of variance from one another: 1) variance shared within each ROI across imaging modalities (i.e., the ROI integrity factors); 2) variance shared within each imaging modality across ROIs (i.e., the method factors); 3) variance unique to each ROI in each modality (i.e., residual variance).

Table 5 Unique effects of age, education, and sex on latent episodic memory and ROI integrity factors

Covariate	Effect on EM, β /std. β (SE)	Effect on PFC β /std. β (SE)	Effect on HC β /std. β (SE)	Effect on PHG β /std. β (SE)	Effect on PRE β /std. β (SE)
Age (years)	-0.06/-0.22 (0.01)*	-0.04/-0.28 (0.01)*	-0.09/-0.28 (0.02)*	-0.02/-0.17 (0.01)*	-0.01/-0.06 (0.02)
Education (years)	0.11/0.29 (0.02)*	0.01/0.06 (0.01)	0.01/0.02 (0.02)	-0.002/-0.02 (0.01)	-0.03/-0.12 (0.02)
Sex (female advantage)	0.32*	0.32*	0.69*	0.43*	0.31*

Notes: Standardized regression coefficient estimates and SE of estimate. Outcome variables are standardized to mean = 5 and SD = 2. EM, episodic memory; PFC, prefrontal cortex; HC, hippocampus; PRE, precuneus. * $P < 0.05$, according to Wald statistic

The psychometric viability of the MTMM representation of regional gray matter integrity demonstrates that macro- and micro-structural indicators of gray matter can indeed be combined to yield latent factors of gray matter integrity. In addition, the latent integrity factors formed a positive manifold, indicating that individual differences in gray matter integrity are correlated across regions. By moving away from specific aspects of integrity indicators to the expression of their common variance at the latent level, we pave the way for a deeper understanding of relations between brain structure and cognitive performance.

Older participants tended to show lower values on all ROI-wise latent integrity factors except precuneus. This result is largely in line with previous findings based on single indicators focusing on volume (Raz et al. 2005; Fjell and Walhovd 2011), and MT ratio (Seiler et al. 2014), or MD (Grydeland et al. 2013). We interpret this as suggesting that older individuals tend to have experienced more gray matter deterioration and therefore tend to show lower values in most ROI factors, which reflect a pattern of lower gray matter density in VBM, lower MT ratio, and higher MD. Also, all method factors showed age differences in the direction of less beneficial values in older participants. This suggests that the variance in each of the ROI factors is not capturing all age differences. There are still age differences in the methods' unique variances. In other words, older individuals tend to show less gray matter integrity (across modalities) in prefrontal cortex, hippocampus, and parahippocampal gyrus, and in addition, they tend to have smaller volumes, lower MT, and higher MD across ROIs.

To examine whether latent gray matter integrity factors are related to episodic memory, we tested their associations with episodic memory ability at the latent level. Episodic memory showed first-order associations with the structural integrity factors of all ROIs except precuneus, but not with the modality-specific method factors. When adjusted for age differences, hippocampus and parahippocampal gyrus continued to be associated with episodic memory. That is, while prefrontal cortex's first-order association with episodic memory could be accounted for by age differences in both gray matter structure and performance, the integrity of hippocampal and parahippocampal gray matter not only reflected individual differences collinear with chronological age, but also associations with episodic memory performance over and above age. Adjusting for interindividual differences in years of education did not substantially affect the associations, with hippocampus and parahippocampal gyrus still showing the strongest associations with episodic memory. When adjusting for sex differences in episodic memory and ROI integrity, hippocampus and parahippocampal gyrus were still significantly associated with episodic memory. This is corroborated in the regression model,

with hippocampus showing the numerically largest unique effect. Overall, this result strongly supports the hypothesis that maintained structural integrity of the hippocampus is germane to preserved episodic memory ability in old age (de Chastelaine et al. 2011; Nyberg et al. 2012; Cabeza et al. 2018; Nyberg and Lindenberger 2020; Nyberg and Pudas 2019).

