Behavioral/Cognitive

Age-Related Increases in Posterior Hippocampal Granularity Are Associated with Remote Detailed Episodic Memory in Development

Bridget Callaghan,[1*] **Camille Gasser,**[2*] **Jennifer A. Silvers,**[1] **Michelle VanTieghem,**[2] **Tricia Choy,**[3] **Kaitlin O'Sullivan,**[4] **Alexa Tompary,**[5] **Lila Davachi,**[2,6†] **and Nim Tottenham**[2†]

[1]University of California, Los Angeles, California 90095, [2]Columbia University, New York, New York 10027, [3]University of California, Riverside, CA, 92521, [4]Georgetown University, Washington, DC 20057, [5]University of Pennsylvania, Philadelphia, Pennsylvania 19104, and [6]Nathan Kline Institute for Psychiatric Research, Orangeburg, NY, 10962

Episodic memory is critical to human functioning. In adults, episodic memory involves a distributed neural circuit in which the hippocampus plays a central role. As episodic memory abilities continue to develop across childhood and into adolescence, studying episodic memory maturation can provide insight into the development and construction of these hippocampal networks, and ultimately clues to their function in adulthood. While past developmental studies have shown that the hippocampus helps to support memory in middle childhood and adolescence, the extent to which ongoing maturation within the hippocampus contributes to developmental change in episodic memory abilities remains unclear. In contrast, slower maturing regions, such as the PFC, have been suggested to be the neurobiological locus of memory improvements into adolescence. However, it is also possible that the methods used to detect hippocampal development during middle childhood and adolescence are not sensitive enough. Here, we examine how temporal covariance (or differentiation) in voxel representations within anterior and posterior hippocampus change with age to support the development of detailed recollection in male and female developing humans. We find age-related increases in the distinctiveness of temporal activation profiles in the posterior, but not anterior, hippocampus. Second, we show that this measure of granularity, when present during postencoding rest periods, correlates with the recall of detailed memories of preceding stimuli several weeks postencoding, suggesting that granularity may promote memory stabilization.

Key words: development; fMRI; hippocampus; intervoxel similarity; memory; multivariate

---

Significance Statement

Studying hippocampal maturation can provide insight into episodic memory development, as well as clues to episodic functioning in adulthood. Past work has shown evidence both for and against hippocampal contributions to age-related improvements in memory performance, but has relied heavily on univariate approaches (averaging activity across hippocampal voxels), which may not be sensitive to nuanced developmental change. Here we use a novel approach, examining time signatures in individual hippocampal voxels to reveal regionally specific (anterior vs posterior hippocampus) differences in the distinctiveness (granularity) of temporal activation profiles across development. Importantly, posterior hippocampus granularity during windows of putative memory stabilization was associated with long-term memory specificity. This suggests that the posterior hippocampus gradually builds the capacity to support detailed episodic recall.

Introduction

Episodic memory, or the ability to form and retrieve detailed event information, is central to the human experience, contributing to autobiographical timelines (Bauer et al., 2007), academic outcomes (Mirandola et al., 2011; Blankenship et al., 2015), and emotional health (Phelps, 2004). Despite the importance of memory to human functioning, it has not yet been established how the construction and stabilization of episodic memories are supported by underlying brain maturation.

In adulthood, episodic memory encoding is supported by a distributed neural network (Spaniol et al., 2009; Davachi and

Danker, 2013; Preston and Eichenbaum, 2013), with the hippocampus playing a central role in associative binding (i.e., linking together elements of an experience). Current models posit that representations of event details and contextual information, routed through perirhinal and parahippocampal cortex (PhC), respectively, converge in the hippocampus to create integrated memories (Davachi et al., 2003; Davachi, 2006; Eichenbaum et al., 2007). Importantly, the ability to form detailed episodic recollections, and integrate new memories with existing knowledge structures, continues to develop throughout childhood and adolescence (Billingsley et al., 2002; Schneider et al., 2002; Brainerd et al., 2004; Ghetti and Angelini, 2008; Ghetti et al., 2011; Hayne and Imuta, 2011; Ghetti and Bunge, 2012; Brod et al., 2017). Studying the maturation of neural systems supporting episodic memory can thus provide insight into episodic memory development, as well as clues to its functioning in adulthood.

Age-related improvements in episodic memory suggest that hippocampal involvement in learning and memory continues to mature throughout childhood and adolescence. Evidence supporting that notion, however, remains inconclusive. Specifically, univariate hippocampal activation during memory encoding and retrieval has been shown to increase across childhood and adolescence in some studies (Ghetti et al., 2010; DeMaster and Ghetti, 2013; DeMaster et al., 2016; Riggins et al., 2016; Sastre et al., 2016; Selmeczy et al., 2019); but not others (Ofen et al., 2007; Selmeczy et al., 2019), and has been observed as early as toddlerhood (Prabhakar et al., 2018).

To better understand the developmental construction of the hippocampus, it is necessary to move beyond the common approach of averaging across voxels to assess mean hippocampal activation. Instead, leveraging multivariate analyses enables assessments of temporal activity fluctuations of single voxels. For example, a recent paper shows that the distinctiveness in temporal activation profiles across hippocampal voxels reveals distinct hippocampal subregions (Brunec et al., 2018). While the functional significance of this metric is not completely understood, one hypothesis is that distinctiveness or "granularity" in voxel time courses is associated with finer-grained information processing and higher detail in memory representations. Given that early memories contain less detail (Willoughby et al., 2012), it is possible that hippocampal granularity emerges gradually, which we predict would be associated with recall of detailed information.

In the current study, we applied this multivariate analysis approach (called "intervoxel similarity") to development, asking whether hippocampal representational "granularity" increases with age and is related to detailed episodic memory. Given past work, showing hippocampal resting activity is associated with memory performance, often independently from encoding activity (Tambini et al., 2010; Tambini and Davachi, 2013, 2019; Murty et al., 2017), we looked for age-related changes in hippocampal representational granularity during episodic encoding, and postencoding rest periods (i.e., a putative window of memory stabilization/consolidation). We then assessed detailed episodic memories immediately after learning and at a 2 week delay. Given prior work demonstrating different rates of structural and functional development along the hippocampal long axis (Gogtay et al., 2006; DeMaster and Ghetti, 2013; DeMaster et al., 2014); we looked at granularity in the anterior and posterior hippocampus separately. As prior work in adults had demonstrated higher granularity in the posterior than anterior hippocampus (Brunec et al., 2018),

we hypothesized that developmental increases in granularity would be most pronounced in the posterior hippocampus. We secondly hypothesized that posterior hippocampal granularity during both encoding and postencoding rest/potential consolidation windows would be associated with long-term memory retention.

Materials and Methods

Participants
In the current study, our initial subject pool comprised $N = 57$ participants, who ranged in age from 6 to 17 years (mean age = 11.43 years). There were $N = 31$ females and $N = 26$ males. Two participants exited the scanner before the necessary scans had started, and another participant completed the initial resting scan and the memory scan but then exited the scanner before the final resting scan. Hence, the final sample with all three scans was $N = 54$. In terms of behavioral data, we had $N = 49$ participants who completed the immediate memory test (of which $N = 46$ also had complete MRI data), $N = 35$ who completed the delay memory test (of which $N = 32$ had complete MRI data), and $N = 32$ who completed both the immediate and delay test (of which $N = 29$ had complete MRI data).

All procedures were approved by the Institutional Review Board at Columbia University. Parents provided informed consent for their children/adolescents, and children/adolescents provided assent for the study procedures.

Procedure
Participants completed two visits for this study. In the first visit, participants completed a range of questionnaires and a cognitive assessment, none of which are discussed here. On the second visit, participants completed the MRI scan. Before entering the scanner on this second visit, all participants were given instructions for the task, and were given an opportunity to practice responding. In the scanner, participants performed an initial rest scan (baseline rest), followed by the task (encoding), and then a second rest scan (postencoding rest). A high-resolution anatomic scan was collected after the second rest scan.

Following the completion of the scanning sequences, participants were administered a memory test outside of the scanner, which they were informed of before scanning. There were $N = 8$ participants who did not complete the immediate memory test because of noncompliance or running out of time. A subset of participants ($N = 33$) also completed a post-learning questionnaire, in which they had to indicate what they had thought about during the rest scans. Missing data resulted from participant refusal to complete the post-learning questionnaire or need to leave quickly after the scan. Specifically, participants were asked whether any toys or faces from the game "popped into their head" while they were resting after the learning task, and whether they felt very sleepy or thought they had fallen asleep. There was one participant who reported possibly falling asleep during one of the three scans; but as they were performing above chance levels (40% correct; chance level was 33% correct) for coarse episodic memory at the immediate test, we kept this participant in our analyses. We created a binary code for sleepiness during the scans (e.g., the participant who reported possibly falling asleep was given the sleepiness positive code) and controlled for sleepiness in the potential confound analysis.

Participants were administered a second memory test 5-28 d after the scan, except for one participant who completed the second memory test 102 d after the scan. Including all participants, the mean time to complete the second memory test was 13.15 d (mode = 7 d, median = 9 d, SD = 16 d). This test was completed online from each participant's home (via Qualtrics) with the assistance of the parent, who had no knowledge of the task content and therefore could not systematically bias their child's responses. As the second memory test was identical to the first, participants were well versed in the task instructions, reducing the possibility that children did not understand the at-home task instructions. (This point is confirmed by the high performance of youth in both the in-laboratory and at-home task.)

Encoding task

Participants performed one block of an associative encoding task in the scanner, presented on E-Prime (version 2.0). Visual stimuli for this task were projected onto a screen at the rear of the scanner, which was viewed through a mirror attached to the participant's head coil. The task comprised 41 trials. On each trial, a scene appeared on the screen for 500 ms, and then an item (toy or face) appeared in a location within that scene and remained on the screen for another 2500 ms. In total, there were 20 different items (10 toy objects and 10 child faces) embedded in seven different scenes. Thus, each of the seven scenes was paired with multiple toys/faces (2-4 unique toys/faces) with at least one toy object and one face paired with each of the seven scenes on different trials. Each of the item-scene pairs was presented between 1 and 4 times, yielding a total of 41 trials. Repetition and item interference within the scenes were embedded in the design to test the effects of repetition and interference on memory retention, which are the focus of analyses outside of this paper. As the current analyses do not involve comparison of trials, we do not control for repetition or interference in the current set of analyses. Each unique object and face were paired with only one scene, and each trial included only one item-scene pair. The object and face trials were presented in a fixed random order.

For the encoding task, participants were told that they would be playing a "hide-and-seek game" where they had to "help a little girl/boy [matched to the participant's sex] find their toys and friends." Participants were told to pay attention so that we could see what they remembered when they finished the scan. The toy objects were selected to be familiar to children and adolescents, including balls, musical instruments, sports equipment, and mechanical toys (e.g., truck or plane). The faces were photographs of 5 male and 5 female children of diverse races (see Stimuli). Each scene was roughly divided into four quadrants, and each of the objects/faces paired with a particular scene were located in a unique quadrant of the scene. The memory encoding task lasted for 6.9 min.

Each trial started with a 1 s fixation cross on a white screen. In between every trial, there was an active baseline task, which was jittered in duration (average jitter = 9 s, range = 2-18 s). During this active baseline period, participants saw a series of arrows (presented one at a time) pointing to the left or right (random selection of direction). Each arrow remained on the screen for 1 s before disappearing and being replaced by another arrow (until the end of the jittered interval). Participants were told to press a button every time they saw an arrow, using a 5 button box held in their dominant hand. There was no expectation of a meaningful response with the button presses during the active baseline task. Instead, the button presses were used to provide participants with a form of engagement during what was otherwise a passive viewing task. No other behavioral response was required during encoding.

Memory test

Participants performed two identical memory tests after exiting the scanner. The first was performed immediately after exiting the scanner on a laptop computer (via E-Prime), assisted by a researcher. The second was performed at a delay of ~2 weeks (mean = 13 d, mode = 7 d, median = 9 d, SD = 16 d). This test was completed online from each participant's home (via Qualtrics) with the assistance of the parent, who had no knowledge of the task content. The same stimuli were tested at the immediate and delay intervals. However, there was no feedback on whether the participants' responses were correct or incorrect at the immediate or delay tests.

Each memory test was self-paced. Children responded either verbally or by pointing to the screen, while the researcher/parent recorded their answer using the computer. Adolescents made their own responses on the computer (with the researcher sitting beside them during the immediate test). The test was composed of two sections: a recognition section (not discussed here) and an associative memory section (presented here). In the associative test section, each of the 20 trials was split into two parts (Fig. 1A). In Part A, participants saw a scene from the scanner task with three objects or faces below it, and were asked to pick which of the toys/faces had been paired with the scene ("which was hiding here?"). Each of the toys/faces had been shown during the encoding task, but only one of the three had been presented with that scene (i.e., was a

correct pair). Hence, each of the toys/faces shown during encoding acted as a target for their own scene, and as a distractor for the other scenes. (Each object/face was presented once as a target and two times as a distractor, aside from 1 object and 1 face that were presented once as a target and three times as a distractor).

