Hippocampal maturity promotes memory distinctiveness in childhood and adolescence

Attila Keresztesa, Andrew R. Bendera, Nils C. Bodammera, Ulman Lindenbergerab, Yee Lee Shingbcd, and Markus Werkle-Bergnera,b

*aCenter for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, 14195, Germany; bMax Planck University College London Centre for Computational Psychiatry and Ageing Research, Berlin, 14195, Germany; and cDivision of Psychology, University of Stirling, Stirling, FK9 4LA, United Kingdom

Edited by Daniel L. Schacter, Harvard University, Cambridge, MA, and approved July 10, 2017 (received for review June 14, 2017)

Adaptive learning systems need to meet two complementary and partially conflicting goals: detecting regularities in the world versus remembering specific events. The hippocampus (HC) keeps a fine balance between computations that extract commonalities of incoming information (i.e., pattern completion) and computations that enable encoding of highly similar events into unique representations (i.e., pattern separation). Histological evidence from young rhesus monkeys suggests that HC development is characterized by the differential development of intrahippocampal subfields and associated networks. However, due to challenges in the in vivo investigation of such developmental organization, the ontogenetic timing of HC subfield maturation remains controversial. Delineating its course is important, as it directly influences the fine balance between pattern separation and pattern completion operations and, thus, developmental changes in learning and memory. Here, we relate in vivo, high-resolution structural magnetic resonance imaging data of HC subfields to behavioral memory performance in children aged 6–14 y and in young adults. We identify a multivariate association of age-related differences in intrahippocampal subfields and show that HC maturity as captured by this pattern is associated with age differences in the differential encoding of unique memory representations.

Many years ago, the Swiss developmentalist Jean Piaget noted an imbalance between assimilation and accommodation during early and middle childhood in the sense that children tend to extract schematic knowledge at the expense of learning and recollecting specific events (1, 2). This imbalance has resurfaced in computational models of memory (3), and later an imbalance between pattern completion and pattern separation, processes linked to computational properties of subfields within the hippocampus (HC) (4–6). Understanding the developmental organization of HC subfields is therefore crucial to understand how associated changes in HC-subfield computations drive concomitant changes in learning and memory.

An important step toward unraveling controversies about human hippocampal maturation (7, 8) is to acknowledge that the HC is not a homogeneous structure, but rather is composed of cytoarchitectonically and functionally distinct subfields (9). The availability of high-resolution, in vivo magnetic resonance imaging (MRI) of the HC permits the study of specific contributions of different HC subfields in humans (10–12). Computational and rodent models of HC function and high-resolution MRI studies in humans have sought to establish the contributions of individual HC subfields to specific mnemonic functions. For example, the dentate gyrus (DG) has been closely linked to pattern separation (6). Developmental findings from animal models (13) and initial evidence from human studies (14) suggest that the DG matures later than other HC subfields. Likewise, memory functions associated with pattern separation, such as recollection (6), show a protracted course of development that extends well into middle childhood (15). Thus, the DG is a candidate region of interest (ROI) for investigating developmental associations between HC and pattern separation.

However, a sole focus on DG is not warranted. Hippocampal subfields are intricately interconnected (13, 16), and their independent demarcation on MRI images remains imperfect (17). Moreover, extant high-resolution MRI studies in human samples are inconsistent in assigning specific HC computations to specific subfields. For instance, pattern separation has been linked not only to DG, but also to the adjacent area 3 of Cornu Ammonis (CA3; e.g., refs. 10 and 18), entorhinal cortex (EC; ref. 19) and in some cases, data have suggested links to the subiculum (Sub; ref. 14). Thus, it appears oversimplified to assign computations, such as pattern separation and completion, to specific parts of the HC in a one-to-one manner (6). Rather, the HC network may be relatively biased toward more pattern separation or more pattern completion, reflecting differential contributions of its constituent parts (6, 10). In a similar vein, based on domain-specific pattern separation signals within the human EC, Reagh and Yassa (19) suggested a conceptual model of interference resolution in the medial temporal lobe (MTL), whereby incremental decrease in representational overlap is reached by pattern separation in domain-specific, parallel pathways already upstream of DG, including EC.

In sum, there is a clear tension between the desire to link structure and function at the level of specific subfields to behavior and the presence of massive connections and interactions within the HC network. Moreover, existing animal and human data suggest that all HC subfields undergo maturational changes during early development (13, 14, 20), albeit along different trajectories. Both of these observations call for a multivariate approach to investigate HC development during childhood and adolescence.

Significance

Children tend to extract schematic knowledge at the expense of learning and recollecting specific events. Our findings allow us to speculate that the heterogeneous development of subregions within the hippocampus—a brain region crucial for laying down novel memories—contributes to this developmental lag in memory. Specifically, we used in vivo high-resolution structural MRI and memory tests in a large sample of children aged 6–14 years and young adults to characterize hippocampal development. We show that hippocampal maturity as expressed in the multivariate pattern of age-related differences in hippocampal subregions is specifically related to the ability to lay down highly specific memories.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1Y.L.S. and M.W.-B. contributed equally to this work.

