Contents lists available at ScienceDirect

Developmental Cognitive Neuroscience

Longitudinal developmental trajectories do not follow cross-sectional age associations in hippocampal subfield and memory development

Attila Keresztes [a,b,c,*], Laurel Raffington [c,d], Andrew R. Bender [c,e], Katharina Bögl [f], Christine Heim [g,h,1], Yee Lee Shing [i,1]

[a] Brain Imaging Centre, Research Centre for Natural Sciences, Eötvös Loránd Research Network (ELKH), Budapest, Hungary
[b] Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary
[c] Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany
[d] Department of Psychology, University of Texas at Austin, TX, USA
[e] Departments of Epidemiology and Biostatistics & Neurology and Ophthalmology, Michigan State University, MI, USA
[f] Faculty of Life Sciences, Berlin School of Mind and Brain and Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany
[g] Charité – Universitätsmedizin Berlin, Freie Universität Berlin and Humboldt-Universität zu Berlin, Insitute of Medical Psychology, Berlin, Germany
[h] Department of Biobehavioral Health, Pennsylvania State University, University Park, PA, USA
[i] Department of Psychology, Goethe University Frankfurt, Frankfurt am Main, Germany

ARTICLE INFO

ABSTRACT

Cross-sectional findings suggest that volumes of specific hippocampal subfields increase in middle childhood and early adolescence. In contrast, a small number of available longitudinal studies reported decreased volumes in most subfields over this age range. Further, it remains unknown whether structural changes in development are associated with corresponding gains in children's memory. Here we report cross-sectional age differences in children's hippocampal subfield volumes together with longitudinal developmental trajectories and their relationships with memory performance. In two waves, 109 participants aged 6–10 years (wave 1: M_{Age}=7.25, wave 2: M_{Age}=9.27) underwent high-resolution magnetic resonance imaging to assess hippocampal subfield volumes (imaging data available at both waves for 65 participants) and completed tasks assessing hippocampus dependent memory processes. We found that cross-sectional age-associations and longitudinal developmental trends in hippocampal subfield volumes were discrepant, both by subfields and in direction. Further, volumetric changes were largely unrelated to changes in memory, with the exception that increase in subiculum volume was associated with gains in spatial memory. Longitudinal and cross-sectional patterns of brain-cognition couplings were also discrepant. We discuss potential sources of these discrepancies. This study underscores that children's structural brain development and its relationship to cognition cannot be inferred from cross-sectional age comparisons.

1. Introduction

A prominent branch of the quest to advance our understanding of children's memory development attempts to identify maturational trajectories of the hippocampal network (Lee et al., 2017). This interest is fueled by the hippocampus' key functions in both learning and remembering (Scoville and Milner, 1957) that render it a brain region of particular interest to research on individual development and school-based education. Recently developed high-resolution magnetic resonance imaging (MRI) techniques provide researchers with new tools

to delineate hippocampal subfields in vivo (Bakker et al., 2008; Eldridge et al., 2005). Volumetric measures of hippocampal subfields obtained using these techniques provide the current best in vivo proxy for assessing maturity of intrahippocampal networks in children (Keresztes et al., 2018). However, developmental trends inferred from cross-sectional comparisons are not always supported – and rather are often contradicted – by data from longitudinal studies. In this study, we examine hippocampal subfield development and its relationship with children's memory performance by assessing both cross-sectional age-differences and longitudinal change in hippocampal subfield volumes

---

and memory. To foreshadow our conclusions, the data presented underscores the necessity to refrain from making inferences on developmental trajectories based on cross-sectional data.