To note, we conceptualize episodic memory on a relatively broad level. Our current focus on the latent factor, which captures the shared between-person variance across these four different tasks, implies that we abstract from the details of the tasks and focus on the commonalities when interpreting associations with gray matter integrity in the ROIs. The chosen tasks are heterogeneous in terms of stimulus material (VLMT: verbal, OL, SE: figural, FP: both), sensory modality of presentation (VLMT: auditory, OL, SE, FP: visual), type of memory (VLMT, SE: item memory; FP, OL: associative memory), or retrieval type (free recall: VLMT, OL; recognition: FP and SE). Performance may be differentially influenced by component processes such as familiarity and recognition (Yonelinas 1994). However, in our view, the heterogeneity of tasks across all these dimensions can also be thought of as a strength (Little et al. 1999). By virtue of the method, the latent factor is void of the specifics of the tasks and extracts what is common to them, and thereby allows us to examine associations to ROIs at the general level of episodic memory.

The aims of this study were to establish a gray matter integrity factor model and validate it by associating its latent factors with episodic memory performance. We did not previously plan to investigate sex differences in measurement models or associations. Only after observing sex differences both in estimates of average gray matter integrity and episodic memory, we ran additional post hoc exploratory analyses to compare the models across men and women. We found the associations between hippocampal and parahippocampal gray matter integrity to be restricted to men. We provide details on these additional analyses and a short discussion in the [Supplementary Material](#). In consequence of this finding, we note that we might wrongly generalize across sexes when interpreting the analyses that are not considering sex differences in associations. Still, the associations may be present in both sexes, however, the processes that lead to large-enough interindividual differences to detect these associations evolve earlier, on average, in men, then in women. In essence, we still hypothesize that the hippocampal and parahippocampal integrity in older adults are relevant for episodic memory performance irrespective of sex. At this point, we can only speculate that there might be more men than women who have already experienced some gray matter integrity deterioration with consequences for memory functioning, possibly related to men carrying a higher metabolic risk with detrimental effects

for both gray matter integrity and episodic memory (Raz et al. 2005; Raz and Rodrigue 2006; Yates et al. 2012). This could also be a reason for the observation that men show on average lower integrity in all ROIs and in episodic memory (Supplementary Table 4). Further elaboration and investigation of these sex differences would exceed the scope of this study and should be pursued in future studies based on longitudinal data.

Our results also suggest that the combination of multimodal data yields information about general properties of gray matter tissue that differ between younger and older individuals above 60 years of age, and are relevant for older adults' episodic memory performance. This raises the important question of which physiological aspects of gray matter are captured by the common variance of regional brain integrity as estimated by VBM, MT, and MD. Given that MD and MT ratio load on the same factor as VBM, it seems worthwhile to consider physiological factors that affect the physical properties of the tissue and its overall size. Normal aging is marked by the loss of dendritic spines, dendritic arbors, synaptic density, and myelinated axons (Hof and Morrison 2004; Morrison and Baxter 2012); in addition, normal aging also involves loss of glia and small blood vessels (Raz and Daugherty 2018). All of these processes can be assumed to lead to a reduction in average tissue density as captured by MD and MT, and to a concomitant decrease in overall volume as captured by VBM. In terms of relative contributions to variations in the MR signal that affect MD, MT, and VBM in a correlated manner, we surmise that individual differences in cortical myelin might play a prominent role. Given that histochemical staining of myelin has shown that myelin coverage is more extensive in deeper relative to superficial cortical layers (Timmler and Simons 2019), one way to follow up on this proposition would be to test for differences in myelin content between layers using structural imaging methods with laminar resolution (Peters and Kemper 2012; Waehnert et al. 2014). Note that the hippocampus is a relatively small structure with complex shape, structure, and function. Given its complex geometry, we cannot rule out that embedded white matter and CSF might contribute to the ROI-specific estimates. Moreover, the hippocampus has functionally distinct subfields (Yassa and Stark 2011), which could not be set apart with the imaging protocols that we used. When interpreting integrity estimates for any given ROI, and the hippocampus in particular, one must bear in mind that such estimates represent aggregates over more or less heterogeneous structures. The primary aim of the present study is to demonstrate the feasibility of a latent factor approach to capturing individual differences in brain integrity at the ROI level. The content validity of this approach awaits further scrutiny. For instance, future work may be able to define ROIs at the resolution of hippocampal subfields, which might reduce confounds due to white matter and CSF while increasing content validity and specificity (e.g., Keresztes et al. 2018).