2To whom correspondence should be addressed. Email: werkle@mpib-berlin.mpg.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1710654114/-/DCSupplemental.
A specific multivariate profile of HC subfields is associated with age.

Methods

We used multivariate statistical techniques to estimate individual maturation profiles of HC anatomy and examined the association between HC maturity and different behavioral measures previously associated with HC. The behavioral measures included a bias score for pattern separation versus pattern completion as the primary target of investigation. In addition, we also included indicators of age-sensitive mnemonic processes such as source memory, associative memory, and item memory that also rely on extra-hippocampal areas (21, 23–26). The inclusion of the latter measures was exploratory and served the purpose of probing the specificity of a potential association between pattern separation/completion and HC subfield maturation. For instance, previous studies reported age-related differences in functional and structural HC contributions to source memory (27, 28) and associative inference (29) along the longitudinal HC axis. However, the relationship between age differences in source memory and HC subfield development remains elusive.

We expected to replicate initial evidence for the relatively late maturation of the DG (14). In addition, we reasoned that individual differences in a multivariate index of HC maturity would predict individual differences in processes that support the specific encoding of unique events such as pattern separation. Furthermore, we expected that this index of HC maturity would be only weakly associated with, or unrelated to, memory processes that rely on early maturing aspects of the HC, such as familiarity, or late-maturing memory processes that are less exclusively HC-dependent and also heavily dependent on extra-hippocampal areas, such as source memory (23).

Results and Discussion

Age-Related Differences in Hippocampal Structures Suggest Differential Development That Extends Well into Middle Childhood and Beyond. In an initial set of analyses, we examined whether individual HC subregions (Methods) show evidence for maturation across childhood. We identified age-related neuroanatomical differences between 6 and 27 y of age by regressing ROI volumes on age (Fig. 1B). In the total sample, only the ROI including the DG and the CA3 (DG-CA3) subfields showed a significant linear age trend ($R^2_{\text{adjusted}} = 0.05, p_{\text{adj}} = 0.023$). Adding a quadratic term revealed a significant age trend in Sub ($R^2_{\text{adjusted}} = 0.14, p_{\text{adjlinear}} < 0.001, p_{\text{adjquadratic}} < 0.001$). When the same analyses were restricted to children, we found significant linear age trends for DG-CA3 ($R^2_{\text{adjusted}} = 0.11, p_{\text{adj}} = 0.005$), Sub ($R^2_{\text{adjusted}} = 0.17, p_{\text{adj}} < 0.001$), as well as CA1-2 ($R^2_{\text{adjusted}} = 0.08, p_{\text{adj}} = 0.02$). F tests on $\Delta R^2$ values showed that adding quadratic terms did not result in significant increments in explained variance.

These univariate results suggest a protracted development of HC subfields, including the CA1-2, DG-CA3, and the Sub, until late middle childhood and early adolescence. In addition, they also show age-related changes in DG-CA3 volume until young adulthood and an onset of volume decrements in Sub around adolescence. The results are in partial agreement with one earlier study by Lee et al. (14) who investigated age differences in HC subfields in four arbitrarily defined age groups in a sample of 8- to 14-y-old children and observed significant quadratic age trends in DG-CA3 and CA1-2. Differences in results may reflect subtle variations in tracing protocols between the two studies [see SI Methods in relation to the delineation of CA1-2 and Sub (cf. ref. 17)], the grouping of the individuals into age groups by Lee et al., or both. Despite these differences, the two studies provide converging evidence for a protracted maturation of HC subfields through middle childhood. This pattern of maturation is followed by later volumetric reductions that may extend into young adulthood, at least for the DG-CA3.

A Specific Multivariate Profile of HC Substructures Is Associated with Age. The HC subfields form a highly interconnected hard-wired processing circuitry (16). Therefore, if maturation would only affect a single substructure (with no effect on others), it would
most likely harm the fine balance between pattern separation and completion operations required for a flexible and adaptive memory system (6). Age-graded changes in the balance between pattern separation and completion are likely to result from a multivariate pattern of subfield changes. Therefore, we examined the age-graded link of HC-subfields maturity to different memory processes from a multivariate perspective. Using partial least squares correlation analysis (PLSC), we extracted a single composite score that captures individual differences in the structural maturity of HC subfields. For simplicity, we refer to this score as “HC-maturity score” (Methods). Our PLSC analysis identified a single reliable latent variable (LV; \( P = 0.038 \)) that optimally represents the association between participants’ age and ROI volumes (\( r = 0.29 \)). Bootstrap ratios (BSR) indicated an age-associated increase of DG-CA3 (BSR = 2.3) volume, and a decrease of EC volume (BSR = −1.98) as the two stable components of the LV expressing the largest amount of information common to both age and the multivariate pattern of ROI volumes (Fig. 2).