All participants then progressed to Part B. During Part B, participants viewed the same scene from Part A, but with no toys/faces underneath. Overlaid on the scene were numbers showing the potential locations of toys/faces that had been paired with that scene, written in yellow text and surrounded by a yellow circle. Participants were asked to choose the location of the toy/face they had selected during Part A in the scene with the prompt: "where was it hiding?" Each potential location showed where one of the items paired with that scene had been located, but only one location was correct for the target item. Trials were presented in a fixed-random order.

For the associative memory test, we calculated two metrics: coarse episodic memory, which was the percent of correctly identified toys/faces for the scene in Part A; and detailed episodic memory, which was the percent of trials where both Part A (the toy/face) and Part B (the location) were correct. Chance performance for the coarse episodic memory metric was 33.33%. Chance performance for the detailed episodic memory metric ranged from 8.33% for scenes paired with four objects/faces (i.e., 33.33% ∗ 25% for four potential locations) to 16.67% for scenes paired with two objects/faces (i.e., 33.33% ∗ 50% for two potential locations). As there were multiple instances where children scored incorrectly on the recognition memory test, but then correctly for that same item in the associative version of the test, we analyzed all trials from the associative memory test, regardless of whether the participant correctly recognized the item in the recognition memory portion of the test.

Stimuli

All scene stimuli were high-resolution cartoons of different scenes (three indoor and four outdoor) in full color that were taken from the Internet, selected to be engaging for children. The indoor scenes were a bedroom, living room, and kitchen, and the outdoor scenes were a residential street, a park, a forest, and a candy land. All scene images were comparable in terms of vibrancy and color. Importantly, the scenes were devoid of face and toy stimuli in the background, other than the faces and toys that were presented as part of the task. The face stimuli were high-resolution photographs of children with the background removed that were taken from the Internet. The children in the photographs were all smiling, were in full color, and were oriented at different angles (e.g., facing directly forward, oriented slightly to the left or right). The faces were male or female gendered, looked to be in middle childhood to early adolescence, and represented a range of races. The faces used as foils (in the recognition memory test) matched the approximate age, sex, and racial distribution of the target face stimuli.

All toy stimuli were high-resolution, cartoon, color images of toys collected from the Internet, also with the background removed. The toys were age-appropriate and familiar (e.g., truck, crayons, hockey stick), and none of the toys had faces (e.g., there were no teddy bears or dolls). The toys used as foils (in the recognition memory test) were conceptually related to the target toys (e.g., there was an airplane target and a paper plane foil, a rapping drum target and a bongo drum foil) but were not systematically different in any other way. For both the face and toy items, the same set of stimuli were used as targets versus foils for each child. Rather than selecting the stimuli from standardized picture sets, stimuli were selected to contrast with the scene background (as they were presented embedded in the scene), and to be engaging even for young children. Children were given two practice trials outside of the scanner, and another practice trial in the scanner before the start of the encoding scan sequence. Stimuli used for the practice trials were unique and were not presented again in either the training phase (within the scanner) or during the test.

Rest scans

During both rest scans, participants were instructed to close their eyes and think about anything they wanted. However, to avoid the possibility
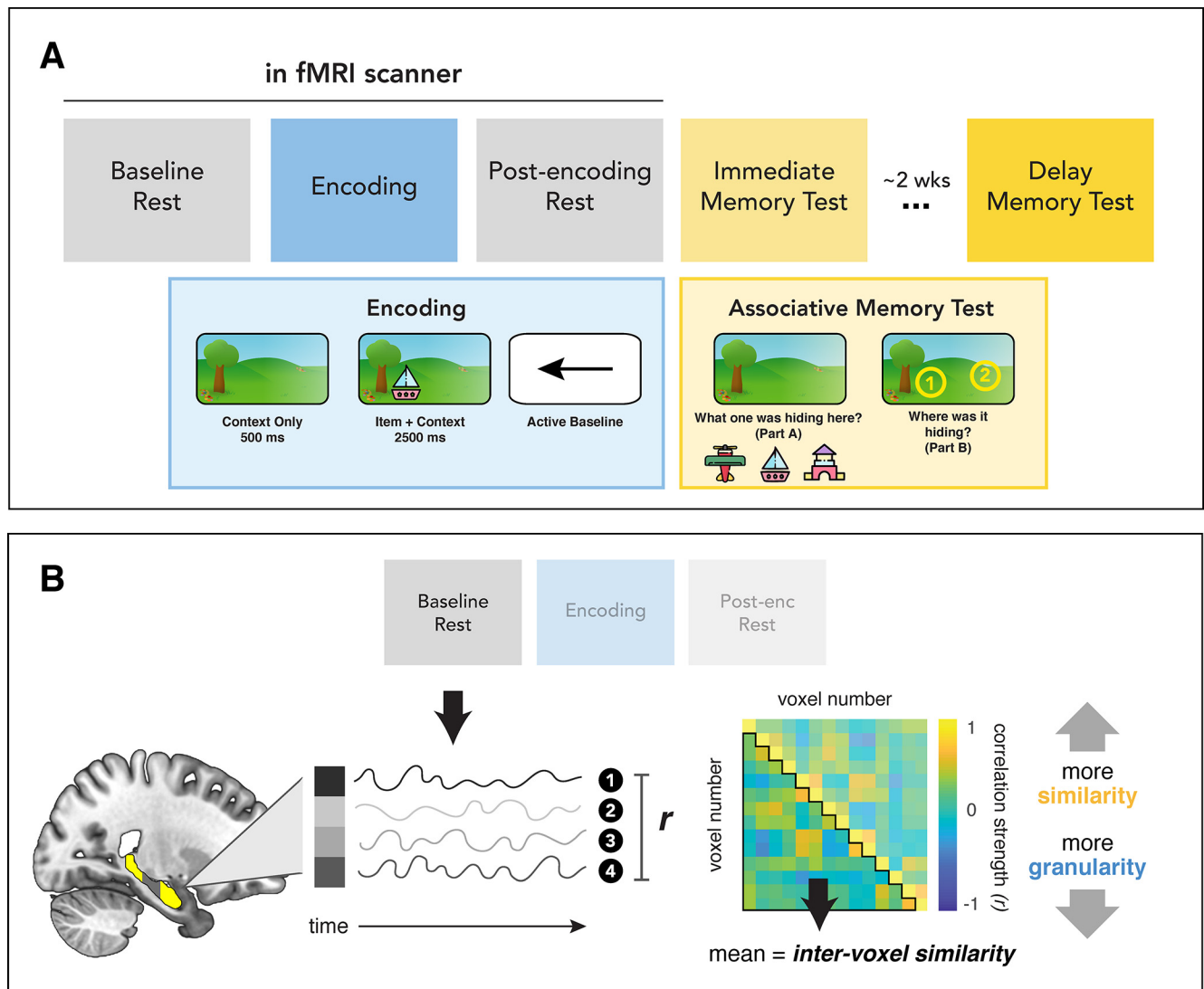
Figure 1. Task design and intervoxel similarity procedure. **A**, During encoding, participants learned a series of item-context associations. Next, they completed an episodic memory test for all learned items at two time points: immediately after learning and ∼2 weeks (wks) later. Both memory tests involved two parts. In Part A, participants had to indicate which of the following probe items was found within a given scene (correct responses on Part A constituted our measure of "coarse episodic memory"). In Part B, participants had to indicate where within the scene the selected item was located (correct responses on both Parts A and B of the episodic memory test constituted our measure of "detailed episodic memory"). **B**, Intervoxel similarity was computed in each Region of Interest (ROI) during each of the three scans. First, we extracted the time courses of all voxels within each ROI across the entire scan. Next, we calculated the correlations between all unique pairs of voxels within each ROI (within and across hemispheres). These resulting (Fisher *R*-to-*Z* transformed) correlation values were averaged together to get a single measure of intervoxel similarity for each scan and ROI. Higher values of intervoxel similarity indicate greater levels of representational similarity, whereas lower values indicate greater levels of representational granularity.

that rehearsal of content might produce similar patterns of activity during the post-task rest period as seen during learning, participants were told to try not to think about the hide-and-seek game they just played. Each rest scan lasted for 5 min.

fMRI data acquisition and preprocessing
Images were acquired with a 3T General Electric MRI system with a whole head Nova 32-channel coil. Functional data were collected using a gradient-EPI sequence (TR = 2s, TE = 30 ms, FOV = 20 cm, slices oriented 200 clockwise from AC-PC on the oblique plane, flip angle = 75°, voxel size = 3.125 × 3.125 × 4 mm, slice thickness = 4 mm, slice spacing = 0 mm). High-resolution T1-weighted Brain Volume Imaging–BRAVO images were acquired for anatomic visualization. Functional data were preprocessed and analyzed using the FMRIB Software Library (FSL; version 5.0.11) and custom MATLAB scripts (version 2019a). In all ensuing analyses, volumes where the framewise displacement (FWD) value exceeded 0.9 mm were censored by

creating a unique regressor for that volume. As an additional robustness check, we also reran all primary analyses using a FWD cutoff value of 0.5 mm. As results did not change with this more conservative FWD value, all analyses in the main text use 0.9 mm to preserve more of the data. The data were first corrected for differences in slice timing (interleaved), followed by motion correction (registration to the first volume). Functional data were high pass filtered (cut off = 111 s), and were not spatially smoothed. Each scan remained in participant native space but were aligned to each participant's structural data.

fMRI analysis
ROI definition. For the majority of participants with fMRI data, right and left hippocampal ROIs were extracted using FreeSurfer's automatic subcortical segmentation tool (N = 52). For the remaining participants, FreeSurfer's segmentation failed, and FSL's FMRIB's Automated Segmentation Tool was used instead (N = 2). Each participant's anatomically defined hippocampus was then segmented along its longitudinal

axis by dividing the number of coronal slices in each hemisphere evenly into three sections. Based on prior work (Dandolo and Schwabe, 2018), we expected to see a gradual change in functionality along the hippocampal long axis, and thus opted to divide the hippocampus into thirds, analyzing only the most anterior and posterior segments (leaving out the middle third) to characterize the extreme ends of the functional gradient. We refer to the anterior third of this division as anterior hippocampus, and the posterior third as posterior hippocampus.

mPFC and posterior medial cortex (PMC) ROIs were taken from a probabilistic atlas defined via resting-state functional connectivity (Shirer et al., 2012) (available at https://findlab.stanford.edu/functional_ROIs.html) as there are no clear anatomic landmarks for these structures. An ROI corresponding to bilateral PhC was taken from a probabilistic atlas of the medial temporal lobe (MTL) subregions (Ritchey et al., 2015) (https://identifiers.org/neurovault.collection:3731). Finally, an additional occipitotemporal cortex (OTC) region was defined from the Harvard-Oxford Cortical Probabilistic Atlas included with FSL.

We also extracted white matter (WM) and CSF masks for each participant to serve as nuisance ROIs. For the majority of participants ($N = 52$), these masks were defined using FreeSurfer's segmentation tool. For the remaining participants, FreeSurfer failed, and FMRIB's Automated Segmentation Tool was used in its place ($N = 2$). All resulting masks were eroded with FSL's default kernel and thresholded at a probability value of 0.9, before being aligned to each participants' functional data for each scan (baseline rest, encoding, and postencoding rest). We then applied an additional erosion step to all WM masks, such that only voxels for which at least two-thirds of the surrounding voxels were labeled WM remained in the mask. This second erosion step ensured that WM signal extracted from these masks was not contaminated by signal from gray matter. We checked each of the masks via visual inspection, and for one participant decided to perform one further erosion step, such that only voxels for which at least half of the surrounding voxels were labeled CSF remained in the mask.

Nuisance signals. To control for the influence of potential confounds on our measure of intervoxel similarity, we considered a number of additional nuisance variables: participant sex, mean FWD during each scan, the number of voxels in each ROI, temporal signal-to-noise ratio (tSNR) in each ROI, and univariate reactivity during encoding.

We used FSL's Motion Outliers to determine the mean FWD during each scan for each participant. Mean FWD was calculated as the average of rotation and translation parameter differences using weighted scaling (Power et al., 2012). Across all participants, the mean FWD (in millimeters) between adjacent time points was 0.17 (SD = 0.20) during baseline rest, 0.23 (SD = 0.25) during encoding, and 0.22 (SD = 0.27) during postencoding rest.

