It's all in the details: Relations between young children's developing pattern separation abilities and hippocampal subfield volumes

Running title: Pattern separation in children

Kelsey L. Canada[a], Chi T. Ngo[b], Nora S. Newcombe[b], Fengji Geng[a], Tracy Riggins[a]

[a] Department of Psychology, University of Maryland, College Park, Maryland, 20742, USA

[b] Department of Psychology, Temple University, Philadelphia, Pennsylvania, 19122, USA

Corresponding Author: Tracy Riggins, Ph.D., Department of Psychology, University of Maryland, College Park, Maryland, 20742, 301.405.5905, riggins@umd.edu

Abstract

The ability to keep similar experiences separate in memory is critical for forming unique and lasting memories, as many events share overlapping features (e.g., birthday parties, holidays). Research on memory in young children suggests their memories often lack high-resolution details, i.e., show impoverished pattern separation. Recently developed assessments of pattern separation suitable for children allow us to relate the formation of distinct, detailed memories to development of the hippocampus, a neural structure critical for this ability in adults. The hippocampus displays a protracted developmental profile and underlies the ability to form detailed memories. This study examined age-related differences in hippocampal subfield volumes in 4- to 8-year-old children and relations with performance on a mnemonic similarity task (MST) designed to index memory specificity. Results revealed age-moderated associations between MST performance and CA2-4/DG subfields. Specifically, age-related differences in the ability to form detailed memories tracked with normative patterns of volume increases followed by reductions over this age range. That is, greater volume correlated with better performance in younger children, whereas smaller volume correlated with better performance in older children. These findings support the hypothesis that developmental differences in hippocampal circuitry contribute to age-related improvements in detailed memory formation during this period.

Keywords

development, early childhood, hippocampus, memory, mnemonic discrimination

Recall of detailed memories from early in life is lower than predicted by normal forgetting functions and is one aspect of the phenomenon of childhood amnesia (Hamond and Fivush 1991; Rubin 2000; Peterson et al. 2017). Many investigators have proposed that this lack of details for early childhood memories is due to the protracted development of the hippocampus (HPC; Bauer 2006; Josselyn and Frankland 2012; Lavenex and Lavenex 2013). However, empirical data from human children to support these theories is lacking. HPC is a heterogeneous subcortical structure within the medial temporal lobe (MTL) that consists of the dentate gyrus (DG), cornu ammonis (CA) 1-4, and the subiculum (Sub) (Duvernoy 1998; Insausti et al. 1998; Ding and Van Hoesen 2015). These subfields differ in their developmental trajectories, with the most protracted development thought to occur in DG and its connections to CA3 (Lavenex and Lavenex 2013; Riggins et al. 2018).

Keeping similar experiences separate in memory is one critical element in successful autobiographical memory because many life events share overlapping features (e.g., birthday parties, holidays). Computational theories posit that memory systems overcome potential interference using pattern separation (PS), a neural computation in which similar memories are assigned distinct representations during encoding, thus reducing the overlap between similar inputs (Complementary Learning Systems, Norman and O'Reilly 2003). Recent research has suggested that this computation is a function of DG and, in some circumstances, CA3 (Bakker et al. 2008; Neunuebel and Knierim 2014). Although PS cannot be measured directly in humans, it has been suggested that behavioral tasks that place a high demand on mnemonic discrimination between similar stimuli may serve as an index of this computation (e.g., Kirwan and Stark 2007; Lacy et al. 2011; Stark et al. 2013). These tasks are variants of traditional recognition memory tasks, with the inclusion of perceptually similar exemplars of the studied items as lures at test.

The degree to which an individual can discriminate between studied items and lures, i.e., lure discrimination, provides an index of PS ability. In adults, functional magnetic resonance imaging (fMRI) of the HPC during a mnemonic discrimination task showed activation in CA3/DG predicted lure discrimination (Lacy et al. 2011; Reagh and Yassa 2014; reviewed in Yassa and Stark 2011). Similarly, structural magnetic resonance imaging (sMRI) research suggests that variations in DG and CA3 volumes relate to individual differences in PS in adult humans (Doxey and Kirwan 2015; Stark and Stark 2017; although see Deuker et al. 2014). In aging, hippocampal subfields are altered, and both volumetric and functional differences of HPC subfields have been linked to diminished PS. Aging adults show both decreased PS ability and CA3/DG volumes, and that the age-related decrements in PS relate to CA3/DG volume among the elderly (Stark and Stark 2017). Additionally, activity in CA3/DG has been linked to mnemonic discrimination deficits occurring during aging. Older adults show both disruption in the hippocampal circuitry due to over-excitation of CA3/DG and age-related rigidity in their ability to resolve interference between similar objects, requiring more dissimilarity between target and lures before CA3/DG shows evidence of separation-like activation (Yassa and Stark 2011).

Much less is known about development of CA3/DG and PS early in life. Recent studies have shown age-related differences in multiple hippocampal subfield volumes from 4- to-13 years of age (Lee et al. 2014; Tamnes et al. 2014; Daugherty et al. 2017; Riggins et al. 2018). These findings are consistent with neuroanatomical studies in non-human primates that suggest DG and CA3 exhibit a prolonged developmental time course, with maturity emerging during early childhood (Seress 2001; Lavenex and Lavenex 2013).

The developmental timeline of CA3/DG coincides with the development of PS in children, according to non-human primate studies. A recent behavioral study examining PS using

a child-friendly mnemonic discrimination task identified 4 to 6 years as a time of significant improvement in PS (Ngo et al. 2017). However, no studies to date have directly examined associations between CA3/DG and PS during early childhood. The goal of the present study was to address this gap. We tested the hypothesis that age-related improvements in PS relate to differences in CA3/DG hippocampal subfield volumes.