Several previous reports (8, 30, 31) have suggested that the HC and associated memory functions reach maturity by middle childhood. In contrast, other studies have provided evidence for prolonged maturation of HC-dependent memory process until adolescence (7). In our view, these apparent contradictions can be overcome by acknowledging the heterogeneous course of HC maturation (e.g., refs. 13, 14, and 32). In particular, in light of the subfields’ different maturational trajectories, HC subfield data can yield a more fine-grained picture of age-graded volumetric differences than total HC volume can. With a whole HC analysis run on standard resolution MRI data, subtle maturational effects detectable by high-resolution MRI derived subfield data may go unnoticed. To check this claim empirically, we ran two additional analyses. First, to mimic whole HC analyses, we aggregated DG-CA3, CA1-2. Sub volumes, and computed an analogous PLSC with age, whole HC, and EC. In contrast to the original analysis based on subfield volumes, the analysis with age, whole HC, and EC failed to extract a significant and generalizable latent variable (\( P = 0.17 \)). Second, a voxel-based morphometry (VBM) analysis revealed widespread age-related differences in gray matter volume (GM) over the cortex, but not in the MTL (see Fig. S2 for results and description of the VBM methods). In sum, previous research may have failed to find age-related differences in HC during middle childhood because their total HC target measure collapsed regions with heterogeneous maturational trajectories and/or did not include EC.

Multidimensional Structural Maturity of the Hippocampal Formation Is Associated with Memory Processes Enabling the Unique Encoding of Similar Representations. Next, we assessed the association between individuals’ HC-maturity scores and memory performance. Memory development across childhood is characterized by an overall improvement of mnemonic functions (21, 24). Nevertheless, the developmental timing and interdependence of different mnemonic operations remains controversial (33–36), especially in relation to age, whole HC, and EC. In contrast to the aforementioned, in this present study, we comprehensively assessed memory processes potentially associated with HC maturity: pattern separation/completion, source memory, item, and associative memory.

A mnemonic similarity task adapted from (10, 37) was used to behaviorally assess pattern separation versus pattern completion bias (Fig. 3A and SI Methods). From this task, we computed a pattern separation/completion bias index that expresses the degree to which mnemonic similarity judgments are biased toward pattern separation (or against pattern completion) (10, 37). Several studies have corroborated the suggestion that this index is a reasonable estimate of the relative strength of HC pattern separation (37–39). Using two alternating contexts during learning blocks of the same task, we also assessed source memory. A second task adapted from Naveh-Benjamin et al. (40) was used to assess item and associative memory (Fig. 3C and SI Methods).

To assess the dependence of each memory process on hippocampal maturity, we ran bivariate correlation analyses between the PLSC-derived HC maturity scores and the behavioral indicators. As predicted, the pattern separation/completion bias score correlated positively with the HC maturity score (\( r = 0.26, P = 0.013 \)), revealing a moderate shift toward pattern separation with increasing HC maturity (Fig. 3B). In addition, false recognition of item memories (Methods and Fig. 3C) showed a significant negative association (Fig. 3D), \( r = −0.33, P < 0.001 \), with HC maturity. No other memory measure revealed significant associations with HC-maturity scores (Table S1). Importantly, the strength of correlations between HC maturity and pattern separation/completion and between HC maturity and false recognition of item memories were significantly stronger than the nonsignificant correlations between HC maturity and the other memory measures [indicated by a significant contrast among correlated correlation coefficients (41)], \( z = 2.67, P = 0.004 \).

Both mnemonic similarity judgments and the rejection of foils in a recognition memory task involving highly similar items crucially depend on the orthogonalization of overlapping feature sets in representational space. Therefore, our results suggest that the multidimensional maturity of structures in the HC is specifically related to processes that enable the construction of unique
mnemonic representations of highly overlapping feature sets during memory encoding. Conversely, the present results also suggest that age-associated differences in item memory, source memory, and associative memory performance (Fig. S1) depend less on HC maturity than the age-associated changes in the disambiguation of highly similar events. Clearly, performance on item memory, source memory, and associative memory relates to hippocampal functioning (42). However, the demand characteristics of these tasks, under most conditions at least, presumably depend less on pattern separation than the demand characteristics of making mnemonic similarity judgments and rejecting highly similar foils. In addition, source memory and associative memory are likely to require prefrontally mediated control processes such as monitoring during source memory decisions (23) and inhibition of combinations of familiar items (25) during old/new decisions in an associative memory task (26). This enhanced prefrontal dependence may inject additional age-related variance into task performance, which may weaken or mask potential associations with HC maturity. To test these assumptions, we applied the PLSC approach used to construct the HC-maturity score to also establish a maturity score for frontal ROIs based on GM measures obtained from our VBM analyses (see Fig. 4 for results, and see Methods and Fig. S2 for methods). In support of our considerations, the frontal maturity score correlated significantly with source memory \( (r = 0.26, P = 0.009) \) but not with any other memory measure except for hits in item memory \( (r = 0.40, P < 0.001; \) Table S1). The strength of the two significant correlations significantly differed from the strength of the nonsignificant ones \( (z = 2.95, P = 0.002) \).