We also recorded the number of voxels in each participant's hippocampal ROIs (after they were aligned to the functional space of each of the three scans). Across all participants and scans, the mean number of voxels was 59.6 (SD = 14.4) in the bilateral anterior hippocampus and 48.4 (SD = 8.5) in the bilateral posterior hippocampus. Looking separately at left and right hemispheres, the mean number of voxels was 29.6 (SD = 8.7) in right anterior hippocampus, 30.0 (SD = 8.4) in left anterior hippocampus, 24.4 (SD = 5.4) in right posterior hippocampus, and 23.9 (SD = 5.5) in left posterior hippocampus.

Next, we computed tSNR within each hippocampal ROI for each scan. Using the preprocessed encoding and rest data, tSNR was calculated for each voxel as the mean BOLD signal divided by the SD across each scan. tSNR estimates for each voxel were then averaged together within each ROI (anterior vs posterior hippocampus). Across all participants and scans, the mean tSNR was 101.6 (SD = 38.7) in the bilateral anterior hippocampus, 136.6 (SD = 45.9) in the bilateral posterior hippocampus, 105.2 (SD = 42.6) in right anterior hippocampus, 135.2 (SD = 45.2) in right posterior hippocampus, 97.9 (SD = 37.3) in left anterior hippocampus, and 138.1 (SD = 46.4) in left posterior hippocampus.

Finally, for each participant, we computed the strength of hippocampal reactivity to encoding trials, as described in Univariate analysis: encoding task. This measure allowed us to control for the potential influence of univariate hippocampal activity on intervoxel similarity during the encoding task.

Univariate analysis: encoding task. In order to examine univariate BOLD signals in the hippocampus, we performed a first-level (i.e., within-participant) univariate GLM analysis using the FMRI Expert Analysis Tool (FEAT) in FSL. The standard six head motion parameters estimated by the MCFLIRT tool in FSL, their six temporal derivatives, and individual stick function regressors corresponding to time points with above-threshold FWD (as described in fMRI data acquisition and preprocessing) were added as confound explanatory variables in the model. The mean signal time courses extracted from the CSF and WM masks were also added as confound explanatory variables. Finally, we modeled reactivity during trials (scene and item + scene presentation) relative to the active baseline (arrows). Using the participant-specific hippocampal ROI masks, we extracted the parameter estimates from each participant's trial > active baseline contrast.

Multivariate intervoxel similarity analysis: encoding task and rest scans. For both rest scans and the encoding scan, intervoxel similarity was computed using the following procedure (Fig. 1B). Preprocessed data from each scan (baseline rest, encoding, and postencoding rest) were entered into a first-level GLM in each participant's native space using FSL's FMRI Expert Analysis Tool. In order to account for nuisance signals introduced by head motion, WM tissue, and CSF, we included the standard six head motion parameters estimated by MCFLIRT, their temporal derivatives, individual stick function regressors corresponding to time points with above-threshold FWD (as described previously), and the mean time courses extracted from CSF and WM masks as confound explanatory variables in each model.

Using the residuals of these GLMs, we then extracted the BOLD activity time course of each voxel in each ROI across each of the encoding and rest scans. Following the procedure outlined by Brunec et al. (2018), the resulting time courses were first z-scored within each voxel, scan, and participant. We then calculated the Pearson correlation between each pair of voxels to compute a multivoxel correlation matrix (Tambini and Davachi, 2013) for each ROI in each scan (for examples of these multivoxel correlation structure (MVCS) matrices, see Fig. 2). These matrices included pairwise correlations of voxels both within and across hemispheres (but using only within-hemisphere or across-hemisphere correlations yielded highly similar results). We then applied a Fisher R-to-Z transformation to each matrix, and took the average of all unique pairwise correlations in these matrices (i.e., values above the diagonal) to get a single measure of intervoxel similarity for each ROI during each scan.

Experimental design and statistical analysis

For all behavioral analyses, we used linear regression models to test the relationship between memory performance and participant age, after accounting for participant sex and the number of days between training and test (for the delay memory test only). Effect sizes for significant regressors were calculated using partial r^2. Next, we calculated across-subject Pearson correlations to examine the relationship between participant age and intervoxel similarity in each ROI (e.g., anterior and posterior hippocampus) during each scan (i.e., baseline rest, encoding, and postencoding rest). To account for any influence of potential confounds on significant age × similarity correlations, we performed multiple linear regression to test the relationship between age and hippocampal intervoxel similarity after controlling for the nuisance signals discussed above (e.g., motion, ROI size). Additionally, to examine whether similarity × age correlations differed as a function of scan or of hippocampal ROI, we used Williams's test (Williams, 1959), a method for testing the difference between dependent correlations, implemented via the psych package in R version 1.8.12 (Revelle, 2020).

We next considered the relationship between hippocampal intervoxel similarity and participant memory behavior, again using linear regression. Separate linear models were created to examine the relationship between episodic memory performance (during either the immediate or delay test) and intervoxel similarity in each hippocampal ROI and during each scan. All such models included participant age as a covariate. All statistical tests were implemented using a combination of SPSS (version 25) and R (version 3.6.1).

Subject A (9.1 years old)
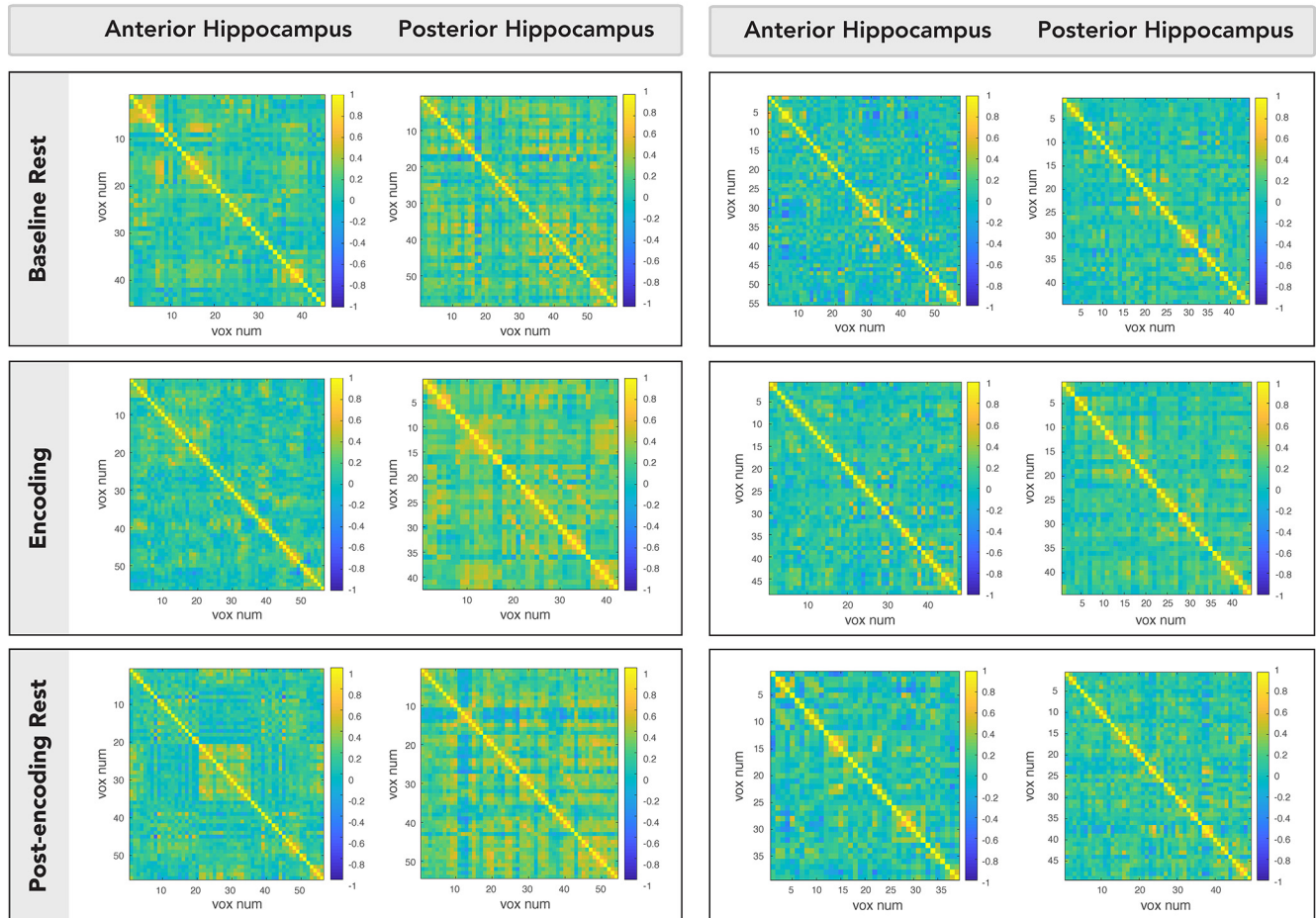
Subject B (17.3 years old)

Figure 2. Example multivoxel correlation structure (MVCS) matrices. MVCS matrices from anterior and posterior hippocampus for two model participants: one child (9.1 years old) and one adolescent (17.3 years old). Vox num represents the voxel number for each region of interest, for every scan for each participant.

Additionally, to complement these statistical tests (which all assume a Frequentist approach to statistical inference), we implemented a series of Bayesian linear models to further probe the relationships between age, intervoxel similarity, and behavior. These models were implemented via the brms package, version 2.10.0 (Bürkner, 2017, 2018) in R. All models included weakly informative priors for both slope and intercept terms [normal (0, 5) and normal (0, 10), respectvely]. Models were fit using four Markov chains, each with 10,000 iterations (warm-up iterations = 5000). Age and intervoxel similarity variables were standardized (i.e., z-scored) across participants before inclusion in each Bayesian model. In models including hippocampal subregion as a regressor, anterior and posterior hippocampi were effect-coded as −1 and 1, respectively. For all reported results, 95% CI indicates the corresponding 95% Bayesian credibility interval.

Results

Episodic memory performance
During encoding, participants studied a series of items (i.e., images of toys and faces), each of which was associated with a specific location in a larger scene. Memory for these item-scene associations was then assessed through two questions (Parts A and B; see Fig. 1A). In Part A, on each trial, participants were shown a single scene plus three items from the task underneath it. They were then asked to identify the target item (i.e., the item that had been paired with that scene during encoding; termed coarse episodic memory). In Part B, participants were asked where the item they had identified as the target in Part A

was located within the scene (2-4 potential locations were presented for each scene). Correct identification of the target item (in Part A) and the location of that item (in Part B) was taken as our metric of detailed episodic memory.

Coarse episodic memory
First, we performed linear regressions to determine whether participant age was related to item-scene memory performance (i.e., coarse episodic memory) at the immediate or delay intervals, controlling for participant sex and the number of days between training and test (for delay test performance only). For immediate coarse episodic memory ($N = 49$), there was no effect of age ($\beta = -0.01$, $t_{(46)} = -1.3$, $p = 0.20$) or participant sex ($\beta = 0.04$, $t_{(46)} = 0.94$, $p = 0.35$) on performance. For delay coarse episodic memory ($N = 33$), there was no effect of age ($\beta = 0.002$, $t_{(29)} = 0.20$, $p = 0.85$) or sex ($\beta = -0.03$, $t_{(29)} = -0.46$, $p = 0.65$) on performance, but there was a negative association between train-test interval and performance ($\beta = -0.01$, $t_{(29)} = -2.90$, $p = 0.007$, partial $r^2 = 0.22$), such that memory for associations was worse at longer delays.

Importantly, when calculating a retention score (delay − immediate memory performance), the effect of age was marginal, albeit not significant ($\beta = 0.16$, $t_{(28)} = 2.03$, $p = 0.052$, partial $r^2 = 0.10$), indicating some evidence that adolescents may have retained more of their initially learned information at the delay interval than children. There was no effect of participant sex on retention ($\beta = -0.03$, $t_{(29)} = -0.57$, $p = 0.573$), but again there

was a negative association between train-test interval and performance ($\beta = -0.005$, $t_{(29)} = -2.94$, $p = 0.006$, partial $r^2 = 0.20$).

We also used chance performance levels of $< 33\%$ of items correct on the coarse episodic memory component at the immediate test to exclude participants performing at below-chance levels, and then recalculated all of the behavioral analyses with those participants excluded. This resulted in the removal of three participants (all 14-year-old females). None of the behavioral results was changed (in terms of direction of the association or significance) for the immediate or delayed tests when those individuals were excluded.