The MTMM model of regional gray matter integrity introduced in this article reflects correlated traits and correlated methods, and properly accounts for the nonindependent structure of measurement errors in our data. Similar to other confirmatory factor analysis variants of structural equation modeling, it links the measurement model (ROI integrity factors, method factors, and residuals) to the structural model (associations between latent factors). The structural part of the model allows researchers to explore relations to behavior, and their modulation by covariates.

The proposed models are certainly not the only ways to model associations between brain structure and cognition

(Kievit et al. 2012). Hence, we would like to encourage researchers to adopt the MTMM approach whenever they have multiple measures for a given construct of interest. Future work may include a larger number of ROIs, a more fine-grained parcellation of ROIs into subregions, interhemispheric differences and commonalities, a larger number of indicators, or additional cognitive domains to investigate the domain-specificity of associations. Furthermore, the general approach can be expanded to include factors of white-matter integrity, neurochemistry, or brain activity as assessed by functional MRI.

In this study, we have used a statistical approach to model the common variance across multiple indicators of gray matter integrity in latent factors for each ROI. At this point, we can only speculate about the physiological basis of individual differences in gray matter integrity captured with the MTMM approach. To overcome these ambiguities, the field needs a stronger coalition between animal models and human research, with structural MRI serving as a critical link (Lerch et al. 2017).

One may wonder whether age-related artifacts present in each of the imaging modalities might account for age differences in the latent ROI and method factors. To reiterate, we modeled the variability in the ROI-wise data from each modality as a combination of a ROI-specific, modality-general part; a modality-specific, ROI-general part; and a residual part that is both ROI- and modality-specific. As ROI factors represent common variance across methods in a specific ROI, older individuals tended to show lower volumes, lower MT and higher MD in these three ROIs. This suggests that older individuals tend to possess lower gray matter integrity in the ROIs, and/or that age-associated artifacts of the methods play out to a similar degree in the different methods in these ROIs. We are not able to tease apart these two possible causes of the age-ROI associations with the current modeling approach. However, here, the differences between what is measured by the three imaging modalities renders it unlikely that correlated age-associated measurement error dominates the common variance to such an extent that it would explain the emergence of a factor structure. In addition, it is actually a strength of the present approach that it allows to estimate age differences in the methods factors, and statistically adjust for them if deemed meaningful.

Furthermore, the present findings are based on cross-sectional data. The observed associations with age represent the joint outcome of individual differences in normal aging and more stable individual differences that were present in early adulthood (Hertzog 1985). It remains to be seen how changes in individual differences in latent patterns of brain integrity map onto changes in episodic memory. Hence, the present analyses need to be extended to longitudinal investigations that examine individual differences in latent brain integrity changes and their correlation with cognitive changes (for methodological work in developmental psychology, see Geiser et al. 2010).

Also note that this paper focuses on the association structure of individual differences in a healthy older population. Our results might not hold for all subgroups. Plausibly, there are hidden heterogeneities in the association structures that should be elucidated by follow-up studies. For instance, associations between dopamine availability and cognition have been found to differ between subgroups in a latent class analysis (Lövdén et al. 2017). Another data-driven way to identify hidden heterogeneities in associations are decision trees (Strobl et al. 2009), which can be usefully combined with structural equation models in structural equation modeling trees (Brandmaier

et al. 2013). In addition to the structural integrity measures we investigated in this study, a comprehensive understanding of maintenance may further benefit from the integration of additional imaging modalities such as white matter integrity, neurochemical, and connectivity measures.