**General Discussion**

Using multivariate correlational techniques on high-resolution structural MRI data of the MTL in a sample of 6- to 27-y-old individuals, we identified a multivariate profile of developmental differences in HC substructures that expresses the structural maturity of the HC. We then showed that HC maturity is specifically related to the development of memory processes promoting the unique encoding of overlapping memory representations. Our results suggest that key contributors of this specific connection between HC maturity and memory are age-associated changes in the DG-CA3 and the EC. HC maturity scores did not reveal a robust association with any of the other memory measures, although these measures also showed age-associated improvements (Fig. S1). Compared with the mnemonic similarity task, our additional recognition measures (item and associative recognition memory and source memory) apparently were less sensitive to shifts in pattern separation/completion bias. The mnemonic similarity task has been specifically designed to assess this bias on a continuous scale between separation and completion (37), whereas performance on the other memory measures may more heavily depend on extrahippocampal areas not incorporated in our HC-maturity score (23, 25, 26).

Our observation that the association between HC maturity and memory is restricted to age-related increases in specificity may reflect one or both of the following underlying processes. First, the development of memory processes that require less specificity with regard to unique feature combinations may depend more strongly on age-related changes in extrahippocampal areas. As discussed above, maturation of prefrontal cortex can, in part, drive improvements in both associative recognition memory and source memory (23, 25, 26), possibly moderated by increases in demands on strategic processes rather than associative memory operations (34). This conjecture is in part supported by our analyses showing that maturation of frontal areas was significantly correlated with source memory, but not with pattern separation/completion bias. However, we should note that, based on standard-resolution MRI, some studies have found age differences in the functional division along the longitudinal axis of the HC (27, 28) that may also contribute to age differences in source memory ability (43). Second, pattern completion may be relatively mature by middle childhood despite ongoing structural changes in HC, whereas computations underlying specificity are still developing, thus promoting the observed age-graded shift in bias from pattern completion to pattern separation.

Our results complement earlier findings (14) demonstrating age-associated differences in HC subfields in middle childhood and extend those observations to a large sample of children aged 6 to 14 y. In addition, we provide an initial picture of HC subfield development in middle childhood. This picture highlights the presence of subfield-specific, heterogeneous maturational tracks. By demonstrating that estimates of whole HC volumes failed to detect age-associated differences, our study also help to resolve conflicting observations, with some studies suggesting that HC maturation levels off early in middle childhood (8, 13, 31, 34) and others suggesting that HC maturation extends well into adolescence, and possibly beyond, this period (7). Previously available standard resolution MRI techniques may not be sensitive enough to reveal extended HC maturation.

Our study revealed effects that complement earlier studies linking the DG and CA3 to pattern separation (6, 10, 18). Beyond the crucial role of the DG-CA3 region for providing separable inputs to downstream HC subfields, the development of memory specificity appears associated with a common maturational process that potentially affects all HC subfields to varying degrees. Our finding that EC development is a key component of HC maturity associated with pattern separation fits nicely with observations in animals that layer 2 and 5 of the lateral EC follows DG development (20), and with human data suggesting that lateral EC may perform pattern separation on overlapping object representations before passing its input onto the DG (19). It is worth noting that EC by itself did not show significant age-related differences in the present sample. The contribution of EC to HC maturity was revealed only when applying a multivariate approach that expresses the common variance between individual differences in HC subfield volumes and age. Methodologically, our approach follows the longstanding claim to conceptualize and analyze developmental change from a multivariate perspective (44). Earlier work has shown that multivariate composites of individual differences in
brain anatomy can serve as a summary description of biological maturity (45). The dimensionality reduction associated with these methods helps to test and refine theories of age-related changes in brain–behavior relations.