Detailed episodic memory
We next performed linear regressions to determine whether participant age was related to item-scene-location memory (i.e., detailed episodic memory) at the immediate or delay intervals, controlling for participant sex and the number of days between training and test (for delay test performance only; see Fig. 3A). For immediate detailed episodic memory ($N = 49$), there was no effect of age ($\beta = -0.003$, $t_{(46)} = -0.31$, $p = 0.76$) or sex ($\beta = 0.02$, $t_{(46)} = 0.35$, $p = 0.73$) on performance. For delay test detailed episodic memory ($N = 33$), there was no effect of age ($\beta = -0.002$, $t_{(29)} = -0.21$, $p = 0.83$), sex ($\beta = 0.04$, $t_{(29)} = -0.58$, $p = 0.57$), or train-test interval ($\beta = 0.004$, $t_{(29)} = -1.76$, $p = 0.09$) on performance.

For the retention score (delay − immediate detailed memory performance), there was no effect of age ($\beta = 0.005$, $t_{(29)} = 0.71$, $p = 0.485$), sex ($\beta = -0.03$, $t_{(29)} = -0.52$, $p = 0.608$), or train-test interval ($\beta = -0.003$, $t_{(29)} = -2.00$, $p = 0.055$).

In sum, while some, but not all, prior work has shown benefits in memory recall with age, the paradigm used here does not lead to striking age differences. Thus, the task we use here is well suited to examine how functional brain changes occur across development when behavioral performance is matched across age.

Association between age and intervoxel similarity
We next addressed our main question of whether hippocampal representational granularity varies across age. To this end, we computed the correlation between participant age and intervoxel similarity (average cross-voxel temporal correlations, the inverse of granularity) within each of our hippocampal ROIs (posterior and anterior), separately for each scan (baseline rest, encoding, postencoding rest; see Fig. 4). In posterior hippocampus, we found that intervoxel similarity decreased with age in all three scans: baseline rest ($r_{(52)} = -0.43$, $p = 0.001$), encoding ($r_{(52)} = -0.43$, $p = 0.001$), and postencoding rest ($r_{(52)} = -0.35$, $p = 0.010$). By contrast, in anterior hippocampus, there was no significant association between intervoxel similarity and age during any of the scans: baseline rest ($r_{(52)} = -0.09$, $p = 0.51$), encoding ($r_{(52)} = -0.13$, $p = 0.33$), or postencoding rest ($r_{(52)} = -0.23$, $p = 0.10$). For the corresponding posterior probabilities of these outcomes, as given by Bayesian linear models, see Bayesian statistics. This pattern of results also replicated when considering left and right hippocampus separately. In both the left and right hemispheres, the correlation between age and intervoxel similarity in posterior hippocampus was significant during baseline rest (left: $r_{(52)} = 0.43$, $p = 0.001$; right: $r_{(52)} = -0.31$, $p = 0.025$), encoding (left: $r_{(52)} = -0.36$, $p = 0.008$; right: $r_{(52)} = 0.40$, $p = 0.003$), and postencoding rest (left: $r_{(52)} = -0.36$, $p = 0.008$; right: $r_{(52)} = -0.32$, $p = 0.019$). In contrast, no significant correlations were found in either hemisphere when looking within anterior hippocampus (all p values > 0.08).

We then examined whether the relationship between age and intervoxel similarity in posterior hippocampus varied across scans. The intervoxel similarity values during baseline rest, encoding, and postencoding rest did not significantly differ from each other (all p values > 0.4), suggesting that the association between posterior hippocampal granularity and age is not fully task-dependent.

Next, we tested the specificity of this effect to posterior versus anterior hippocampus. Indeed, during baseline rest, the correlation between age and intervoxel similarity was significantly stronger (i.e., more negative) in posterior relative to anterior hippocampus ($t_{(51)} = -2.73$, $p = 0.009$, $d = 0.74$). However, this difference was only marginal during encoding ($t_{(51)} = -1.99$, $p = 0.052$, $d = 0.54$), again in the direction of greater (negative) correlation strength in posterior hippocampus, and nonsignificant during postencoding rest ($t_{(51)} = -0.93$, $p = 0.36$). Together, these results provide evidence for a task-independent negative association between age and intervoxel similarity in posterior hippocampus, which during baseline rest was significantly stronger than the effect in anterior hippocampus.

Previous work examining intervoxel similarity in the hippocampus in adults has reported greater similarity (i.e., less representational granularity) in anterior relative to posterior hippocampus (Brunec et al., 2018). Interestingly, in our developmental population, intervoxel similarity was greater in posterior relative to anterior hippocampus during all three scans (baseline rest: $t_{(53)} = -5.16$, $p < 0.001$; encoding: $t_{(53)} = -4.09$, $p < 0.001$; postencoding rest: $t_{(53)} = -6.26$, $p < 0.001$). However, we also found some evidence that the difference in intervoxel similarity between these subregions (posterior − anterior hippocampus) was negatively correlated with age (baseline rest: $r_{(52)} = -0.39$, $p = 0.003$; encoding: $r_{(52)} = -0.26$, $p = 0.058$; postencoding rest: $r_{(52)} = -0.20$, $p = 0.14$). These results lend further support to the notion that the representational specificity of different hippocampal subregions changes dynamically throughout development, reaching a pattern of maturity only after adolescence.

As referenced previously, to supplement the results presented in this section, we also ran a series of Bayesian linear models investigating the relationship between age and hippocampal intervoxel similarity (see Bayesian statistics). In brief, these models provided strong support for an age-related increase in hippocampal granularity (i.e., a decrease in intervoxel similarity) in posterior hippocampus during both task and rest, while also suggesting that the same effect may be present to a weaker degree in anterior hippocampus. In light of these results indicating a stronger age-related effect in posterior hippocampus, as well as the fact that significant age × similarity correlations were observed only in posterior hippocampus when using Frequentist statistics, as described above, we focused specifically on posterior hippocampus when asking how intervoxel similarity may relate to subsequent memory behavior.

Stability of intervoxel similarity within participants
The consistency of the observed association between intervoxel similarity and age across scans suggests that such similarity is relatively stable within individuals. To test this observation more directly, we next assessed the stability of participants' intervoxel similarity values across all scans. Indeed, we found significant correlations between participants' intervoxel similarity values during baseline rest and encoding ($r_{(52)} = 0.72$, $p < 0.001$), baseline rest and postencoding rest ($r_{(52)} = 0.49$, $p < 0.001$), and encoding and
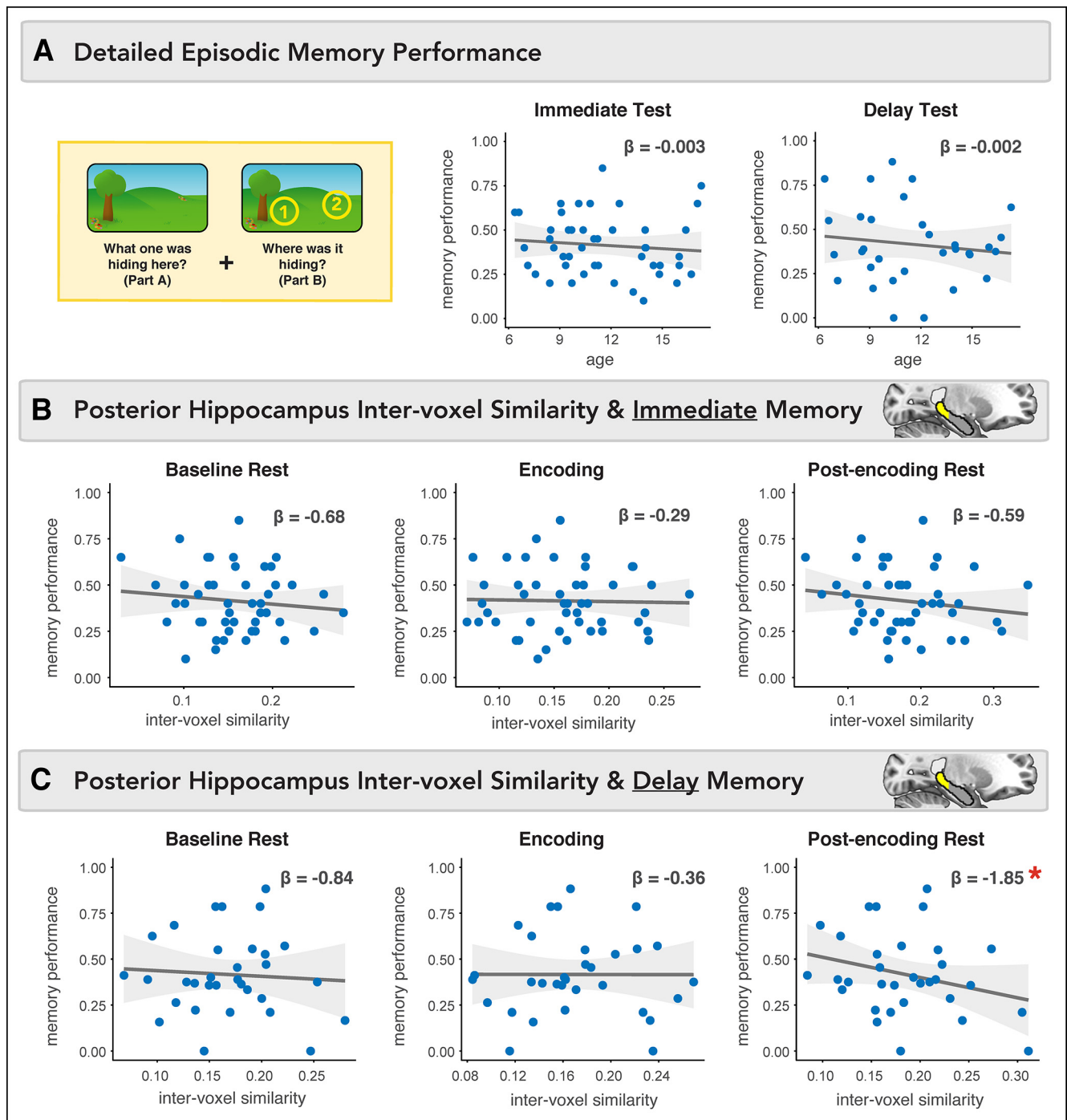
Figure 3. Intervoxel similarity in posterior hippocampus is associated with delay memory performance. ***A***, Relationship between detailed episodic memory performance (i.e., item-context-location associations) and age. Following the statistics reported in the main text, β values indicate unstandardized regression coefficients corresponding to the effect of age on memory performance, after controlling for sex (in the case of the immediate test) or sex and the number of days between encoding and delay retrieval sessions (in the case of the delay test). ***B***, Relationship between immediate detailed episodic memory and intervoxel similarity in posterior hippocampus during baseline rest (left), encoding (middle), and postencoding rest (right). ***C***, Relationship between delay detailed episodic memory and intervoxel similarity in posterior hippocampus during baseline rest (left), encoding (middle), and postencoding rest (right). ***B***, ***C***, β values indicate the unstandardized regression coefficients corresponding to the effect of intervoxel similarity on memory performance, after controlling for age. Although covariates were used in the analyses as described, in these plots, we show the raw data, rather than the partial correlations after controlling for covariates. Gray lines indicate the best fit line representing the relationship between the two specified variables. Gray ribbons represent 95% confidence intervals (CIs). ∗$p < 0.05$.

postencoding rest ($r_{(52)} = 0.59$, $p < 0.001$). However, when testing for differences between these correlation values, we found that, within participants, intervoxel similarity was more stable (i.e., showed a stronger correlation) from baseline rest to encoding versus baseline rest to postencoding rest ($t_{(51)} = 2.56$, $p = 0.01$, $d = 0.70$). No other significant differences

were observed between across-scan correlations (all p values > 0.19). This pattern of results suggests that intervoxel similarity is both an individual difference measure as it is relatively stable within individuals, but is also influenced by task demands, particularly the transition from encoding to memory stabilization/consolidation.

Table 1. Influence of confound variables on intervoxel similarity in posterior hippocampus[a]

Sex	β	SE	t	tSNR	β	SE	t
Baseline rest				*Baseline rest*			
Age	−0.007	0.002	−3.398**	Age	−0.007	0.002	−3.287**
Sex	−0.003	0.006	−0.540	tSNR	−0.000	0.000	−0.270
Encoding				*Encoding*			
Age	−0.007	0.002	−3.336**	Age	−0.007	0.002	−2.930**
Sex	−0.002	0.006	−0.369	tSNR	−0.000	0.000	−0.053
Postencoding rest				*Postencoding rest*			
Age	−0.007	0.003	−2.590*	Age	−0.007	0.003	−2.594*
Sex	−0.009	0.008	−1.065	tSNR	0.000	0.000	0.010
No. of voxels				Motion (FWD)			
Baseline rest				*Baseline rest*			
Age	−0.007	0.002	−3.222**	Age	−0.007	0.002	−3.229**
No. of voxels	−0.001	0.001	−0.806	FWD	0.027	0.032	0.848
Encoding				*Encoding*			
Age	−0.007	0.002	−3.120**	Age	−0.008	0.002	−3.132**
No. of voxels	−0.001	0.001	−0.945	FWD	−0.013	0.030	−0.435
Postencoding rest				*Postencoding rest*			
Age	−0.007	0.003	−2.333*	Age	−0.007	0.003	−2.589*
No. of voxels	−0.001	0.001	−0.670	FWD	−0.005	0.031	−0.175
Sleepiness				Univariate activation			
Baseline rest				*Encoding*			
Age	−0.007	0.003	−2.503*	Age	−0.008	0.002	−3.791***
Sleepiness	0.017	0.020	0.819	Univariate	0.039	0.016	2.502*
Encoding				Spontaneous recall			
Age	−0.008	0.003	−3.013**				
Sleepiness	0.029	0.019	1.512				
Postencoding rest				*Postencoding rest*			
Age	−0.008	0.004	−2.103*	Age	−0.009	0.004	−2.333*
Sleepiness	0.033	0.028	1.182	Spontaneous retrieval	−0.037	0.023	−1.579

[a] β values represent unstandardized regression coefficients. No. of voxels = number of voxels. tSNR represents temporal signal to noise ratio. SE represents the standard error. t represents t values from the regression.
***p < 0.001; **p < 0.01; *p < 0.05.