By applying MTMM modeling to data from a large sample of BASE-II participants, we established latent factors of gray matter integrity in hippocampus, parahippocampal gyrus, prefrontal cortex, and precuneus, which represented the shared variance of VBM, MT, and MD for each of these regions. Further, we found that older adults with greater structural integrity in hippocampus and parahippocampal gyrus also showed higher levels of episodic memory performance, with hippocampus showing the largest unique association. Our results are consistent with the hypothesis that maintained structural integrity of the hippocampus helps to preserve episodic memory in old age. Future research needs to corroborate the content validity of the latent brain factors, and extend the present approach to longitudinal observations.

Supplementary Material

Supplementary material can be found at *Cerebral Cortex* online.

Notes

We are grateful for the assistance of the MRI team at the Max Planck Berlin Institute for Human Development consisting of Sonali Beckmann, Nils Bodammer, Thomas Feg, Sebastian Schröder, and Nadine Taube, for the team leading the cognitive tests, and for all participants of BASE-II. *Conflict of interest:* None declared.

Funding

European Commission as part of the Lifebrain Consortium (grant number 732592) within the Horizon 2020 programme; German Federal Ministry of Education and Research (grant number 01GQ1421B); the MINERVA program of the Max Planck Society (to M.C.S.).

Data Availability Statement

Data can be requested from the steering committee of the Berlin Aging Study II. Further information regarding the application can be found under <https://www.base2.mpg.de/en>.