The present study has several limitations, which can guide future research in the field. Given that development is a process unfolding in ontogenetic time, repeated within-subjects assessments are needed to directly capture longitudinal relationships between neural and behavioral variables of interest (46). For this reason, we refrained from using hierarchical linear regression models with age as independent variable, memory processes as dependent variables, and HC subfields as mediator variables. It has been shown analytically that these methods may fail to detect longitudinal mediation when it is present (false negatives) and detect mediation when it is absent (false positives) (47; see also ref. 48). A second limitation is related to the bias score used in this study, which pits pattern separation against pattern completion. Future studies need to obtain measures that separately index age differences in the efficiency of pattern separation and pattern completion mechanisms. The restriction of our analyses to HC body is a third limitation. Previous studies found both structural and functional age-related differences in source memory contributions of the HC head and tail, but not the body (27, 28). Investigating subfield contributions along the full anatomical extent of HC could therefore refine our understanding of how HC subfield and memory development are related (see ref. 29; however, this study also highlights the controversies regarding methods for identifying HC subfield in the head and tail, see ref. 49). Fourth, recent fMRI findings suggest that pattern separation may not be restricted to the HC (50, 51). In the present study, we selected a task that aims at studying age differences in pattern separation performed by the HC, but one that is not well-suited for examining pattern separation, and age differences therein, in other brain areas, such as visual cortex. Future studies need to address the maturational course of pattern separation in other brain areas and their contributions to behavioral development. Last, we devised this study to test the suggestion that HC and related mnemonic functions may develop beyond the onset of middle childhood, but had no a priori reason to postulate that this development may continue beyond middle childhood. Therefore, we did not include individuals aged 15–18 y in the present sample. Also, we did not include children below 6 y of age, reflecting practical limitations when conducting MRI studies with young children. Our results should encourage future research to explore HC subfields and related mnemonic development in a more extended age range.

We found that age-related shifts from pattern completion toward pattern separation are associated with maturational changes in HC subfields. If corroborated by longitudinal evidence from tasks directly measuring some form of knowledge extraction from invariances (i.e., category learning), this result has fundamental implications for theories of episodic memory development: It leads us to speculate that the extraction of invariance across a range of different experiences may precede the encoding, consolidation, and retrieval of detail for reasons that are rooted, at least in part, in the uneven maturational course of substructures within the HC.

Returning to Piaget (1, 2), we conclude that this décalage, or developmental lag may be developmentally advantageous, as it helps children to recognize regularities, form stable representations of recurring episodes, predict the structure of future events, and build semantic knowledge.

Methods
Participants. Seventy children (35 girls; age range: 6–14 y; M = 9.80, SD = 2.39 y), and 33 young adults (18 women; age range: 18–27 y; M = 23.21, SD = 2.5 y) participated in the study. Participants provided written informed consent, also signed by the primary caregiver for all children. Participants were right-handed and had no history of neurological or psychiatric disorders. They completed the study in two sessions lasting 2 h each and were paid 40 €. Behavioral data were not available for one child and one young adult because of technical issues. The Ethics Committee of the German Psychological Society (Deutsche Gesellschaft für Psychologie) approved the study.

Delineating ROIs in the MTL. Four ROIs were manually demarcated bilaterally (Fig. 1A) by two expert tracers on coronal slices of the high-resolution structural MR volume (voxel size: 0.4 mm × 0.4 mm × 2 mm). The segmentation protocol included three ROIs (Sub, CA1-2, and DG-C3A) segmented along the full range of the HC body, excluding head and tail, and the EC segmented in six slices anterior to the HC body. Bilateral ROIs were collapsed across hemisphere and adjusted for intracranial volume (ICV) for all following analyses (SI Methods).

Assessing the Multivariate Relationship Between ROI Volumes and Age Using PLSC. We chose to use PLSC (52, 53) to assess HC maturation on conceptual and retrieval of detail for reasons that are rooted, at least in part, in the uneven maturational course of substructures within the HC.

Approximations of robustness (Fig. 2C). BSRs are comparable to conventional z values, where a value larger/smaller than ±1.96 is treated as reliably robust. We also obtained a summary measure of each participant’s robust expression of the estimated LVs’s profile, a within-person HC-maturity score, by multiplying the model-based vector of ROI weights (V) by each subject’s vector of ROI volume estimates (Q), producing a single within-subject value, the HC-maturity score = VQ (Figs. 2C and 3 B and D).

A comparable procedure was used to derive “frontal maturity score” for frontal analyses. Control of volumetric measures of HC subfields, we used VBM derived gray matter volume (GM) estimates in six frontal ROIs of the lpba40 Atlas (54). Results of this analysis are presented in Fig. 4. Whole-brain PLSC analyses, probing the associations between GM and age are reported in Fig. S1.

Voxel-Based Morphometry. We used the standard preprocessing pipeline of the CAT12 toolbox (dbm.neuro.uni-jena.de/cat) run in SPM12 (Wellcome Trust Centre for Neuroimaging, www.fil.ion.ucl.ac.uk/spm) (version 6906) to obtain voxelwise and ROI specific GM estimates (see Fig. S2 for details). For ROI analyses (Fig. 4), GM estimates were collapsed across hemispheres, and ICV corrected using the same approach as for our HC analysis (SI Methods).

Acknowledgments. We thank the students and technical assistants of the Cognitive and Neuronal Dynamics of Memory across the Lifespan (CONMEM) project for their support in data collection; the participants for their cooperation; and Julia Delius, Yana Fandakova, Naftali Raz, Michael D. Rugg, and Myriam C. Sander for valuable discussions. This study was conducted within the project CONMEM at the Center for Life and Mind, Psychologische Forschungsinstitute Max Planck (MPIB). The research was partially financed by the Max Planck Society and by the 2010 Leibniz award of the German Research Foundation (U.L.). M.W.-B.’s work is supported by German Research Foundation Grant WE 4286/3-1 (Y.L.S., as co-principal investigator) and an Early Career Research Fellowship 2017–2019 awarded by the Jacobs Foundation. A.K.’s position is supported by the MPIB, and a Minerva Research Group was awarded by the Max Planck Society (to Y.L.S.).