Role of potential confound variables
We next aimed to test whether the observed negative relationship between intervoxel similarity in posterior hippocampus and age might be an artifact of potential confound variables, including the following: participant sex, motion in the scanner, number of voxels in each ROI, tSNR in each ROI, univariate activity in the posterior hippocampus during encoding, sleepiness during the scans, and spontaneous recall of task information during the postencoding rest period. Importantly, the relationship between age and intervoxel similarity in posterior hippocampus during each scan held after controlling for participant sex, motion (i.e., mean FWD during each scan), tSNR in the posterior hippocampus, the number of voxels in the posterior hippocampal ROI for each scan, and participant-reported sleepiness (Table 1). For the encoding scan, the relationship between age and intervoxel similarity in posterior hippocampus also held after controlling for univariate task-based activity in the same ROI. For the postencoding rest scan, age remained a significant predictor of intervoxel similarity in posterior hippocampus when controlling for self-reported spontaneous recall of task information during postencoding rest, whereas posterior hippocampus intervoxel similarity during postencoding rest and spontaneous recall were not significantly associated.

Of note, we did observe a significant difference in tSNR between anterior and posterior hippocampus during baseline rest (bilateral: $t_{(53)} = -8.75$, $p < 0.001$; left: $t_{(53)} = -9.30$, $p < 0.001$; right: $t_{(53)} = -6.99$, $p < 0.001$), encoding (bilateral: $t_{(53)} = -10.07$, $p < 0.001$; left: $t_{(53)} = -10.26$, $p < 0.001$; right: $t_{(53)} = -8.32$, $p < 0.001$),

and postencoding rest (bilateral: $t_{(53)} = -9.18$, $p < 0.001$; left: $t_{(53)} = -9.50$, $p < 0.001$; right: $t_{(53)} = -6.94$, $p < 0.001$). However, as mentioned above and reported in Table 1, tSNR in the posterior hippocampus did not explain the relationship between age and intervoxel similarity.

Together, these findings suggest that representational granularity in the posterior hippocampus, during both task and rest, increases throughout middle childhood and adolescence. This relationship, moreover, is not explained by participant sex, motion, ROI size, univariate task-based hippocampal activity, tSNR, sleepiness, or spontaneous recall.

Specificity of age × intervoxel similarity relationship to hippocampus
In order to assess the specificity of the observed relationship between age and representational granularity to the hippocampus, we also tested age × intervoxel similarity associations in four additional cortical ROIs known to contribute to episodic memory function (for visual stimuli): PhC, mPFC, PMC, and OTC (Fig. 5). No significant correlations between age and intervoxel similarity were observed during any scan in mPFC (baseline rest: $r_{(52)} = -0.04$, $p = 0.77$; encoding: $r_{(52)} = -0.14$, $p = 0.31$; postencoding rest: $r_{(52)} = -0.01$, $p = 0.96$), PMC (baseline rest: $r_{(52)} = 0.02$, $p = 0.88$; encoding: $r_{(52)} = 0.24$, $p = 0.08$; postencoding rest: $r_{(52)} = -0.26$, $p = 0.06$), or OTC (baseline rest: $r_{(52)} = -0.08$, $p = 0.56$; encoding: $r_{(52)} = 0.15$, $p = 0.27$; postencoding rest: $r_{(52)} = 0.05$, $p = 0.71$).

Table 2. Relationship between detailed episodic memory performance and intervoxel similarity (controlling for age) in anterior hippocampus[a]

Immediate detailed episodic memory				Delay detailed episodic memory			
Anterior hippocampus				Anterior hippocampus			
	β	SE	t		β	SE	t
Baseline rest				Baseline rest			
Intervoxel similarity	−0.805	0.624	−1.290	Intervoxel similarity	−0.874	1.051	−0.831
Age	−0.007	0.008	−0.793	Age	−0.011	0.013	−0.865
Encoding				Encoding			
Intervoxel similarity	−0.412	0.450	−0.824	Intervoxel similarity	−0.195	0.820	−0.238
Age	−0.007	0.008	−0.792	Age	−0.009	0.013	−0.715
Postencoding rest				Postencoding rest			
Intervoxel similarity	−1.084	0.518	−2.093*	Intervoxel similarity	−1.266	1.018	−1.243
Age	−0.010	0.008	−1.159	Age	−0.012	0.013	−0.979

[a] β values represent unstandardized regression coefficients. SE represents the standard error. t represents the t values from the regression.
*$p < 0.05$.

Table 3. Relationship between intervoxel similarity in anterior and posterior hippocampus and coarse memory performance[a]

Immediate coarse episodic memory				Delay coarse episodic memory				Retention: coarse episodic memory			
Anterior hippocampus				Anterior hippocampus				Anterior hippocampus			
	β	SE	t		β	SE	t		β	SE	t
Baseline rest				Baseline rest				Baseline rest			
Intervoxel similarity	0.005	0.587	0.009	Intervoxel similarity	−0.912	1.056	−0.864	Intervoxel similarity	0.222	0.811	0.274
Age	−0.012	0.008	−1.594	Age	−0.003	0.013	−0.226	Age	0.016	0.009	1.754
Encoding				Encoding				Encoding			
Intervoxel similarity	0.001	0.465	0.003	Intervoxel similarity	0.112	0.825	0.136	Intervoxel similarity	0.685	0.587	1.166
Age	−0.012	0.008	−1.582	Age	−0.001	0.013	−0.043	Age	0.017	0.009	1.837
Postencoding rest				Postencoding rest				Postencoding rest			
Intervoxel similarity	−0.527	0.495	−1.064	Intervoxel similarity	−1.133	1.029	−1.100	Intervoxel similarity	0.448	0.749	0.599
Age	−0.014	0.008	−1.820	Age	−0.004	0.013	−0.295	Age	0.017	0.009	1.828
Posterior hippocampus				Posterior hippocampus				Posterior hippocampus			
Baseline rest				Baseline rest				Baseline rest			
Intervoxel similarity	−0.624	0.518	−1.206	Intervoxel similarity	−1.579	0.930	−1.698	Intervoxel similarity	−0.832	0.712	−1.168
Age	−0.017	0.008	−1.988	Age	−0.013	0.014	−0.931	Age	0.010	0.010	0.984
Encoding				Encoding				Encoding			
Intervoxel similarity	−0.247	0.530	−0.467	Intervoxel similarity	−0.941	0.923	−1.020	Intervoxel similarity	−0.034	0.750	−0.045
Age	−0.014	0.009	−1.652	Age	−0.007	0.014	−0.528	Age	0.016	0.010	1.543
Postencoding rest				Postencoding rest				Postencoding rest			
Intervoxel similarity	−0.424	0.392	−1.082	Intervoxel similarity	−1.765	0.786	−2.245*	Intervoxel similarity	−0.999	0.577	−1.732
Age	−0.015	0.008	−1.897	Age	−0.016	0.014	−1.191	Age	0.007	0.010	0.743

[a] β values represent unstandardized regression coefficients. SE represents the standard error. t represents the t values from the regression.
*$p < 0.05$.

However, there was some evidence that intervoxel similarity in PhC also decreases with age: specifically, significant correlations between these measures were observed during baseline rest ($r_{(52)} = -0.27$, $p = 0.045$) and encoding ($r_{(52)} = -0.31$, $p = 0.023$), but not postencoding rest ($r_{(52)} = -0.20$, $p = 0.15$), a finding that is perhaps unsurprising, given the strong anatomic and functional connectivity between the posterior hippocampus and PhC (Furtak et al., 2007; Aggleton et al., 2012; Ranganath and Ritchey, 2012). Ultimately, these results suggest that the age-related decrease in representational granularity we observed in the posterior hippocampus is relatively specific to the posterior MTL but may extend outside of the hippocampus proper in a manner consistent with known MTL pathways.

Relationship between intervoxel similarity in posterior hippocampus and memory performance

Given the observed relationship between representational granularity in posterior hippocampus and age, we next tested whether intervoxel similarity in this subregion was associated with participants' detailed episodic memory performance (i.e., memory for item-context-location associations). Detailed episodic memory was chosen as the behavioral variable here for two reasons: (1) given this measure's additional requirement of correct retrieval of specific item locations, it reflects a greater demand on memory function; and (2) because we hypothesized that intervoxel similarity might be related to the specificity of associative memory representations. Additionally, we decided to analyze immediate and delay memory behavior separately, rather than using a combined retention measure, as we had more complete data for immediate ($N = 49$) and delay ($N = 35$) memory than for retention ($N = 32$) memory. In other words, there was a subset of participants who completed the delay memory test who did not have corresponding immediate test data.

To this end, we performed linear regressions in which memory performance was predicted by intervoxel similarity in posterior hippocampus, separately for each scan (baseline rest, encoding, and postencoding rest). All regressions included

Table 4. Relationship between detailed episodic memory performance and intervoxel similarity in cortical regions of interest (ROIs)[a]

Immediate detailed episodic memory				Delay detailed episodic memory			
mPFC				mPFC			
	β	SE	t		β	SE	t
Baseline rest				Baseline rest			
Intervoxel similarity	0.401	0.814	0.492	Intervoxel similarity	−0.928	1.167	−0.795
Age	−0.005	0.008	−0.615	Age	−0.009	0.012	−0.722
Encoding				Encoding			
Intervoxel similarity	0.586	0.896	0.654	Intervoxel similarity	0.571	1.150	0.496
Age	−0.004	0.009	−0.485	Age	−0.008	0.013	−0.595
Postencoding rest				Postencoding rest			
Intervoxel similarity	0.104	0.730	0.143	Intervoxel similarity	−1.891	1.106	−1.710
Age	−0.006	0.008	−0.669	Age	−0.008	0.013	−0.584
PMC				PMC			
Baseline rest				Baseline rest			
Intervoxel similarity	0.259	0.383	0.676	Intervoxel similarity	0.292	0.904	0.323
Age	−0.006	0.008	−0.693	Age	−0.010	0.013	−0.760
Encoding				Encoding			
Intervoxel similarity	0.658	0.486	1.354	Intervoxel similarity	1.203	0.639	1.883
Age	−0.008	0.008	−0.927	Age	−0.017	0.013	−1.319
Postencoding rest				Postencoding rest			
Intervoxel similarity	−0.119	0.339	−0.351	Intervoxel similarity	−0.726	0.696	−1.042
Age	−0.006	0.009	−0.742	Age	−0.009	0.012	−0.736
OTC				OTC			
Baseline rest				Baseline rest			
Intervoxel similarity	0.160	0.481	0.331	Intervoxel similarity	0.234	0.796	0.294
Age	−0.005	0.008	−0.629	Age	−0.009	0.013	−0.670
Encoding				Encoding			
Intervoxel similarity	0.022	0.678	0.032	Intervoxel similarity	0.369	0.973	0.379
Age	−0.006	0.008	−0.671	Age	−0.010	0.013	−0.785
Postencoding rest				Postencoding rest			
Intervoxel similarity	−0.878	0.501	−1.752	Intervoxel similarity	−0.809	0.845	−0.957
Age	−0.005	0.008	−0.651	Age	−0.008	0.012	−0.619
PhC				PhC			
Baseline rest				Baseline rest			
Intervoxel similarity	−0.444	0.488	−0.910	Intervoxel similarity	0.465	0.797	0.583
Age	−0.008	0.009	−0.901	Age	−0.006	0.013	−0.488
Encoding				Encoding			
Intervoxel similarity	0.258	0.340	0.759	Intervoxel similarity	0.575	0.495	1.163
Age	−0.003	0.009	−0.381	Age	−0.004	0.013	−0.306
Postencoding rest				Postencoding rest			
Intervoxel similarity	−0.163	0.417	−0.392	Intervoxel similarity	−0.051	0.666	−0.077
Age	−0.006	0.009	−0.732	Age	−0.009	0.013	−0.707

[a] β values represent unstandardized regression coefficients. SE represent standard errors. t represents t values from the regressions. mPFC medial prefrontal cortex. PMC posterior medial cortex. OTC occipital temporal cortex. PhC parahippocampal cortex.

participant age as a covariate. See Table 2 for results in anterior hippocampus. Also see Table 3 for results obtained when considering coarse rather than detailed episodic memory performance. For completeness, we also searched for potential relationships between behavior and intervoxel similarity in our cortical control ROIs (mPFC, PMC, OTC, and PhC), but no significant effects were found (Table 4).