References

- Abe O, Yamasue H, Aoki S, Suga M, Yamada H, Kasai K, Masutani Y, Kato N, Ohmoto K. 2008. Aging in the CNS: comparison of gray/white matter volume and diffusion tensor data. *Neurobiol Aging*. 29:102–116.
- Aribisala BS, Royle NA, Maniega SM, Valdés Hernández MC, Murray C, Penke L, Gow A, Starr JM, Bastin ME, Deary IJ et al. 2014. Quantitative multi-modal MRI of the Hippocampus and cognitive ability in community-dwelling older subjects. *Cortex*. 53:34–44.
- Ashburner J, Friston KJ. 2000. Voxel-based morphometry—the methods. *Neuroimage*. 11:805–821.
- Asperholm M, Högman N, Rafi J, Herlitz A. 2019. What did you do yesterday? A meta-analysis of sex differences in episodic memory. *Psychol Bull*. 145:785–821.
- Bartrés-Faz D, Arenaza-Urquijo EM. 2011. Structural and functional imaging correlates of cognitive and brain reserve hypotheses in healthy and pathological aging. *Brain Topogr*. 24:340–357.
- Becker N, Laukka EJ, Kalpouzos G, Naveh-Benjamin M, Bäckman L, Brehmer Y. 2015. Structural brain correlates of associative memory in older adults. *Neuroimage*. 118:146–153.
- Benoit RG, Schacter DL. 2015. Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. *Neuropsychologia*. 75:450–457.
- Bertram L, Böckenhoff A, Demuth I, Düzel S, Eckardt R, Li S-C, Lindenberger U, Pawelec G, Siedler T, Wagner GG et al. 2013. Cohort profile: The Berlin Aging Study II (BASE-II). *Int J Epidemiol*. 43:703–712.
- Brandmaier AM, Oertzen v T, McArdle JJ, Lindenberger U. 2013. Structural equation model trees. *Psychol Methods*. 18:71–86.
- Cabeza R, Albert M, Belleville S, Craik FIM, Duarte A, Grady CL, Lindenberger U, Nyberg L, Park DC, Reuter-Lorenz PA et al. 2018. Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nat Rev Neurosci*. 19:701–710.
- Campbell DT, Fiske DW. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull*. 56:81–105.
- Carlesimo GA, Cherubini A, Caltagirone C, Spalletta G. 2010. Hippocampal mean diffusivity and memory in healthy elderly individuals. *Neurology*. 74:194.
- Cavanna AE, Trimble MR. 2006. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*. 129:564–583.
- Chalmers RP, Flora DB. 2015. faoutlier: an R package for detecting influential cases in exploratory and confirmatory factor analysis. *Appl Psychol Measur*. 39:573–574.
- de Chastelaine M, Wang TH, Minton B, Muftuler LT, Rugg MD. 2011. The effects of age, memory performance, and callosal integrity on the neural correlates of successful associative encoding. *Cereb Cortex*. 21:2166–2176.
- de Frias CM, Lövdén M, Lindenberger U, Nilsson L-G. 2007. Revisiting the dedifferentiation hypothesis with longitudinal multi-cohort data. *Intelligence*. 35:381–392.
- Dickerson BC, Eichenbaum H. 2009. The episodic memory system: neurocircuitry and disorders. *Neuropsychopharmacology*. 35:86–104.
- Düzel S, Münte TF, Lindenberger U, Bunzeck N, Schütze H, Heinze H-J, Düzel E. 2010. Basal forebrain integrity and cognitive memory profile in healthy aging. *Brain Res*. 1308:124–136.
- Düzel S, Schütze H, Stallforth S, Kaufmann J, Bodammer N, Bunzeck N, Münte TF, Lindenberger U, Heinze H-J, Düzel E. 2008. A close relationship between verbal memory and SN/VTA integrity in young and older adults. *Neuropsychologia*. 46:3042–3052.
- Düzel S, Voelkle MC, Düzel E, Gerstorff D, Drewelies J, Steinhagen-Thiessen E, Demuth I, Lindenberger U. 2016. The Subjective Health Horizon Questionnaire (SHH-Q): assessing future time perspectives for facets of an active lifestyle. *Gerontology*. 62:345–353.
- Eichenbaum H. 2017. Prefrontal–hippocampal interactions in episodic memory. *Nat Rev Neurosci*. 18:547–558.
- Eid M, Nussbeck FW, Geiser C, Cole DA, Gollwitzer M, Lischetzke T. 2008. Structural equation modeling of multitrait-multimethod data: different models for different types of methods. *Psychol Methods*. 13:230–253.