9216 | www.pnas.org/cgi/doi/10.1073/pnas.1710654114
Keresztes et al.
Supporting Information

Keresztes et al. 10.1073/pnas.1710654114

SI Methods

Imaging Data Acquisition and Preprocessing. High-resolution images of the MTL were acquired by using a T2-weighted turbo spin echo sequence on a 3 T Siemens Magnetom TrioTim syngo MRI scanner. Two separate images [field of view (FOV): 206 mm; repetition time (TR): 6,500 ms; echo time (TE): 16 ms; number of slices: 30; voxel size: 0.4 mm × 0.4 mm × 2 mm] were acquired in an oblique direction, perpendicular to the longitudinal axis of the right HC to cover the full bilateral HC. The two images were coregistered by using FLIRT, Oxford Centre for Functional MRI of the Brain’s Linear Image Registration Tool (55), and averaged to provide one high-resolution structural image with improved quality for manual tracing.

Delineating ROIs in the MTL. Subfields were traced on all volumes by both tracers (A.K., and A.R.B.) using Analyze 11.0 (AnalyzeDirect). The segmentation protocol included three ROIs (Sub), a region covering CA1 and CA2 (CA1-2), and a region covering the dentate gyrus and CA3 (DG-CA3) segmented along the full range of the HC body, excluding head and tail. An additional ROI in the EC was segmented in six slices anterior to the HC body, ending on the first HC body slice on which the other regions were traced. First, one tracer traced half of the slices for a given volume, and then the other tracer traced the remaining slices. Slices within volumes and order of tracers were randomized, and both raters traced 50% of slices per volume and traced as first tracers for 50% of volumes. Raters were blind to participant age. In demarcation of the ROIs, we followed suggestions by refs. 32 and 56 with slight modifications. Specifically, in comparison with Shing et al. (58), we placed the boundary between Sub and CA1-2 more laterally compared with the most medial point of DG-CA3. For the current study, the placement of this boundary separating Sub and CA1-2 was defined as half of the distance between the most medial aspect of the DG-CA3 and the most lateral aspect of CA1-2, and instead of only 3 slices for DG-CA3, CA1-2, and Sub, we traced the full body. We defined ranges for the body similarly to Lee et al. (14). Intraclass correlation coefficients [ICC (2)] showed high interrater reliability for all ROIs overall volume bilaterally [ICC (2) > 0.85], and collapsed across hemisphere [ICC (2) > 0.9]. Bilateral ROIs were collapsed across hemisphere for all following analyses. To account for differences in ROI volumes because of differences in head size, we used the analysis of covariance approach (59, 60) with a slight modification to avoid neglecting age effects on ROIs. We adjusted all ROI volumes in two steps. First, we regressed intracranial volume (ICV) on age to calculate an individual age-predicted ICV. Second, we adjusted each ROI volume by correcting for differences in ROIs volume due to differences in actual and predicted ICV. The adjusted volumetric data are used for all ROIs throughout the present report. ICV estimates were obtained by using the brain extraction tool in FSL 5.0 (61) using procedures described in Bender et al. (62).

Behavioral Procedures and Measures. In the mnemonic similarity task (Fig. 3d) adapted from Bakker et al. (10) and Stark et al. (37), participants incidentally encoded a series of pictures depicting objects drawn from several online databases (10, 37, 63) in a 3 T Siemens Tim-Trio scanner (BOLD data acquired during scanning is not presented in this article). They saw 432 pictures, each for 2,000 ms (with an intertrial interval jittered with a left-skewed distribution between 1,000 and 2,500 ms), with 144 objects shown only once, 72 objects repeated once with exactly the same picture, and 72 objects repeated with a slightly different picture. Repetition lags for repeated objects were 15, 20, 25, and 30 trials. The encoding phase was organized in six blocks with two contexts alternating for a later source memory task. Participants were instructed to decide, within a 3,000-ms response window, whether the displayed objects did or did not fit into one of two placeholders (a treasure box and a water bucket) that were consistently matched to one of the two contexts (a “fire” planet, and a “water” planet). Contexts were presented by a short video before each block. Out of the score of ±15 min later, participants performed a surprise recognition test that involved making mnemonic similarity judgments. Again, they saw a series of pictures depicting objects, 108 targets (i.e., same objects with exactly the same picture), 108 lures (i.e., same objects with a slightly different picture), and 108 foils (i.e., new objects), and their task was to identify targets, highly similar lures, and novel items (foils), choosing from one of three response options (old, similar, or new). Each trial was shown for 1,000 ms, with an ITI of 2,000 ms for young adults, and 5,000 ms for children, with a total response window of 3,000 ms and 6,000 ms, respectively. Assignment of pictures to conditions and blocks to contexts, and order of presentation, was pseudorandomized for each participant with the above constraints.