Immediate memory performance
Memory performance on the immediate test was not associated with intervoxel similarity in posterior hippocampus during any of the three scans (Fig. 3B): baseline rest, $\beta = -0.68$, $t_{(43)} = -1.22$, $p = 0.23$ (effect of age: $\beta = -0.01$, $t_{(43)} = -1.14$, $p = 0.26$), encoding, $\beta = -0.29$, $t_{(43)} = -0.50$, $p = 0.62$ (effect of age: $\beta = -0.01$, $t_{(43)} = -0.82$, $p = 0.42$), and postencoding rest, $\beta = -0.59$, $t_{(43)} = -1.40$, $p = 0.17$ (effect of age: $\beta = -0.01$, $t_{(43)} = -1.13$, $p = 0.27$).

Delay memory performance
Memory performance on the delay test was also not associated with intervoxel similarity in posterior hippocampus during baseline rest, $\beta = -0.84$, $t_{(29)} = -0.88$, $p = 0.39$ (effect of age: $\beta = -0.02$, $t_{(29)} = -1.06$, $p = 0.30$) or encoding, $\beta = -0.36$, $t_{(29)} = -0.39$, $p = 0.70$ (effect of age: $\beta = -0.01$, $t = -0.81$, $p = 0.43$). However, we did find that intervoxel similarity during postencoding rest was significantly associated with delay memory performance, $\beta = -1.85$, $t_{(29)} = -2.38$, $p = 0.024$, partial $r^2 = 0.16$ (effect of age: $\beta = -0.03$, $t_{(29)} = -1.87$, $p = 0.07$), such that lower intervoxel similarity (i.e., greater granularity) during the postencoding rest period was linked to better delay memory (Fig. 3C). A similar pattern of results was obtained when using Bayesian linear models rather than Frequentist statistics, as detailed in the following section. Moreover, results were largely unchanged when examining left and right posterior hippocampus separately: the association between delay memory
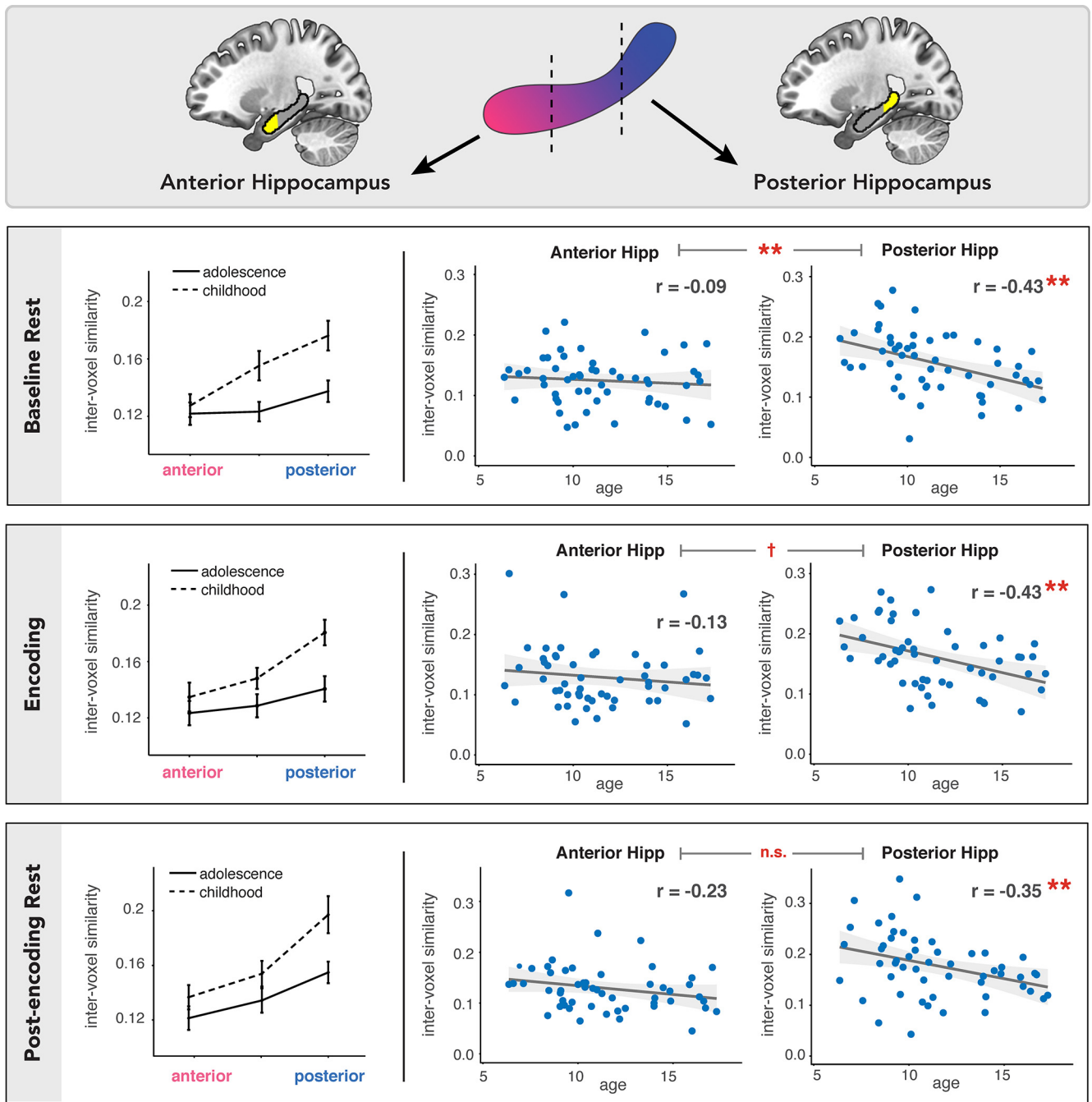
Figure 4. Intervoxel similarity in posterior hippocampus (hipp) decreases with age. In posterior hippocampus, we observed an age-related decrease in intervoxel similarity during both encoding and rest periods. This effect was not seen in anterior hippocampus. Leftmost column plots, Mean intervoxel similarity in anterior, middle, and posterior thirds of the hippocampus for participants older ("adolescence") and younger ("childhood") than 11 years (for visualization purposes). Error bars indicate standard error of the mean (SEM). In the two rightmost columns of the figure, gray lines on each plot indicate the best fit line representing the relationship between intervoxel similarity and age. Gray ribbons represent 95% confidence intervals (CIs). r values represent the correlation between the variables on the X and Y axis of the plots. Tests for significant differences between anterior and posterior hippocampus correlations are represented on the top of each of the three right panels. $**p < 0.01$. $†p < 0.1$. n.s. not significant.

performance and intervoxel similarity during postencoding rest was significant in the left hippocampus, $\beta = -1.37$, $t_{(29)} = -2.46$, $p = 0.020$, partial $r^2 = 0.17$ (effect of age: $\beta = -0.02$, $t_{(29)} = -1.75$, $p = 0.09$) and marginal in the right hippocampus, $\beta = -1.39$ $t_{(29)} = -1.77$, $p = 0.09$, partial $r^2 = 0.10$ (effect of age: $\beta = -0.02$, $t_{(29)} = -1.57$, $p = 0.13$), while no significant effects were found during baseline rest or encoding in either hemisphere (all p values > 0.2).

In order to examine whether the observed relationship between delay memory performance and intervoxel similarity during postencoding rest was statistically stronger than

that observed during other scans, we next looked for significant differences in the partial correlations between intervoxel similarity and memory performance (after controlling for age) across scans. Indeed, the correlation between intervoxel similarity and performance was marginally stronger (i.e., more negative) during postencoding rest relative to encoding ($t_{(29)} = 1.95$, $p = 0.06$), but not relative to baseline rest ($t_{(29)} = 1.45$, $p = 0.16$).

Next, we checked whether any of the confound variables considered in the relationship between age and intervoxel similarity
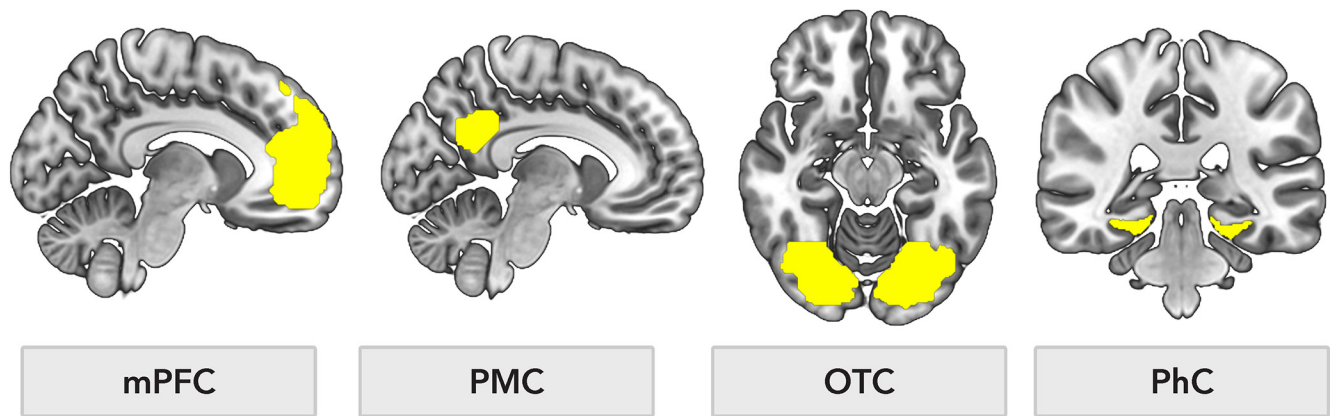
Figure 5. Nonhippocampal regions of interest (ROIs). Visualization of cortical control regions used in the current study. From left to right: medial prefrontal cortex (mPFC), posterior medial cortex (PMC), occipital temporal cortex (OTC), and parahippocampal cortex (PhC).

Table 5. Influence of confound variables on the relationship between detailed delay episodic memory performance and intervoxel similarity in posterior hippocampus[a]

Sex	β	SE	t	tSNR	β	SE	t
Postencoding rest				*Postencoding rest*			
Intervoxel similarity	−1.849	0.786	−2.351*	Intervoxel similarity	−1.816	0.775	−2.345*
Age	−0.025	0.014	−1.843	Age	−0.027	0.013	−1.991
Sex	0.015	0.037	0.417	tSNR	0.001	0.001	1.038
No. of voxels				Motion (FWD)			
Postencoding rest				*Postencoding rest*			
Intervoxel similarity	−1.724	0.807	−2.137*	Intervoxel similarity	−1.712	0.759	−2.256*
Age	−0.026	0.014	−1.911	Age	−0.027	0.013	−2.082*
No. of voxels	0.003	0.005	0.630	FWD	−0.355	0.220	−1.611
Sleepiness				Univariate encoding activity			
Postencoding rest				*Postencoding rest*			
Intervoxel similarity	−1.947	0.834	−2.335*	Intervoxel similarity	−1.765	0.787	−2.242*
Age	−0.020	0.014	−1.380	Age	−0.024	0.013	−1.814
Sleepiness	0.024	0.100	0.241	Univariate	0.074	0.096	0.771
Spontaneous recall				Train-test interval[b]			
Postencoding rest				*Postencoding rest*			
Intervoxel similarity	−1.343	0.695	−1.933	Intervoxel similarity	−1.804	0.860	−2.098*
Age	−0.009	0.013	−0.665	Age	−0.022	0.013	−1.695
Recall	0.164	0.079	2.080	Train-test interval	0.012	0.006	1.804

[a]β values represent unstandardized regression coefficients. No. of voxels represents the number of voxels in the region of interest. tSNR represents the temporal signal to noise ratio. FWD represents the framewise displacement in motion.
[b]One outlier participant was excluded from this model with train-test interval > 1 month.
*$p < 0.05$.

could explain the association between similarity and delay memory performance. As shown in Table 5, the effect of intervoxel similarity in posterior hippocampus during postencoding rest on memory performance remained significant (when also accounting for age) after controlling for participant sex, motion (i.e., FWD), tSNR, number of voxels in each ROI, univariate activity in the posterior hippocampus during encoding, reported sleepiness, and the train-test interval. When controlling for self-reported spontaneous recall of task information during the postencoding rest scan, the effect of intervoxel similarity became marginal ($p = 0.068$). However, it is worth noting that lack of power may have contributed to the lack of a significant effect in this final model, given that not all participants completed the self-report questionnaire (see Materials and Methods). Ultimately, these results suggest that increased representational

granularity in the posterior hippocampus after learning may facilitate the ability to retrieve detailed episodic memories after a delay.