- Fjell AM, Walhovd KB. 2011. Structural brain changes in aging: courses, causes and cognitive consequences. *Rev Neurosci*. 187–221.
- Fjell AM, Walhovd KB. 2016. Multimodal imaging of the aging brain. In: Cabeza R, Nyberg L, Park DC, editors. *Cognitive neuroscience of aging: linking cognitive and cerebral aging*. 2nd ed. Oxford (UK): Oxford University Press, p. 131–182.
- Fjell AM, Walhovd KB, Fennema-Notestine C, McEvoy LK, Hagler DJ, Holland D, Brewer JB, Dale AM. 2009. One-year brain atrophy evident in healthy aging. *J Neurosci*. 29: 15223–15231.
- 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*. 25:940–918.
- Freton M, Lemogne C, Bergouignan L, Delaveau P, Lehericy S, Fossati P. 2014. The eye of the self: precuneus volume and visual perspective during autobiographical memory retrieval. *Brain Struct Funct*. 219:959–968.
- Geiser C, Eid M, Nussbeck FW, Courvoisier DS, Cole DA. 2010. Analyzing true change in longitudinal multitrait-multimethod studies: application of a multimethod change model to depression and anxiety in children. *Dev Psychol*. 46:29–45.
- Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RS. 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage*. 14:21–36.
- Gorbach T, Pudas S, Lundquist A, Orädd G, Josefsson M, Salami A, de Luna X, Nyberg L. 2017. Longitudinal association between hippocampus atrophy and episodic-memory decline. *Neurobiol Aging*. 51:167–176.
- Grydeland H, Walhovd KB, Tamnes CK, Westlye LT, Fjell AM. 2013. Intracortical myelin links with performance variability across the human lifespan: results from T1- and T2-weighted MRI myelin mapping and diffusion tensor imaging. *J Neurosci*. 33:18618–18630.
- Hedden T, Schultz AP, Rieckmann A, Mormino EC, Johnson KA, Sperling RA, Buckner RL. 2016. Multiple brain markers are linked to age-related variation in cognition. *Cereb Cortex*. 26:1388–1400.
- Henson RN, Campbell KL, Davis SW, Taylor JR, Emery T, Erzincliglu S, Cam-CAN, Kievit RA. 2016. Multiple determinants of lifespan memory differences. *Sci Rep*. 6: 32527.
- Hertzog C. 1985. An individual differences perspective: implications for cognitive research in gerontology. *Res Aging*. 7:7–45.
- Hof PR, Morrison JH. 2004. The aging brain: Morphomolecular senescence of cortical circuits. *Trends Neurosci*. 27:607–613.
- Jäncke L, Sele S, Liem F, Oschwald J, Merillat S. 2020. Brain aging and psychometric intelligence: a longitudinal study. *Brain Struct Funct*. 225:519–536.
- Jenkinson M, Bannister P, Brady M, Smith S. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*. 17:825–841.
- Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. 2012. *FSL*. *Neuroimage*. 62:782–790.
- Jespersen SN, Kroenke CD, Østergaard L, Ackerman JJH, Yablonskiy DA. 2007. Modeling dendrite density from magnetic resonance diffusion measurements. *Neuroimage*. 34:1473–1486.
- Josefsson M, de Luna X, Pudas S, Nilsson L-G, Nyberg L. 2012. Genetic and lifestyle predictors of 15-year longitudinal change in episodic memory. *J Am Geriatr Soc*. 60: 2308–2312.