The pattern separation/completion bias score was calculated by subtracting the proportion of similar responses to foils from the proportion of correct similar responses to lures that indexed the ability the separate highly similar memories from one another (10). A bias score closer to zero would suggest a bias toward pattern completion, whereas a score closer to one would suggest a bias toward separation. A given score does not allow inferring whether a given value means more pattern separation than completion; however, the difference in two scores suggests differential biases between separation and completion. Therefore, the pattern separation/completion bias score can be used to assess individual differences in the balance between pattern separation and completion.

Trials that were responded either similar or old were followed by a source memory decision trial, where participants had to choose the source of the item during encoding from two response options (source 1 vs. source 2) within an unlimited response window. Here, all target and lure items were pooled together to provide an overall accuracy score that served as our source memory measure.

In the faces and names task (Fig. 3c) adapted from Naveh-Benjamin et al. (40) participants incidentally encoded pictures of faces together with an auditorily presented name. During encoding (36 trials, 3,000 ms each, with a 500 ms ITI) participants saw 18 male and 18 female faces, presented with corresponding high-frequency names. They were instructed to report their subjective decision on whether the name did or did not fit the face. After a 1-min delay, participants performed a surprise recognition test that consisted of three phases, presented in a random order: an item test for faces, an item test for names, and an associative test for face-name pairs. Pairs presented at encoding were randomly assigned to three sets of 12 (3 female and 3 male faces/names). Sets were then randomly assigned to test phases. In each test phase, 12 targets from the assigned set and 12 novel foils were presented, each for 4,000 ms. This design ensured that no item appeared more than once during the test, and also that no item appeared again in a test phase if its paired associate had been presented in another test phase. The associative test included only faces and names presented during encoding that were either paired as during encoding or rearranged. For the item tests, participants were instructed to...
indicate whether an item was old or new. For the associative test, they were instructed to respond old to nonrearranged pairs, and new to rearranged pairs. Results for two item test phases (faces and names) were pooled together for the analyses. We calculated hit rates (percent of target item/pairs correctly identified as old) and false alarm rates (percent of new items/pairs incorrectly identified as old) separately for item and associative recognition.

Processes underlying our derived measures—pattern separation/completion bias, source memory, associative memory, and item memory—have all been linked to hippocampal processing (6, 23, 42). Except for pattern separation, there is also existing data suggesting that these processes undergo some form of development (22). However, the processes vary in their degree of dependence on pattern separation, with our pattern separation/completion bias being the most clearly related to it, and source memory and associative memory—at least in part—being also highly dependent on strategic prefrontal processes during encoding and retrieval (23).

Fig. S1. Related to Fig. 3, univariate analyses reveal a complex picture of age-related differences in various memory processes. Scatterplots showing performance (in percent) on memory scores derived from the two behavioral tasks (Fig. 3 and Methods). Red lines represent linear regression models fit on the relationship of performance with age. Dashed red lines represent 95% confidence intervals. (A) Pattern separation/completion bias showed a marginally significant linear age trend (R² = 0.03, pβ = 0.054), and a significant age trend when a quadratic term was included (R² = 0.1, pβ [linear] = 0.002, pβ [quadratic] = 0.004). (B) Source memory (percent of correct source identifications for items classified as either old or similar) showed a significant linear age trend (R² = 0.09, pβ = 0.001). (C) Hits in item memory recognition (percent of old items correctly identified as old) showed a significant linear age trend (R² = 0.21, pβ < 0.001). Adding a quadratic term to the regression (R² = 0.24, pβ [linear] = 0.005, pβ [quadratic] = 0.03) increased fit significantly. (D) False alarms in item memory recognition (percent of new items incorrectly identified as old) showed a significant negative linear age trend (R² = 0.04, pβ < 0.02). (E) Hits in associative memory (percent of old pairs correctly identified as old) showed a significant age trend when both linear and quadratic terms were included (R² = 0.07, pβ [linear] = 0.008, pβ [quadratic] = 0.015). (F) False alarms in associative memory (percent of rearranged pairs incorrectly identified as old) also showed a significant age trend when both linear and quadratic terms were included (R² = 0.04, pβ [linear] = 0.032, pβ [quadratic] = 0.047).