Materials and Methods

Participants

A total of 97 4- to- 8-year-old children participated in the current study. Of these children, 2 were excluded due to experimenter error, 20 failed to pass training, and 7 were excluded due to poor quality of MRI data that resulted in unusable segmentations. Overall, 68 children (N_{female}= 37, M_{years}= 6.67 ± 1.32, range= 4.19-8.96 years) provided useable behavioral and neuroimaging data and were included in the analyses; 42 of these participants (62%) were included in Riggins and colleagues (2018), which examined relations between hippocampal volume and performance on an episodic memory task (i.e., that did not probe pattern separation). Children were recruited from a major metropolitan area through the use of a University maintained database of families interested in participating in research and the distribution of recruitment flyers. Children were enrolled in a larger, ongoing, longitudinal study on memory and brain development. Children were screened to ensure they were not born premature, had normal or corrected-to-normal vision, and had no diagnoses for any neurological conditions, developmental delays, or disabilities. Informed consent was obtained from parents, and written assent was obtained for children older than 7 years. The majority of the sample was Caucasian, from middle- to high- income households.

Materials

Mnemonic Similarity Task (MST). The MST in the current study was modified by Ngo and colleagues (2017) to be child-friendly and included 88 pairs of common everyday objects. These pairs were selected randomly from 161 object pairs; 46 pairs drawn from an adult MST stimuli database (http://faculty.sites.uci.edu/starklab/mnemonic-similaritytask-mst/) and 115 pairs drawn from the internet based on their appeal to children (e.g., toys, animals) and the likelihood that children would be familiar with the objects (e.g., hat, bicycle; for additional details see Ngo et al. 2017). The selected 88 object pairs were divided into four sets of 22 stimuli and subsequently counterbalanced into versions of study and test lists such that each set of objects was assigned as the targets, lures, foils, and not-tested items an equal number of times across participants (each item in the pair could be a target, lure, foil, or not-tested stimulus).

Procedure

Mnemonic Similarity Task (MST). The MST consisted of an incidental encoding task, training session, and test (if training was passed). In the incidental encoding task, children made indoor/outdoor judgment for individual pictures presented on a computer screen using one of two buttons on a toy button box. When pressed, the buttons played a recording of the word 'indoors' (white button) or 'outdoors' (yellow button). Children were instructed to make decisions quickly because they had three seconds for each decision. Before the encoding task began, two self-paced practice trials (a bird, a spoon) were given to familiarize children with the general rule of the game. During the task, sixty-six objects were presented sequentially in a randomized order for 3 s each followed by a 0.5 s ITI via Eprime 2.0 (Psychology Software Tools, Pittsburgh, PA; see Figure 1). The incidental encoding task lasted approximately 5 minutes.

Following the incidental encoding task, children immediately began the training session where they were introduced to a new game and new toy button box consisting of three buttons

that, when pressed, played a recording of the phrase ‘exactly the same’ (red button), ‘kind of the same’ (blue button), or ‘new picture’ (green button). In the new game, children made memory judgments about pictures, some of which were seen during the indoor/outdoor game, some of which were similar to pictures during the indoor/outdoor game, and some that were new and dissimilar to the studied items. Items assigned as ‘not-tested’ at encoding (e.g., the Lego in Figure 1) were used for the training session to help children understand the task.

Children completed training trials for 4 ‘not-tested’ items, pressing the toy buttons themselves, where they decided if each practice item (shown on the right side of the screen) was ‘exactly the same’, ‘kind of the same’, or a ‘new picture’ compared to a ‘not-tested’ indoor/outdoor picture (shown on left side of the screen). Children then completed a set of six ‘screening’ trials. Children received immediate feedback from the experimenter following the completion of each trial. If they chose an incorrect response, they were corrected and were given additional training. If they chose the correct button for all ‘screening’ trials, they continued onto the test phase. The training session took approximately 3-6 minutes (see Ngo et al. 2017 for additional details).

Immediately after training, children were given a self-paced test on 66 items evenly divided between targets, lures, and foils (22 of each). Targets were identical items to those studied at encoding. Lures were similar exemplars of the studied items. Foils were novel items that were dissimilar from other objects in the stimuli set. For every trial, the experimenter asked, ‘Is this exactly the same, kind of the same, or completely new compared with the ones you saw before?’ Once the children pressed a button on the toy box, the experimenter recorded their responses by pressing keys corresponding to ‘old’ (‘exactly the same’), ‘similar’ (‘kind of the same’), and ‘new’ (‘new picture’) responses on the keyboard. The test image remained on the

screen until the experimenter inputted the children's response. The order of the test items was randomized for each participant. The test phase took approximately 5-6 minutes.

Proportions of memory responses (old, similar, and new) for each item type (target, lure, and foil) were separately calculated for each participant (see Table S1 for average performance across participants). Proportions of old responses to lures ("old" | lure) were subtracted from the proportion of old responses to targets ("old" | target) to create a bias-corrected measure of lure discrimination and served as the behavioral measure of pattern separation (Leal et al. 2014; Loiotile and Courtney 2015). This lure discrimination formula adequately accounts for the relative overlap between old and lure memory strengths, indexing the extent of false endorsement of lures—the hallmark of pattern separation failure, corrected for target recognition. Possible values on this index range from -1 to 1, where positive values denote successful discrimination between targets and lures, negative values denote a higher tendency to over-generalize between two similar items, and value of zero denotes chance-level discrimination. Similar to previous studies, we also calculated general item memory accuracy, assessing children's abilities to discriminate old items from novel and dissimilar items by subtracting the proportion of "old" responses to foils from the proportion of "old" responses to targets (Lacy et al. 2011; Stark et al. 2013, 2015; Ngo et al. 2017).