- Kaup AR, Mirzakhani H, Jeste DV, Eyer LT. 2011. A review of the brain structure correlates of successful cognitive aging. *J Neuropsychiatry Clin Neurosci*. 6–15.
- Keresztes A, Ngo CT, Lindenberger U, Werkle-Bergner M, Newcombe NS. 2018. Hippocampal maturation drives memory from generalization to specificity. *Trends Cogn Sci*. 22(8):676–686.
- Kievit RA, van Rooijen H, Wicherts JM, Waldorp LJ, Kan K-J, Scholte HS, Borsboom D. 2012. Intelligence and the brain: a model-based approach. *Cogn Neurosci*. 3:89–97.
- Kline RB. 2015. *Principles and practice of structural equation modeling*. New York: Guilford publications.
- Kühn S, Düzel S, Eibich P, Krekel C, Wüstemann H, Kolbe J, Mårtensson J, Goebel J, Gallinat J, Wagner GG, et al. 2017. In search of features that constitute an “enriched environment” in humans: associations between geographical properties and brain structure. *Sci Rep*. 7:11920.
- Lerch JP, van der Kouwe AJ, Raznahan A, Paus T, Johansen-Berg H, Miller KL, Smith SM, Fischl B, Sotiropoulos SN. 2017. Studying neuroanatomy using MRI. *Nat Neurosci*. 20:314–326.
- Liem F, Varoquaux G, Kynast J, Beyer F, Kharabian Masouleh S, Huntenburg JM, Lampe L, Rahim M, Abraham A, Craddock RC et al. 2017. Predicting brain-age from multimodal imaging data captures cognitive impairment. *Neuroimage*. 148:179–188.
- Lindenberger U. 2014. Human cognitive aging: corriger la fortune? *Science*. 346:572–578.
- Little TD, Lindenberger U, Nesselrode JR. 1999. On selecting indicators for multivariate measurement and modeling with latent variables: when “good” indicators are bad and “bad” indicators are good. *Psychol Methods*. 4: 192–211.
- Lövdén M, Karalija N, Andersson M, Wählin A, Axelsson J, Köhncke Y, Jonasson LS, Rieckman A, Papenberg G, Garrett DD et al. 2017. Latent-profile analysis reveals behavioral and brain correlates of dopamine-cognition associations. *Cereb Cortex*. 28:3894–3907.
- Morrison JH, Baxter MG. 2012. The ageing cortical synapse: hallmarks and implications for cognitive decline. *Nat Rev Neurosci*. 13:240.
- Nevalainen N, Riklund K, Andersson M, Axelsson J, Ögren M, Lövdén M, Lindenberger U, Bäckman L, Nyberg L. 2015. COBRA: a prospective multimodal imaging study of dopamine, brain structure and function, and cognition. *Brain Res*. 1612:83–103.
- Nyberg L, Lindenberger U. 2020. Brain maintenance and cognition in old age. In: Poeppel D, Mangun G, Gazzaniga M, editors. *The cognitive neurosciences*. 6th ed. Cambridge (MA): MIT Press.
- Nyberg L, Pudas S. 2019. Successful memory aging. *Annu Rev Psychol*. 70:219–243.
- Nyberg L, Lövdén M, Riklund K, Lindenberger U, Bäckman L. 2012. Memory aging and brain maintenance. *Trends Cogn Sci*. 16:292–305.
- Oertzen v T, Brandmaier AM, Tsang S. 2014. Structural equation Modeling with Ω nyx. *Struct Equ Modeling*. 22:148–161.
- Pantel J, Kratz B, Essig M, Schroder J. 2003. Parahippocampal volume deficits in subjects with aging-associated cognitive decline. *Am J Psychiatry*. 160:379–382.
- Persson J, Pudas S, Lind J, Kauppi K, Nilsson LG, Nyberg L. 2012. Longitudinal structure-function correlates in elderly reveal MTL dysfunction with cognitive decline. *Cereb Cortex*. 22:2297–2304.