Memory retention
Finally, we examined whether intervoxel similarity in the posterior hippocampus was associated with the retention of learned associations from the immediate to the delay memory test (calculated as immediate − delay performance). However, memory retention was not associated with intervoxel similarity in posterior hippocampus during baseline rest, $\beta = -0.58$, $t_{(26)} = -0.91$, $p = 0.37$ (effect of age: $\beta = -0.00$, $t_{(26)} = -0.04$, $p = 0.97$), or encoding, $\beta = -0.21$, $t_{(26)} = 0.32$, $p = 0.75$ (effect of age: $\beta = 0.01$, $t_{(26)} = 0.55$, $p = 0.59$). There was a marginal (though not significant) association between memory retention and intervoxel

similarity in the posterior hippocampus during postencoding rest, $\beta = -1.01$, $t_{(26)} = -1.99$, $p = 0.06$, partial $r^2 = 0.13$ (effect of age: $\beta = -0.01$, $t_{(26)} = -0.55$, $p = 0.59$).

Bayesian statistics
To augment the results of the Frequentist linear models described thus far, we also constructed Bayesian linear models to assess relationships between our variables of interest (age, intervoxel similarity, and memory performance).

Association between age and intervoxel similarity
Separate Bayesian linear models were created to assess the relationship between age and intervoxel similarity in each of our three scans (baseline rest, encoding, and postencoding rest) and both hippocampal subregions (anterior and posterior). Looking first in posterior hippocampus, we again found a reliable, negative association between intervoxel similarity and age during baseline rest ($\beta = -0.41$, 95% CI $= [-0.65, -0.17]$), encoding ($\beta = -0.40$, 95% CI $= [-0.64, -0.16]$), and postencoding rest ($\beta = -0.40$, 95% CI $= [-0.70, -0.10]$), such that older participants showed lower intervoxel similarity (i.e., greater representational granularity) than younger participants. As one way to interpret the strength and reliability of this negative association, we can examine directional posterior probabilities, or the probability that each slope term (β) was <0. A posterior probability of $p > 0.5$ indicates evidence in favor of a negative slope, a probability of $p < 0.5$ indicates evidence in favor of a positive slope, and a probability of $p = 0.5$ indicates that we cannot be confident either way. The posterior probability that age was negatively associated with intervoxel similarity was $p > 0.99$ in each of the three scans, suggesting that we can be very confident in the presence of this effect during both task and rest.

In anterior hippocampus, we also see some evidence for a negative association between intervoxel similarity and age, although the strength of this effect varied across scans: baseline rest ($\beta = -0.07$, 95% CI $= [-0.28, 0.14]$), encoding ($\beta = -0.12$, 95% CI $= [-0.37, 0.12]$), and postencoding rest ($\beta = -0.19$, 95% CI $= [-0.43, 0.03]$). The posterior probability that $\beta < 0$ was 0.74 during baseline rest, 0.84 during encoding, and 0.95 during postencoding rest, suggesting a notably weaker effect than that observed in posterior hippocampus.

In order to more explicitly test for a difference in this age × intervoxel similarity effect between anterior and posterior hippocampus, we next created new linear models for each scan, in which intervoxel similarity was predicted by age, hippocampal subregion, and their interaction. We found evidence for a negative interaction between age and hippocampal subregion, such that age was more strongly associated with greater representational granularity (i.e., lower intervoxel similarity) in posterior relative to anterior hippocampus during all three scans: baseline rest (interaction: $\beta = -0.17$, 95% CI $= [-0.32, -0.01]$; age: $\beta = -0.24$, 95% CI $= [-0.39, -0.08]$; subregion: $\beta = 0.30$, 95% CI $= [0.15, 0.46]$), encoding (interaction: $\beta = -0.14$, 95% CI $= [-0.31, 0.04]$; age: $\beta = -0.26$, 95% CI $= [-0.43, -0.09]$; subregion: $\beta = 0.30$, 95% CI $= [0.13, 0.47]$), and postencoding rest (interaction: $\beta = -0.10$, 95% CI $= [-0.29, 0.08]$; age: $\beta = -0.29$, 95% CI $= [-0.48, -0.11]$; subregion: $\beta = 0.44$, 95% CI $= [0.26, 0.63]$). The posterior probability that the age × subregion interaction term was <0 was 0.98 during baseline rest, 0.94 during encoding, and 0.86 during postencoding rest. Together, these results suggest that, although this age × subregion interaction may weaken after learning, we can be reasonably confident in the claim that age has a stronger

negative relationship with intervoxel similarity in posterior versus anterior hippocampus.

Relationship between intervoxel similarity in posterior hippocampus and memory performance
Next, we investigated whether greater representational granularity (i.e., lower intervoxel similarity) in posterior hippocampus was associated with detailed episodic memory performance, either on the immediate or delay test. Specifically, Bayesian linear models were constructed with memory performance predicted by intervoxel similarity, while controlling for the effect of age.

First, focusing on the immediate test, we found that memory performance was not clearly associated with intervoxel similarity in posterior hippocampus during encoding ($\beta = -0.02$, 95% CI $= [-0.08, 0.05]$; effect of age: $\beta = -0.02$, 95% CI $= [-0.08, 0.03]$) but showed weak evidence for a negative association during baseline rest ($\beta = -0.04$, 95% CI $= [-0.10, 0.03]$; effect of age: $\beta = -0.03$, 95% CI $= [-0.09, 0.02]$) and postencoding rest ($\beta = -0.03$, 95% CI $= [-0.08, 0.02]$; effect of age: $\beta = -0.03$, 95% CI $= [-0.08, 0.02]$). The posterior probability that the β value for intervoxel similarity was <0 was 0.88 during baseline rest, 0.69 during encoding, and 0.91 during postencoding rest.

Next, we turned to the delay test. Most notably, we found evidence for a negative association between memory performance and intervoxel similarity during postencoding ($\beta = -0.10$, 95% CI $= [-0.19, -0.01]$; effect of age: $\beta = -0.07$, 95% CI $= [-0.16, 0.01]$), such that greater granularity (i.e., lower intervoxel similarity) after learning was associated with better memory ~2 weeks later. In this model, the posterior probability that the intervoxel similarity β value was <0 was 0.99. During baseline rest, in contrast, there was weak evidence for a negative association between memory performance and intervoxel similarity (intervoxel similarity: $\beta = -0.05$, 95% CI $= [-0.15, 0.06]$; age: $\beta = -0.05$, 95% CI $= [-0.14, 0.04]$). During encoding, there was again no clear evidence for an association between memory performance and intervoxel similarity (intervoxel similarity: $\beta = -0.02$, 95% CI $= [-0.12, 0.09]$; age: $\beta = -0.03$, 95% CI $= [-0.12, 0.06]$). The posterior probabilities that the intervoxel similarity β values in these models were <0 were 0.80 and 0.65, respectively.

Ultimately, these results provide further support for the claim that greater granularity in posterior hippocampus during rest, particularly after learning, is associated with better detailed episodic memory (i.e., memory for item-scene-location associations) within this age group.

Discussion
We report here age-related differences in hippocampal granularity along the anterior-posterior axis, whereby granularity in the posterior hippocampus increased with age. Interestingly, although this relationship was present during all three phases of our experiment (baseline rest, encoding, and postencoding rest), granularity during the postencoding rest scan only, a putative memory stabilization period, was associated with detailed episodic memory performance. Hence, posterior hippocampus may represent information at high levels of granularity to support the recall of specific event details.

Traditional models of memory development have often identified cortical structures, such as mPFC, which mature at slower rates (relative to the hippocampus), as the loci of maturational improvements in memory after early childhood (DeMaster and Ghetti, 2013). Almost diametrically opposed to this idea is the suggestion that the hippocampus is functionally dormant, and does not contribute to memory in childhood (Foster and

Burman, 2010). In contrast to both these accounts are data demonstrating that the hippocampus does indeed contribute to changes in memory across development (Ghetti et al., 2010; Lee et al., 2014; DeMaster et al., 2016; Riggins et al., 2016; Sastre et al., 2016; Schlichting et al., 2017; Selmeczy et al., 2019), gradually maturing into its role as a critical node within the larger memory circuit (Foster and Burman, 2010; Travaglia et al., 2016). By using multivariate analyses to reveal subtle variations in hippocampal activity patterns across age, our results provide evidence that, at least in posterior regions, hippocampal function continues to mature across middle childhood and adolescence.

In the current study, we did not observe age-related improvements in memory on the detailed associative task, although there was a trending association between age and memory retention. While this is not inconsistent with some past work, which reveals a nuanced picture of memory maturation across childhood and adolescence (both supporting, e.g., DeMaster and Ghetti, 2013; and not supporting, e.g., Ghetti et al., 2010; Lee et al., 2014; such continued maturation), it does prevent us from making strong inferences about the association of hippocampal granularity and associative memory maturation per se However, as behavioral differences can complicate the interpretation of developmental effects in brain data (Schlaggar et al., 2002), in this study we do have the advantage of interpreting age-related changes in the brain while behavioral performance is held constant. Nonetheless, future work should aim to replicate these findings on a similar task that is susceptible to age-related improvements in memory.

It is interesting to note that we only observed a significant relationship between behavior and intervoxel similarity during the postencoding period (i.e., a potential window for memory stabilization). While it must be acknowledged that issues of power may have contributed to this result, if the brain-behavior associations we report here are indeed specific to the postencoding rest period, it would support and extend past work demonstrating that hippocampal activity during offline, postencoding rest periods (e.g., connectivity with lateral occipital cortex, as well as encoding pattern replay/reinstatement) is associated with memory performance, often independently from encoding activity (Tambini et al., 2010; Tambini and Davachi, 2013, 2019; Murty et al., 2017).

One strength of this study is the presence of both recent (immediate) and remote (~2 week) retention intervals, a feature not commonly seen in memory research in either youth or adult populations (but see Dandolo and Schwabe, 2018; Sekeres et al., 2018). Using this design, we found that postencoding granularity in posterior hippocampus was important for detailed remote (but not immediate) memory recall. This distinction could stem from several factors, not necessarily mutually exclusive. Perhaps most obvious is that the postencoding period may support processes related to long-term memory stabilization/consolidation, and thus would not necessarily relate to immediate performance. Although systems consolidation occurs hours to days after learning, such processes can be initiated soon after learning in the form of memory stabilization (Squire et al., 2015). It is also possible that tests of recent memory might not be sensitive markers of age-related brain change during middle childhood and adolescence, when learning is already strong. Instead, more challenging long-term memory tests might be better suited to reveal subtle developmental gradients within the brain at these ages. Another possibility could be that representational granularity in the posterior hippocampus after learning indicates the presence of stronger, more robust memory representations, which are more likely

to be remembered at remote intervals. Finally, it is also worth noting that, as the immediate and delay tests were identical, the delay test includes a source memory component not present in the immediate test. Specifically, on the delayed test, participants were required to remember a scene-item pairing from the initial study phase and to reduce interference from the immediate test. Assessing how post-learning rest periods map onto ongoing memory stabilization processes, and what aspect of this process intervoxel similarity represents, could be the focus of future work.

Another factor worth considering in understanding the contributions of hippocampal granularity to detailed memory was that our detailed memory measure relied strongly on the recall of object location within a scene, thereby pointedly taxing spatial memory processes. Prior work has shown that posterior parahippocampal gyrus is more active during viewing and successful remembering of scene stimuli. Current models of hippocampal function posit that perirhinal, entorhinal, and parahippocampal inputs to hippocampus (Davachi, 2006; Eichenbaum et al., 2007; Awipi and Davachi, 2008; Staresina et al., 2011) may bias different regions of the hippocampus toward the encoding and consolidation of different memory content. Interestingly, we observed increases in intervoxel granularity with age in both posterior hippocampus and PhC, perhaps reflecting the known anatomic and functional interconnectedness of these two regions, as well as the spatial nature of the task. However, only intervoxel similarity in posterior hippocampus was associated with memory behavior. This pattern of results is consistent with the idea that, while PhC is important in the processing and remembering of spatial details, it is the hippocampus that is specifically critical for binding this information with item-level information (Ranganath and Ritchey, 2012), as was required for memory in the current task. Future work can test this hypothesis by examining age-related changes in memory for nonassociative and nonspatial (e.g., item or social) information.