MRI. All participants completed training in a mock scanner before MR data acquisition in order to become acclimated to the scanner environment and receive motion feedback. Participants were scanned in a Siemens 3.0-T scanner (MAGNETOM Trio Tim System, Siemens Medical Solutions, Erlangen, Germany) using a 32-channel coil. An initial structural scan was acquired using a high-resolution T1 magnetization-prepared rapid gradient-echo (MPRAGE) sequence consisting of 176 contiguous sagittal slices (0.9 mm isotropic; 1900 ms TR; 2.32ms

TE; 900ms inversion time; 9° flip angle; pixel matrix= 256 x 256). This was used to measure intracranial volume (ICV) and isolate the HPC for a subsequent ultra-high resolution structural scan using a T2-weighted fast spin echo sequence (TR=4120ms, TE=41ms, 24 slices, 149° flip angle, voxel size .4mm x .4mm x 2mm).

Subfields. Hippocampal subfield volumes were identified in the head and body of the HPC in both left and right hemispheres using an existing protocol (La Joie et al. 2010) based on Harding and colleagues (1998) and Duvernoy (1998). The protocol was selected after existing protocols for manual tracing of hippocampal subfields were reviewed (n = 21, see Yushkevich et al. 2015). Protocols developed for T2-weighted images with resolution similar to our data and collected from 3T scanners were compared. Although several exist, we selected a protocol (La Joie et al. 2010) that yielded the subfields of interest in both the head and body subregions of the hippocampus at the desired resolution (.4mm x .4mm x 2mm) on a 3T scanner (but see also Winterburn et al. 2013; Berron et al. 2017). This protocol was selected because previous research in children has suggested developmental effects may be present in both the hippocampal head and body (DeMaster et al. 2013; Riggins et al. 2015, 2018). Similar to La Joie and colleagues (2010), seven different slice types were identified from coronal slices and used for manual segmentation (see La Joie et al. 2010; Riggins et al. 2018 for details). Three subfields were identified: subiculum (Sub), CA1, and a combination region of CA2-4/dentate gyrus (CA2-4/DG). Although the latter region combines multiple subfields, it includes both of the subfields implicated in previous studies of PS (CA3 and DG) along with CA2 and CA4, both of which are small in size relative to the other subfields. Details regarding identification of internal and external boundaries are reported in Riggins et al. (2018).

Two raters (FG, TR) blinded to the age and sex of the subjects independently traced 10 cases (2 from each of the 5 age groups) bilaterally. Dice Similarity Coefficients (DSC) were calculated to determine overlap between raters and are as follows for each subfield: Sub = .74, CA1 = .73, CA2-4/DG = .85. DSC values above 0.7 are typically considered acceptable for agreement (Zijdenbos et al. 1994), as such, overlap between the two raters indicated agreement. Intra-class correlations (ICC (2,1); Shrout and Fleiss 1979) were also calculated to determine reliability of the volume measurement and are as follows for each subfield: Sub = .93, CA1 = .98, CA2-4/DG = .90. ICC values above .90 are typically considered highly reliable, indicating consistency in the volume measurements.

One rater (FG) then traced an additional 10 cases (again, 2 from each age group). These segmentations were combined with the 10 cases used for manual reliability (i.e., 20 total) and input into Automatic Segmentation of Hippocampal Subfields software (ASHS, Yushkevich et al. 2014) to create a study-specific template. This study-specific template was used to generate hippocampal subfield volumes for the entire sample. All resulting segmentations were checked visually for quality. No manual edits were made, and only subjects yielding high-quality segmentations were included in the present report.

As we did not hold any a priori predictions regarding differences in subfield contribution between hemispheres, we collapsed across hemispheres (Daugherty et al. 2016; Keresztes et al. 2017; Schlichting et al. 2017). This resulted in total bilateral volumes of each subfield.

In order to ensure that any observed effects were not the result of differences in brain size, subfield volumes were subsequently adjusted to control for differences in ICV using an analysis of covariance approach (Jack et al. 1989; Van Petten 2004; Raz et al. 2005). Brain extraction was conducted separately in six toolboxes including ANTs, AFNI, FSL, BSE,

ROBEX, and SPM8. Multiple toolboxes were utilized to maximized anatomical specificity. The voxels extracted by at least four toolboxes were included in the brain mask (see Tillman et al. 2017 for similar approach). Exploration of ICV values indicated significant influences of age (β = 0.345, p = .003) and sex (β = -0.316, p = .007) on total brain size (adjusted R^2 = .16, F (2,65) = 7.83, p = .001). Consequently, both age and sex were used to calculate predicted ICV values for the entire sample (see Keresztes et al. 2017; Riggins et al. 2018 for similar approach). All subfield volumes were adjusted for predicted ICV values (adjusted volume = raw volume – b * (ICV – predicted ICV); see Keresztes et al. 2017). To account for the possibility that any observed effects were simply a product of this adjustment, results were examined for native volumes first and then for adjusted volumes. Only the latter are reported.

Results

Improvements in pattern separation during early childhood

A linear regression analysis was performed to assess the degree to which age predicted lure discrimination performance on the MST ("old" | targets - "old" | lures), controlling for sex. The model was significant (Adjusted R^2 = .28, F (2, 65) = 13.75, p < .001), with age predicting PS performance (β = 0.49, p < .001), even after controlling for sex differences (β = 0.25, p = .021).