- Persson N, Persson J, Lavebratt C, Fischer H. 2017. Effects of DARPP-32 genetic variation on prefrontal cortex volume and episodic memory performance. *Front Neurosci.* 11:289–298.
- Peters A, Kemper T. 2012. A review of the structural alterations in the cerebral hemispheres of the aging rhesus monkey. *Neurobiol Aging.* 33:2357–2372.
- Pierpaoli C, Bassler PJ. 1996. Toward a quantitative assessment of diffusion anisotropy. *Magn Reson Med.* 36:893–906.
- Raz N, Daugherty AM. 2018. Pathways to brain aging and their modifiers: free-radical-induced energetic and neural decline in senescence (FRIENDS) model — a mini-review. *Gerontology.* 64:49–57.
- Raz N, Rodrigue KM. 2006. Differential aging of the brain: patterns, cognitive correlates and modifiers. *Neurosci Biobehav Rev.* 30:730–748.
- Raz N. 2005. The aging brain observed in vivo: differential changes and their modifiers. In: Cabeza R, Nyberg L, Park DC, editors. *Cognitive neuroscience of aging.* Oxford, UK: Oxford University Press.
- Raz N, Lindenberger U, Rodrigue KM, Kennedy KM, Head D, Williamson A, Dahle C, Gerstorf D, Acker JD. 2005. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cereb Cortex.* 15:1676–1689.
- R Core Team. 2019. *R: a language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.
- Ritchie SJ, Booth T, Valdés Hernández M D C, Corley J, Maniega SM, Gow AJ, Royle NA, Pattie A, Karama S, Starr JM et al. 2015. Beyond a bigger brain: multivariable structural brain imaging and intelligence. *Intelligence.* 51:47–56.
- Rosseel Y. 2012. Lavaan: an R package for structural equation modeling. *J Stat Softw.* 48:1–36.
- Rönnlund M, Nyberg L, Bäckman L, Nilsson LG. 2005. Stability, growth, and decline in adult life span development of declarative memory: cross-sectional and longitudinal data from a population-based study. *Psychol Aging.* 20:3–18.
- Rubin DB. 1976. Inference and missing data. *Biometrika.* 63:581–592.
- Rugg MD, Johnson JD, Park H, Uncapher MR. 2008. Encoding-retrieval overlap in human episodic memory: a functional neuroimaging perspective. In: Sossin WS, Lacaille J-C, Castellucci VF, Belleville S, editors. *Progress in brain research.* Vol. 169. Elsevier, p. 339–352.
- Rugg MD, Vilberg KL. 2013. Brain networks underlying episodic memory retrieval. *Curr Opin Neurobiol.* 23:255–260.
- Ruigrok ANV, Salimi-Khorshidi G, Lai M-C, Baron-Cohen S, Lombardo MV, Tait RJ, Suckling J. 2014. A meta-analysis of sex differences in human brain structure. *Neurosci Biobehav Rev.* 39:34–50.
- Schafer JL, Graham JW. 2002. Missing data: our view of the state of the art. *Psychol Methods.* 7:147–177.
- Schaie KW, Maitland SB, Willis SL, Intrieri RC. 1998. Longitudinal invariance of adult psychometric ability factor structures across 7 years. *Psychol Aging.* 13:8–20.
- Schermelleh-Engel K, Moosbrugger H, Müller H. 2003. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol Res.* 8:23–74.
- Seiler S, Ropele S, Schmidt R. 2014. Magnetization transfer imaging for in vivo detection of microstructural tissue changes in aging and dementia: a short literature review. *J Alzheimers Dis.* 42:S229–S237.
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE et al. 2004. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage.* 23(S1):208–219.
- Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM et al. 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage.* 31(4):1487–1505.
- Squire LA. 1992. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol Rev.* 99:195–231.
- Stern Y. 2009. Cognitive reserve. *Neuropsychologia.* 47:2015–2028.
- Strobl C, Malley J, Tutz G. 2009. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods.* 14:323–348.
- Sundgren PC, Dong Q, Gómez-Hassan D, Mukherji SK, Maly P, Welsh R. 2004. Diffusion tensor imaging of the brain: review of clinical applications. *Neuroradiology.* 46:339–350.
- Timmler S, Simons M. 2019. Grey matter myelination. *Glia.* 67:2063–2070.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage.* 15:273–289.
- Waehnert MD, Dinse J, Weiss M, Streicher MN, Waehnert P, Geyer S, Turner R, Bazin P-L. 2014. Anatomically motivated modeling of cortical laminae. *Neuroimage.* 93:210–220.
- Walhovd KB, Westlye LT, Amlien I, Espeseth T, Reinvang I, Raz N, Agartz I, Salat DH, Greve DN, Fischl B et al. 2011. Consistent neuroanatomical age-related volume differences across multiple samples. *Neurobiol Aging.* 32:916–932.
- Ward AM, Mormino EC, Huijbers W, Schultz AP, Hedden T, Sperling RA. 2015. Relationships between default-mode network connectivity, medial temporal lobe structure, and age-related memory deficits. *Neurobiol Aging.* 36:265–272.
- Wolff SD, Balaban RS. 1989. Magnetization transfer contrast (MTC) and tissue water proton relaxation in vivo. *Magn Reson Med.* 10:135–144.
- Yassa MA, Stark CEL. 2011. Pattern separation in the hippocampus. *Trends Neurosci.* 34:515–525.
- Yates KF, Sweat V, Yau PL, Turchiano MM, Convit A. 2012. Impact of metabolic syndrome on cognition and brain: a selected review of the literature. *Arterioscler Thromb Vasc Biol.* 32:2060–2067.
- Yonelinas AP. 1994. Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *J Exp Psychol Learn.* 20:1341.