Keresztes et al. www.pnas.org/cgi/content/short/1710654114
Fig. S2. Related to Figs. 1 and 2. Voxel-based morphometry reveals widespread cortical but no mediotemporal areas showing age-related differences in gray matter volume. Gray matter volume estimates were extracted by using the standard preprocessing pipeline of the CAT12 toolbox (dbm.neuro.uni-jena.de/cat) run in SPM12 (Wellcome Trust Centre for Neuroimaging, London, www.fil.ion.ucl.ac.uk/spm) (version 6906). In brief, T1-weighted (FOV: 256 mm, TR: 2,500 ms; TE: 3.69 ms; voxel size: 1 mm × 1 mm × 1 mm) images of each participant were normalized to template space, then segmented into tissue classes of gray and white matter, and cerebrospinal fluid. After checking sample homogeneity, we excluded one brain that had extremely low average pairwise correlation to all other brains (2 SDs below mean) and visual quality check confirmed that the T1 image was too noisy to be used. Resulting images were then smoothed with a kernel of 8 mm (FWHM) by using SPM12. Voxelwise gray matter volumetric estimates were ICV corrected by using the same approach as for our HC analysis (see description in the main text section Delineating ROIs in the MTL). Similarly to the extraction of HC-maturity and frontal maturity scores, we extracted one significant LV ($P < 0.001$) that optimally (in a least-squares sense) represents the associations between age and voxelwise gray matter volume estimates. BSR calculated for the resulting LV for each voxel in the brain was mapped to and is visualized here on MNI space. The scale represents voxelwise BSR. Importantly, BSR values indicate that despite the widespread age-related differences in several frontal and parietal regions, no voxels showed reliable age-related difference in the MTL. Suggesting that total HC measures do not show age-related change in our sample. A similar analysis restricted to bilateral HC ROIs defined by the lpba40 Atlas (54) was performed to further explore any potential age-related differences in total HC. Again, PLSC failed to extract an optimal and generalizable latent variable structure expressing the association between age and HC ($P = 0.23$).
Table S1. Pairwise correlations (Pearson’s r) between volumetric measures, behavior measures, and age, related to Figs. 1–4

<table>
<thead>
<tr>
<th>Variables</th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
<th>9.</th>
<th>10.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Entorhinal cortex volume</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Subiculum volume</td>
<td>0.23* (0.019)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Dentate gyrus–CA3 volume</td>
<td>−0.05 (0.63)</td>
<td>0.31** (0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. CA1–CA2 volume</td>
<td>−0.11 (0.278)</td>
<td>0.29** (0.003)</td>
<td>0.79*** (0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Pattern separation/completion bias</td>
<td>−0.12 (0.25)</td>
<td>−0.01 (0.914)</td>
<td>0.21* (0.04)</td>
<td>0.19 (0.061)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Source memory</td>
<td>−0.02 (0.809)</td>
<td>0.08 (0.457)</td>
<td>0.14 (0.165)</td>
<td>0.09 (0.378)</td>
<td>0.26* (0.013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Hits (item memory)</td>
<td>−0.17 (0.087)</td>
<td>0.03 (0.8)</td>
<td>0.09 (0.361)</td>
<td>0.07 (0.491)</td>
<td>0.28** (0.006)</td>
<td>0.13 (0.191)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. False alarms (item memory)</td>
<td>0.23* (0.021)</td>
<td>0.19 (0.059)</td>
<td>−0.15 (0.138)</td>
<td>−0.24* (0.017)</td>
<td>−0.09 (0.393)</td>
<td>0.06 (0.588)</td>
<td>0.14 (0.178)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Hits (associative memory)</td>
<td>−0.05 (0.635)</td>
<td>0.08 (0.414)</td>
<td>0.2* (0.044)</td>
<td>0.2* (0.046)</td>
<td>0.12 (0.243)</td>
<td>0.04 (0.695)</td>
<td>0.37*** (0.000)</td>
<td>0.03 (0.75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. False alarms (associative memory)</td>
<td>−0.07 (0.514)</td>
<td>−0.14 (0.173)</td>
<td>0.06 (0.535)</td>
<td>−0.01 (0.924)</td>
<td>−0.12 (0.245)</td>
<td>0.05 (0.603)</td>
<td>−0.01 (0.903)</td>
<td>0.32*** (0.001)</td>
<td>0.1 (0.338)</td>
<td></td>
</tr>
<tr>
<td>11. Age, mo</td>
<td>−0.17 (0.086)</td>
<td>−0.12 (0.221)</td>
<td>0.22* (0.024)</td>
<td>0.13 (0.175)</td>
<td>0.2 (0.054)</td>
<td>0.32** (0.001)</td>
<td>0.46*** (0.000)</td>
<td>−0.23 (0.023)</td>
<td>0.17 (0.082)</td>
<td>−0.13 (0.179)</td>
</tr>
<tr>
<td>12. Frontal-maturity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. HC-maturity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*P < 0.05, **P < 0.01, ***P < 0.001. Exact P values in parentheses (reported up to three decimals). All values are uncorrected for multiple comparisons. Correlations between maturity scores and age, and maturity scores and subfields, are reported in the main text (Figs. 2 and 4) but not shown here because they are the result of the PLSC optimization, and are not interpreted.