Figure 1 about here

Age-related differences in hippocampal subfields during early childhood

To identify age-related neuroanatomical differences in HPC subfields, linear regression analyses were conducted predicting adjusted HPC subfield volumes via age and sex (See Figure 2). The best fitting models indicated significant quadratic age trends in CA2-4/DG (Adjusted R^2 = 0.11, p_β linear = 0.013, p_β quadratic = 0.018) and CA1 (Adjusted R^2 = 0.09, p_β linear = 0.009,

p_β quadratic = 0.011), and, a significant linear age trend in Sub (Adjusted R^2 = 0.05, p_β = 0.035). Sex was non-significant in all models (*p*s > .24). To address the multiple testing for the various hypotheses, the Holm–Bonferroni method controlling for family-wise errors at alpha level (0.05) was applied (Holm 1979).

Figure 2 about here

Age moderates the relation between hippocampal subfield volume and PS

To assess relations between hippocampal subfield volumes and lure discrimination (MST performance), a multiple linear regression analysis was conducted to predict lure discrimination performance using age, subfield, and age by subfield interactions, controlling for sex (see Table S2 for correlations between raw memory measures and subfield volumes).

$$Lure\ Discrimination \sim CA1 + CA2\text{-}4/DG + Sub + CA1 \times age + \mathbf{CA2\text{-}4/DG \times age} + Sub \times age + age + sex$$

This model accounted for 36% of the variance in lure discrimination (Adjusted R^2 = 0.27, F (8, 59) = 4.154, p < .001). The interaction between CA2-4/DG volume and age was the only significant predictor of lure discrimination (β = -0.29, p = .04). Specifically, the relation between CA2-4/DG volume and lure discrimination was positive in younger children, but negative in older children (See Figure 3). In contrast, neither Sub nor CA1 volumes showed main effects (*p*s > .50) or interactions with age (*ps* >.12) in predicting lure discrimination. To ensure that multicollinearity (e.g., high correlations among HPC subfields; see Table S3 for correlations between demographic variables, memory indices, and subfield volumes) was not adversely impacting our regression results, we examined variable inflation factors (VIF) attributed to each predictor in our model. We found VIFs ranged from 1.092 to 2.474 for predictor variables, including 1.684 for the predictor variable of interest (CA2-4/DG × age). Because these values

were within the acceptable range, VIF < 10, predictor variables were not removed to manage high inter-correlation.

Figure 3 about here

To examine the specificity of this finding, we tested an additional model predicting item memory ("old" | targets – "old" | foils) using age, subfield, and age by subfield interactions, controlling for sex. Only age significantly related to item memory ($\beta = 0.53$, $p < .001$). Neither subfields nor age by subfield interactions predicted item memory ($ps > .11$; see Table S4 for comparison of regression models predicting lure discrimination and item memory). For this model, VIFs ranged from 1.095 to 2.588 for predictor variables.

Discussion

The present study provides empirical evidence supporting theories relating protracted development of hippocampal subfields and detailed memory during early childhood. We found an age-moderated relation between CA2-4/DG subfield volume and PS ability in 4- to- 8-year-old children such that larger CA2-4/DG volume predicted better PS in younger children, whereas smaller CA2-4/DG volume related to better PS in older children. This apparent developmental transition is consistent with the hypothesis that maturation of hippocampal circuitry maturation is critical for the formation of high-resolution memories and may contribute to children's poor memories of past events in the early years of life. These results showed specificity in that 1) Sub and CA1 did not relate to lure discrimination performance on the child-friendly MST, and 2) CA2-4/DG did not relate to memory for items in general.

Volume of CA2-4/DG showed a nonlinear association with age, which aligns with previous literature showing decreased volume of subfields into late adolescence (Tamnes et al. 2014; Daugherty et al. 2016; Tamnes et al. 2017; Riggins et al. 2018). This suggests that more

"mature" hippocampal subfield development results in better PS ability. However, what is beneficial for a younger versus older child may differ in the context of development. It may be that older children with smaller CA2-4/DG show better PS due to microstructural development of these subfields (i.e., synaptic pruning and synaptogenesis as suggested by Daugherty et al. 2017), with synaptic connections become more "mature" and memories more stable as they move into adolescence. Comparatively, older children with larger CA2-4/DG volumes may be performing worse due less maturation and refinement of the connectivity in this HPC subfield. This is consistent with work suggesting continual dendritic development and synaptogenesis until at least the age of 5 years (Eckenhoff and Rakic 1991; Seress 2001; Lavenex and Lavenex 2013). Our findings support an inflection point in development where CA2-4/DG subfield maturity reflects stability in the circuitry supporting adult-like memory, specifically in the projections connecting dentate gyrus granule cells to CA3 pyramidal neurons (Seress 2001; Rolls 2013).

These results are consistent with those reported by Ngo and colleagues (2017) in 4- to 6-year-old children and bolster the argument that early childhood is a period of significant development in PS ability. To our knowledge, our study is the first to examine the relation of specific HPC subfields to PS ability during early-childhood. A recent report by Keresztes and colleagues (2017) showed greater HPC maturity predicted improved PS ability in 6- to- 25-year-olds. However, it is difficult to draw a direct comparison of this finding to the current study due to methodological differences. First, Keresztes et al. focused on an older and wider age range (6-25 years versus 4-8 years). Critically, the present study's inclusion of younger children, 4 - 6 years, captures the critical developmental period where robust improvements in PS (Ngo et al. 2017) supported by the functional maturation of DG and CA3 hippocampal subfields are thought

to occur. Second, Keresztes and colleagues (2017) utilized a latent measure of hippocampal maturity including both subfields and entorhinal cortex, whereas the present study solely focused on subfield volumes included in the hippocampus proper, with subfields considered individually.