The associations between intervoxel similarity, brain, and behavior reported here highlight a critical issue: what exactly does intervoxel granularity measure? One possibility is that intervoxel granularity represents some form of sparse coding, akin to a temporal pattern separation process (LaRocque et al., 2013; Favila et al., 2016; Chanales et al., 2017). This process should be differentiated from the more typical spatial pattern separation, which focuses on the creation of maximally orthogonal neural ensembles responding to stimuli with overlapping features, and is associated with the dentate gyrus (Yassa and Stark, 2011). In contrast, a temporal pattern separation pattern would reflect temporal fluctuations across neighboring voxels becoming more distinctive with age, occurring in posterior hippocampus. Another hypothesis is that the continuum from intervoxel similarity to granularity is representative of the principles of "interactive specialization" (Johnson, 2011). Specifically, in this theory, brain regions (typically cortical) are hypothesized to begin development with broad functionality, being activated in a wide range of task contexts, before activity-dependent interactions hone the response properties of the region, increasing its specialization. By this account, increasing posterior hippocampal granularity with age may represent increasing specialization and neural sharpening, building the capacity for more precise behavioral correlates.

In contrast to these more mechanistic interpretations of the data, yet another suggestion might be that granularity instead represents a noisier or less temporally stable signal because of subregional proximity to different vasculature and areas of signal dropout (e.g., the nasal cavity) (Spallazzi et al., 2019). However,

including tSNR as a covariate in our analyses did not change the association of intervoxel similarity with age, and intervoxel similarity was relatively stable within individuals. Together, these results suggest that granularity is a relatively stable individual difference marker, which represents a true developmental signal beyond mere noise.

One point of difference with the intervoxel similarity outcomes reported here and those in past work (e.g., Brunec et al., 2018) concerns the distribution of similarity gradients across the hippocampal long axis. In adults (Brunec et al., 2018), intervoxel similarity was highest in the anterior hippocampus, whereas granularity was more characteristic of posterior hippocampus. Here, we observed that intervoxel patterns were actually most similar in the posterior hippocampus for the majority of development, and reached comparably levels of low similarity as the anterior hippocampus only later in adolescence. This might suggest that the dissociation of intervoxel similarity in anterior and posterior hippocampus reported previously in adults emerges slowly across development, as a result of age decrements in posterior similarity. Future studies could look at this phenomenon across a sample of individuals that cover the ages reported in both this paper (6-17 years) as well as in the previously published study (19-30 years) (Brunec et al., 2018).

In conclusion, the data presented here contribute to an ongoing body of work demonstrating the importance of hippocampal activity in offline windows for memory performance. Specifically, spontaneous posterior hippocampus correlations are inversely related to age and remote memory performance. By taking a developmental approach, we highlight that the posterior hippocampus may gradually build up its capacity to support detailed episodic recall, introducing compelling and novel ideas about the role of sparse neural coding in memory stabilization.

References

Aggleton JP, Wright NF, Vann SD, Saunders RC (2012) Medial temporal lobe projections to the retrosplenial cortex of the macaque monkey. Hippocampus 22:1883–1900.

Awipi T, Davachi L (2008) Content-specific source encoding in the human medial temporal lobe. J Exp Psychol Learn Mem Cogn 34:769–779.

Bauer PJ, Burch MM, Scholin SE, Güler OE (2007) Using cue words to investigate the distribution of autobiographical memories in childhood. Psychol Sci 18:910–916.

Billingsley RL, Smith ML, McAndrews MP (2002) Developmental patterns in priming and familiarity in explicit recollection. J Exp Child Psychol 82:251–277.

Blankenship TL, O'Neill M, Ross A, Bell MA (2015) Working memory and recollection contribute to academic achievement. Learn Individ Differ 43:164–169.

Brainerd CJ, Holliday RE, Reyna VF (2004) Behavioral measurement of remembering phenomenologies: so simple a child can do it. Child Dev 75:505–522.

Brod G, Lindenberger U, Shing YL (2017) Neural activation patterns during retrieval of schema-related memories: differences and commonalities between children and adults. Dev Sci 20:e12475.

Brunec IK, Bellana B, Ozubko JD, Man V, Robin J, Liu ZX, Grady C, Rosenbaum RS, Winocur G, Barense MD, Moscovitch M (2018) Multiple scales of representation along the hippocampal anteroposterior axis in humans. Curr Biol 28:2129–2135.e6.

Bürkner PC (2017) brms: an R package for Bayesian multilevel models using Stan. J Stat Soft 80:28.

Bürkner PC (2018) Advanced Bayesian multilevel modeling with the R package brms. The R Journal, Vol 10, pp 395-411. Available at https://journal.r-project.org/archive/2018/RJ-2018-017/index.html.

Chanales AJ, Oza A, Favila SE, Kuhl BA (2017) Overlap among spatial memories triggers repulsion of hippocampal representations. Curr Biol 27:2307–2317.e5.

Dandolo LC, Schwabe L (2018) Time-dependent memory transformation along the hippocampal anterior-posterior axis. Nat Commun 9:1205.

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

Davachi L, Danker J (2013) Cognitive neuroscience of episodic memory. In: The Oxford handbook of cognitive neuroscience, Vol 1: Core topics (Ochsner SK, ed). Oxfod: Oxford UP.

Davachi L, Mitchell JP, Wagner AD (2003) Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. Proc Natl Acad Sci USA 100:2157–2162.

DeMaster DM, Ghetti S (2013) Developmental differences in hippocampal and cortical contributions to episodic retrieval. Cortex 49:1482–1493.

DeMaster D, Pathman T, Lee JK, Ghetti S (2014) Structural development of the hippocampus and episodic memory: developmental differences along the anterior/posterior axis. Cereb Cortex 24:3036–3045.

DeMaster D, Coughlin C, Ghetti S (2016) Retrieval flexibility and reinstatement in the developing hippocampus. Hippocampus 26:492–501.

Eichenbaum H, Yonelinas AP, Ranganath C (2007) The medial temporal lobe and recognition memory. Annu Rev Neurosci 30:123–152.

Favila SE, Chanales AJ, Kuhl BA (2016) Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. Nat Commun 7:11066.

Foster JA, Burman MA (2010) Evidence for hippocampus-dependent contextual learning at postnatal day 17 in the rat. Learn Mem 17:259–266.

Furtak SC, Wei SM, Agster KL, Burwell RD (2007) Functional neuroanatomy of the parahippocampal region in the rat: the perirhinal and postrhinal cortices. Hippocampus 17:709–722.

Ghetti S, Angelini L (2008) The development of recollection and familiarity in childhood and adolescence: evidence from the dual-process signal detection model. Child Dev 79:339–358.

Ghetti S, Bunge SA (2012) Neural changes underlying the development of episodic memory during middle childhood. Dev Cogn Neurosci 2:381–395.

Ghetti S, DeMaster DM, Yonelinas AP, Bunge SA (2010) Developmental differences in medial temporal lobe function during memory encoding. J Neurosci 30:9548–9556.

Ghetti S, Mirandola C, Angelini L, Cornoldi C, Ciaramelli E (2011) Development of subjective recollection: understanding of and introspection on memory states. Child Dev 82:1954–1969.

Gogtay N, Nugent TF 3rd, Herman DH, Ordonez A, Greenstein D, Hayashi KM, Clasen L, Toga AW, Giedd JN, Rapoport JL, Thompson PM (2006) Dynamic mapping of normal human hippocampal development. Hippocampus 16:664–672.

Hayne H, Imuta K (2011) Episodic memory in 3- and 4-year-old children. Dev Psychobiol 53:317–322.

Johnson MH (2011) Interactive specialization: a domain-general framework for human functional brain development? Dev Cogn Neurosci 1:7–21.

LaRocque KF, Smith ME, Carr VA, Witthoft N, Grill-Spector K, Wagner AD (2013) Global similarity and pattern separation in the human medial temporal lobe predict subsequent memory. J Neurosci 33:5466–5474.

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

Mirandola C, Del Prete F, Ghetti S, Cornoldi C (2011) Recollection but not familiarity differentiates memory for text in students with and without learning difficulties. Learn Individ Differ 21:206–209.

Murty VP, Tompary A, Adcock RA, Davachi L (2017) Selectivity in postencoding connectivity with high-level visual cortex is associated with reward-motivated memory. J Neurosci 37:537–545.

Ofen N, Kao YC, Sokol-Hessner P, Kim H, Whitfield-Gabrieli S, Gabrieli JD (2007) Development of the declarative memory system in the human brain. Nat Neurosci 10:1198–1205.

Phelps EA (2004) Human emotion and memory: interactions of the amygdala and hippocampal complex. Curr Opin Neurobiol 14:198–202.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59:2142–2154.

Prabhakar J, Johnson EG, Nordahl CW, Ghetti S (2018) Memory-related hippocampal activation in the sleeping toddler. Proc Natl Acad Sci USA 115:6500–6505.

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

Ranganath C, Ritchey M (2012) Two cortical systems for memory-guided behaviour. Nat Rev Neurosci 13:713–726.

Revelle W (2020) psych: Procedures for Psychological, Psychometric, and Personality Research. Northwestern University, Evanston, Illinois. R package version 2.0.12. Available at https://CRAN.R-project.org/package= psych.

Riggins T, Geng F, Blankenship SL, Redcay E (2016) Hippocampal functional connectivity and episodic memory in early childhood. Dev Cogn Neurosci 19:58–69.

Ritchey M, Montchal ME, Yonelinas AP, Ranganath C (2015) Delay-dependent contributions of medial temporal lobe regions to episodic memory retrieval. Elife 4:e05025.

Sastre M 3rd, Wendelken C, Lee JK, Bunge SA, Ghetti S (2016) Age- and performance-related differences in hippocampal contributions to episodic retrieval. Dev Cogn Neurosci 19:42–50.

Schlaggar BL, Brown TT, Lugar HM, Visscher KM, Miezin FM, Petersen SE (2002) Functional neuroanatomical differences between adults and school-age children in the processing of single words. Science 296:1476–1479.

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.

Schneider W, Knopf M, Stefanek J (2002) The development of verbal memory in childhood and adolescence: findings from the Munich Longitudinal Study. J Educ Psychol 94:751–761.

Sekeres MJ, Riggs L, Decker A, de Medeiros CB, Bacopulos A, Skocic J, Szulc-Lerch K, Bouffet E, Levine B, Grady CL, Mabbott DJ, Josselyn SA, Frankland PW (2018) Impaired recent, but preserved remote, autobiographical memory in pediatric brain tumor patients. J Neurosci 38:8251–8261.

Selmeczy D, Fandakova Y, Grimm KJ, Bunge SA, Ghetti S (2019) Longitudinal trajectories of hippocampal and prefrontal contributions to episodic retrieval: effects of age and puberty. Dev Cogn Neurosci 36:100599.

Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD (2012) Decoding subject-driven cognitive states with whole-brain connectivity patterns. Cereb Cortex 22:158–165.

Spallazzi M, Dobisch L, Becke A, Berron D, Stucht D, Oeltze-Jafra S, Caffarra P, Speck O, Düzel E (2019) Hippocampal vascularization patterns: a high-resolution 7 Tesla time-of-flight magnetic resonance angiography study. Neuroimage Clin 21:101609.

Spaniol J, Davidson PS, Kim AS, Han H, Moscovitch M, Grady CL (2009) Event-related fMRI studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. Neuropsychologia 47:1765–1779.

Squire LR, Genzel L, Wixted JT, Morris RG (2015) Memory consolidation. Cold Spring Harb Perspect Biol 7:a021766.

Staresina BP, Duncan KD, Davachi L (2011) Perirhinal and parahippocampal cortices differentially contribute to later recollection of object- and scene-related event details. J Neurosci 31:8739–8747.

Tambini A, Davachi L (2013) Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. Proc Natl Acad Sci USA 110:19591–19596.

Tambini A, Davachi L (2019) Awake reactivation of prior experiences consolidates memories and biases cognition. Trends Cogn Sci 23:876–890.

Tambini A, Ketz N, Davachi L (2010) Enhanced brain correlations during rest are related to memory for recent experiences. Neuron 65:280–290.

Travaglia A, Bisaz R, Sweet ES, Blitzer RD, Alberini CM (2016) Infantile amnesia reflects a developmental critical period for hippocampal learning. Nat Neurosci 19:1225–1233.

Williams EJ (1959) Regression analysis. New York: Wiley.

Willoughby KA, Desrocher M, Levine B, Rovet JF (2012) Episodic and semantic autobiographical memory and everyday memory during late childhood and early adolescence. Front Psychol 3:53.

Yassa MA, Stark CE (2011) Pattern separation in the hippocampus. Trends Neurosci 1:515–525.