It is worth noting that none of the hippocampal subfield measures related to memory discrimination for targets from dissimilar and novel foils, i.e., item memory. It is likely that extra-hippocampal cortices may support item memory (e.g., PrC: Davachi 2006). Given that the goal of the present study was to specifically examine the role of hippocampal subfields in PS based on CLS theory's predictions, other brain structures were not included and thus this possibility awaits exploration. It is also possible that although HPC subfield volume did not predict item memory, the activation or the connectivity of HPC might relate (e.g., Tang et al. 2017).

Despite the novelty and important implications of our findings, there are some limitations. First, these findings are correlational and do not directly assess functional mechanisms underlying the associations between hippocampal volume and behavior. Second, the study is cross-sectional and does not address change within individuals. Third, due to the modest sample size within each age subgroup, the current findings should be treated with caution due to the potential vulnerability to inflated effect sizes and/or Type II errors (Button et al. 2013). Finally, due to the specific hypotheses, only one behavioral task indexing PS was administered within a narrow age range.

These limitations are targets of future research. For example, exploration of functional differences to further clarify relations between PS ability, age, and hippocampal subfields would be beneficial. However, assessing possible functional differences will be challenging for this age range, as task-based fMRI paradigms prove difficult for young children because they must

remain still while performing, what is for them, a difficult cognitive operation. In addition, longitudinal studies examining developmental *changes* in detailed memory, using both measures of PS and episodic memory and the corresponding trajectory of HPC subfields, could provide a more fine-grained understanding of why young children's memories lack granularity and how this relates to the offset of childhood amnesia. Finally, incorporating these findings into the larger neural network supporting memory in every day circumstances is a necessary step for full understanding of this important cognitive ability. This would include examination not only of pattern separation (i.e., keeping events separate in memory) but also pattern completion (i.e., connecting related experiences despite minor differences at the time of encoding), as both are necessary in how we remember experiences in our lives.

Funding

This research was supported by the National Institutes of Health (grant number HD079518), awarded to TR.

Acknowledgements

The authors declare no competing financial interest. We thank the members of the Neurocognitive Developmental Lab at the University of Maryland for their support in data collection, the Maryland Neuroimaging Center for their support, and the participants and their families for their cooperation and participation.

References

Bakker A, Kirwan CB, Miller M, Stark CEL. 2008. Pattern Separation in the Human Hippocampal CA3 and Dentate Gyrus. Science. 319:1640–1642.

Bauer PJ. 2006. Constructing a past in infancy: a neuro-developmental account. Trends Cogn Sci. 10:175–181.

Berron D, Vieweg P, Hochkeppler A, Pluta JB, Ding SL, Maass A, Luther A, Xie L, Das SR, Wolk DA, et al. 2017. A protocol for manual segmentation of medial temporal lobe subregions in 7 Tesla MRI. Neuroimage Clin. 15:466–482.

Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR. 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nature Rev Neurosci. 14(5):365–376.

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

Daugherty AM, Flinn R, Ofen N. 2017. Hippocampal CA3-dentate gyrus volume uniquely linked to improvement in associative memory from childhood to adulthood. Neuroimage. 153:75–85.

Davachi L. 2006. Item, context and relational episodic encoding in humans. Curr Opin Neurobiol. 16:693–700.

DeMaster D, Pathman T, Lee JK, Ghetti S. 2013. Structural development of the hippocampus and episodic memory: developmental differences along the anterior/posterior axis. Cereb Cortex. 24:3036–3045.

Deuker L, Doeller C, Fell J, Axmacher N. 2014. Human neuroimaging studies on the hippocampal CA3 region–integrating evidence for pattern separation and completion. Front Cell Neurosci. 8:64.

Ding SL, Van Hoesen GW. 2015. Organization and detailed parcellation of human hippocampal head and body regions based on a combined analysis of cyto- and chemo-architecture. J Comp Neurol. 523:2233–2253.

Doxey CR, Kirwan CB. 2014. Structural and functional correlates of behavioral pattern separation in the hippocampus and medial temporal lobe. Hippocampus. 25:524–533.

Duvernoy HM. 1998. The Human Hippocampus, 2nd ed. Berlin: Springer.

Eckenhoff MF, Rakic P. 1991. A quantitative analysis of synaptogenesis in the molecular layer of the dentate gyrus in the rhesus monkey. Brain Res Dev Brain Res. 64:129–135.

Hamond NR, Fivush R. 1991. Memories of Mickey Mouse: Young children recount their trip to Disneyworld. Cog Dev. 6:433–448.

Harding AJ, Halliday GM, Kril JJ. 1998. Variation in hippocampal neuron number with age and brain volume. Cereb Cortex. 8:710–718.

Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat. 1979. 6:65–70.

Insausti R, Juottonen K, Soininen H, Insausti AM, Partanen K, Vainio P, Laakso MP, Pitka A. 1998. MR Volumetric Analysis of the Human Entorhinal, Perirhinal, and Temporopolar Cortices. Am J Neuroradiol. 19:659–671.

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

Josselyn SA, Frankland PW. 2012. Infantile amnesia: A neurogenic hypothesis. Learn Mem. 19:423–433.

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

Kirwan CB, Stark CEL. 2007. Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe. Learn Mem. 14:625–633.

La Joie R, Fouquet M, Mézenge F, Landeau B, Villain N, Mevel K, Pélerin A, Eustache F, Desgranges B, Chételat G. 2010. Differential effect of age on hippocampal subfields assessed using a new high-resolution 3T MR sequence. Neuroimage. 53:506–514.

Lacy JW, Yassa MA, Stark SM, Muftuler LT, Stark CEL. 2010. Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. Learn Mem. 18:15–18.

Lavenex P, Lavenex PB. 2013. Building hippocampal circuits to learn and remember: Insights into the development of human memory. Behav Brain Res. 254:8–21.

Leal SL, Tighe SK, Jones CK, Yassa MA. 2014. Pattern separation of emotional information in hippocampal dentate and CA3. Hippocampus. 24:1146–1155.

Lee JK, Ekstrom AD, Ghetti S. 2014. Volume of hippocampal subfields and episodic memory in childhood and adolescence. Neuroimage. 94:162–171.

Loiotile RE, Courtney SM. 2015. A signal detection theory analysis of behavioral pattern separation paradigms. Learn Mem. 22:364–369.

Neunuebel JP, Knierim JJ. 2014. CA3 retrieves coherent representations from degraded input: Direct evidence for CA3 pattern completion and dentate gyrus pattern separation. Neuron.

81:416–427.

Newcombe NS, Lloyd ME, Ratliff KR. 2007. Development of episodic and autobiographical memory: A cognitive neuroscience perspective. In: Kail RV, editor. Advances in child development and behavior. San Diego: Elsevier. p. 37–85.

Ngo CT, Newcombe NS, Olson IR. 2017. The ontogeny of relational memory and pattern separation. Dev Sci.

Norman KA, O'Reilly RC. 2003. Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. Psychol Rev. 110:611–646.

Peterson C, Hallett D, Compton-Gillingham C. 2017. Childhood amnesia in children: A prospective study across eight years. Child Dev.

Raz N, Lindenberger U, Rodrigue KM, Kennedy KM, Head D, Williamson A, Dahle C, Gerstorf D, Acker JD. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. Cereb Cortex. 15:1676-1689.

Reagh ZM, Yassa MA. 2014. Object and spatial mnemonic interference differentially engage lateral and medial entorhinal cortex in humans. Proc Natl Acad Sci USA. 111:4264–4273.

Riggins T, Blankenship SL, Mulligan E, Rice K, Redcay E. 2015. Developmental differences in relations between episodic memory and hippocampal subregion volume during early childhood. Child Dev. 86:1710–1718.

Riggins T, Geng F, Botdorf M, Canada K, Cox L, Hancock GR. 2018. Protracted hippocampal development is associated with age-related improvements in memory during early childhood. Neuroimage. 174:127–137.

Rolls E. 2013. The mechanisms for pattern completion and pattern separation in the hippocampus. Front Syst Neurosci. 7:74.

Rubin DC. 2000. The distribution of early childhood memories. Memory. 8:265–269.

Schlichting ML, Guarino KF, Schapiro AC, Turk-Browne NB, Preston AR. 2017. Hippocampal structure predicts statistical learning and associative inference abilities during development. J Cogn Neurosci. 29:37–51.

Seress L. 2001. Morphological changes of the human hippocampal formation from midgestation to early childhood. In: Nelson CA, Luciana M, editors. Handbook of Developmental Cognitive Neuroscience. Cambridge: MIT Press. p. 45–58.

Shrout PE, Fleiss JL. 1979. Intraclass Correlations: Uses in Assessing Rater Reliability. Psychol Bull. 86:420–428.

Stark SM, Stark CEL. 2017. Age-related deficits in the mnemonic similarity task for objects and scenes. Behav Brain Res. 333:109–117.

Stark SM, Yassa MA, Lacy JW, Stark CEL. 2013. A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. Neuropsychologia. 51:2442–2449.

Tamnes CK, Walhovd KB, Engvig A, Grydeland H, Krogsrud SK, Østby Y, Holland D, Dale AM, Fjell AM. 2014. Regional hippocampal volumes and development predict learning and memory. Dev Neurosci. 36:161–174.

Tamnes CK, Herting MM, Goddings AL, Meuwese R, Blakemore SJ, Dahl RE, Güroğlu B, Raznahan A, Sowell ER, Crone EA, et al. 2017. Development of the Cerebral Cortex across Adolescence: A multisample study of inter-related longitudinal changes in cortical volume, surface area, and thickness. J Neurosci. 37:3402–3412.

Tang L, Shafer AT, Ofen N. 2017. Prefrontal Cortex Contributions to the Development of Memory Formation. Cereb Cortex. 1–14.

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

Winterburn J, Pruessner J, Chavez S, Schira M, Lobaugh N, Voineskos A, Chakravarty M. 2013. A novel in vivo atlas of human hippocampal subfields using high-resolution 3 T magnetic resonance imaging. Neuroimage. 74:254–265.

Yassa MA, Stark CEL. 2011. Pattern separation in the hippocampus. Trends Neurosci. 34:515–525.

Yushkevich PA, Amaral RSC, Augustinack JC, Bender AR, Bernstein JD, Boccardi M, Bocchetta M, Burggren AC, Carr VA, Chakravarty MM, et al. 2015. Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in in vivo MRI: Towards a harmonized segmentation protocol. Neuroimage. 111:526–541.

Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. 1994. Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Trans Med Imaging. 13:716–724.

Captions

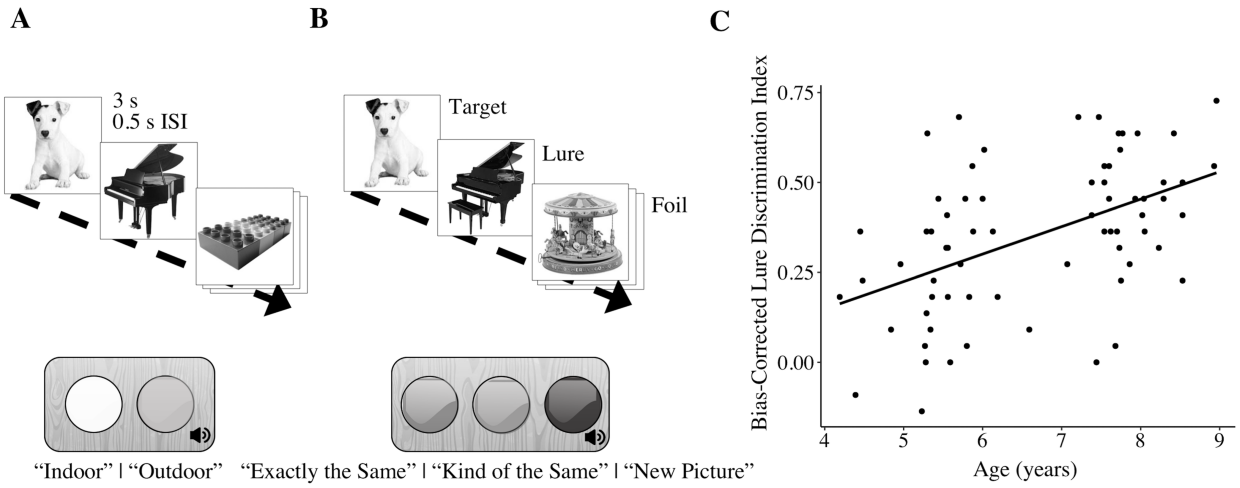

Figure 1. MST performance improves with age.

A schematic depiction of the encoding (A) and test (B) phases of the mnemonic similarity task (see Ngo et al. 2017). (C) Scatterplot of age and bias-corrected lure discrimination index.

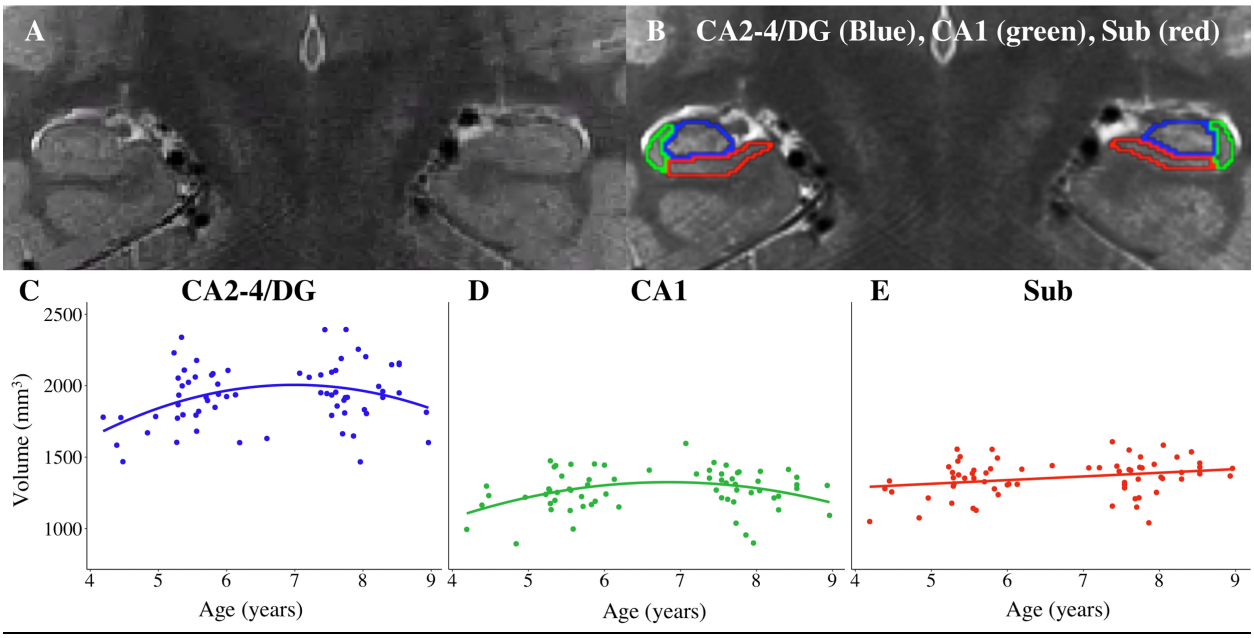

Figure 2. Hippocampal subfields show different developmental trajectories.

A and B) Example tracing of hippocampal subfields: CA2-4/DG in blue, CA1 in green, and subiculum in red from the present study's sample via La Joie and colleagues' protocol (see La Joie et al. 2010). Relations between volume and age for CA2-4/DG (C), CA1 (D), and subiculum (E).

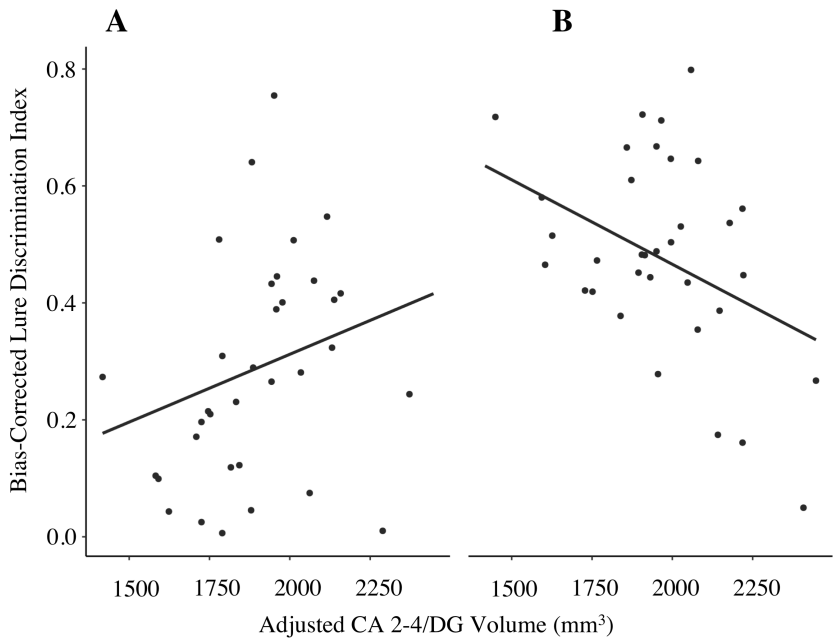

Figure 3. Age-moderates the relation between CA2-4/DG and MST performance.

(A) Relation between CA2-4/DG volume and lure discrimination in "younger" children (i.e., 1 SD below the mean age, or approximately 5.34 years). (B) Relation between CA2-4/DG volume and lure discrimination in "older" children (i.e., 1 SD above the mean age, or approximately 7.98 years).

Supplementary Material

Table S1. Mean proportions of responses to each stimulus type.

	Response Type		
Stimulus Type	Old	Similar	New
Target	0.75 (.22)	0.13 (.14)	0.12 (.14)
Lure	0.40 (.18)	0.37 (.20)	0.23 (.19)
Foil	0.07 (.13)	0.12 (.16)	0.81 (.22)

N=68. Standard deviations in parentheses.

Supplementary Material

Table S2. Pearson correlations between age, raw memory measures, and subfield volumes.

Trial Type	**Targets**			**Lures**			**Foils**		
Response	"old"	"similar"	"new"	"old"	"similar"	"new"	"old"	"similar"	"new"
Exact Age									
Overall	0.51***	-0.23+	-0.56***	0.06	0.43***	-0.54**	-0.32**	-0.06	0.23+
CA2-4/DG									
Overall	0.13	0.01	-.020+	0.17	-0.01	-0.15	0.27*	-0.19	-0.02
Younger	0.26	-0.11	-0.30+	0.15	0.06	-0.21	0.35*	-0.17	-0.10
Older	-0.40*	0.32+	0.21	0.21	-0.22	0.08	0.33+	-0.21	0.01
CA1									
Overall	0.13	-0.06	-0.14	0.14	0.00	-0.13	0.04	-0.04	0.00
Younger	0.24	-0.16	-0.21	0.21	0.12	-0.34+	-0.03	0.05	-0.02
Older	-0.22	0.17	0.12	0.05	-0.21	0.24	0.27	-0.14	-0.02
Sub									
Overall	0.19	-0.06	-0.24+	0.13	0.13	-0.27*	0.19	-0.03	-0.08
Younger	0.27	-0.12	-0.30+	0.37*	0.03	-0.41*	0.27	-0.08	-0.11
Older	-0.14	0.13	0.05	-0.18	0.14	0.01	0.21	0.05	-0.13

N=68. Shows correlation between each hippocampal subfield volume and endorsements of target, lure, and foil stimuli (younger < 6.66 years; n=34, older ≥ 6.66 years; n=34). $^{+}p < 0.10$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. All values are uncorrected for multiple comparisons.

Supplementary Material

Table S3. Pearson correlations between demographic variables, memory indices, and subfield volumes.

Measure	1	2	3	4	5	6	7
1. Age	---						
2. Sex	.155	---					
3. Lure Discrimination	.487**	.318**	---				
4. Item Memory	.548**	.342**	.691**	---			
5. Sub	.270*	.037	.144	.109	---		
6. CA1	.125	-.124	.059	.116	.581**	---	
7. CA2-4/DG	.191	-.044	.039	.020	.477**	.488**	---

N=68. $^{*}p < 0.05$; $^{**}p < 0.01$.

Supplementary Material

Table S4. Summary of linear regression models predicting lure discrimination (Left) and item memory (Right). Note that the regression model for each dependent variable was conducted separately.

	Dependent variable:	
	Lure Discrimination	Item Memory
Predictor	β	β
Age	0.415**(-0.113)	0.495**(-0.108)
Sex	0.239*(-0.110)	0.255*(-0.105)
CA2-4/DG	-0.030(-0.141)	-0.098(-0.134)
CA1	0.158(-0.152)	0.238(-0.145)
Sub	-0.095(-0.151)	-0.154(-0.144)
CA2-4/DG*Age	-0.293*(-0.141)	-0.108(-0.134)
CA1*Age	-0.073(-0.175)	-0.071(-0.167)
Sub*Age	0.270(-0.171)	-0.001(-0.163)
Constant	0(-0.103)	0(-0.99)
Observations	68	68
R^2	0.360	0.418
Adjusted R^2	0.274	0.339
Residual Std. Error	0.852 (df = 59)	0.813 (df = 59)
F Statistic	4.154** (df = 8; 59)	5.289** (df = 8; 59)

*$p < 0.05$; **$p < 0.01$. The values shown are standardized regression coefficients (β). Standard errors in parentheses.