1.1. Developmental trends inferred from cross-sectional comparisons versus longitudinal studies

In line with histological studies of monkeys (Lavenex and Lavenex, 2013), high-resolution structural MRI investigations of children have observed age-related volumetric differences in the hippocampus (for reviews, see Keresztes et al., 2018; Lee et al., 2017) suggesting that hippocampal development extends at least into early adolescence, and potentially into adulthood. Importantly, this may demonstrate a more protracted development than previous human data had indicated (Frotscher and Seress, 2007; Utsunomiya et al., 1999). This line of research also provided evidence for heterogeneous developmental trends across regions within the hippocampus, previously observed using longitudinal standard resolution MRI (Gogtay et al., 2006). Recent high-resolution MRI studies with pediatric samples (Bouyeure et al., 2021; Canada et al., 2019; Daugherty et al., 2017; Keresztes et al., 2020, 2017; Krogsrud et al., 2014; Lee et al., 2014; Riggins et al., 2018; Schlichting et al., 2017; see Table S1) consistently reported volumetric age-differences in the dentate gyrus (DG), cornu ammoni (CA) regions, and subiculum (SUB), but not in the entorhinal cortex (EC). Due to differences in sample size, age range, the longitudinal extent of the hippocampus assessed, as well as the subregions delineated (see Table S1), comparing the available findings is not straightforward. That said significant age–volume associations reported in most studies for regions comprising DG, CA1 to CA4, and SUB were either positive linear or quadratic, with quadratic associations reflecting a positive association for younger children and no association or a negative association for older children (see Table S1).

Two studies (Daugherty et al., 2017; Schlichting et al., 2017) diverged from this overall pattern of results. Daugherty et al. (2017) found a quadratic association for a region comprising CA1 and 2 – with younger children showing a negative association and older children showing a positive association with age – and a negative linear association for a region comprising CA3 and DG. One critical difference between this and all other studies that may drive the divergence of findings is the extent of the hippocampal longitudinal axis assessed. Compared to other studies assessing hippocampal subfield volumes along the full length of the hippocampal body, or head and body, or the full hippocampus, this study only used the three most anterior slices of the hippocampal body to assess hippocampal subfield volumes. Schlichting et al. (2017) reported a negative linear age–volume association for SUB in the hippocampal body, and a negative linear association for CA1 in the head. Although these findings were inconsistent with the pattern of findings observed across other studies, consistently with those, Schlichting and colleagues also found a positive linear association for DG. In this case, the divergent findings may be due to a multitude of differences in methods compared to other studies.

To our knowledge, three previous studies have examined longitudinal trajectories of subfield volumes in children (Canada et al., 2021; Tamnes et al., 2018, 2014, see also Table S1). Critically, their findings stand in stark contrast to cross-sectional results. One early study (Tamnes et al., 2014) found volumetric decrease in all subfields measured within the hippocampus proper, except in the right CA1 from 8 to 21 years of age. Another study from the same lab (Tamnes et al., 2018), but with an independent sample of participants aged 8–26 years, found linear decreases in most subfields, except for CA1 with an increase until 20 followed by a decrease, as well as an increase until age 13–15 followed by decelerated decrease in the SUB. The third study (Canada et al., 2021) found volumetric increases in all subfields, but only in specific age windows and hippocampal sections (body or head) for specific subfields: DG2–4/CA3 and SUB in the hippocampal body showing an increase between 5 and 6 years of age, and CA1 in the

hippocampal head showing an increase between 4 and 5 years of age, within the 4–8 years age-window covered in their study sample.

Sources of discrepancies across and among cross-sectional and longitudinal findings may involve differences in methods of subfield delineation (Yushkevich et al., 2015a) in scanning parameters, in age ranges, in intervals between waves in case of longitudinal studies (Canada et al., 2019), and in sample size. These potential sources of differences aside, however, theoretical accounts, simulations, and comparisons of cross-sectional and longitudinal analyses of the same data caution strongly against making inferences about change based on cross-sectional studies (Bender and Raz, 2015; Lindenberger et al., 2011; Louis et al., 1986; Nyberg et al., 2010; Pfefferbaum and Sullivan, 2015). For instance, Louis et al. (1986) have mathematically shown that linearly modeled cross-sectional and longitudinal slopes from the same sample will only agree under the assumption that age distributions in cross-sectional samples are Gaussian and that change is linear or quadratic nonlinear. However, real life samples may rarely behave this way and can change in various non-linear ways (Ghisletta et al., 2020; Raz and Lindenberger, 2011). One striking example of discrepancy between cross-sectional age difference and longitudinal slope was identified by Nyberg et al. (2010). They showed that change across six years in frontal recruitment in healthy aging was negative which contradicted a plethora of cross–sectional evidence that had suggested increased frontal recruitment in older than in younger adults. Similarly, a recent study of hippocampal volume in a lifespan sample found no cross-sectional age-differences while longitudinal analyses of the same sample detected change (Pfefferbaum and Sullivan, 2015).

1.2. Volume-memory associations across childhood development

An early meta-analysis of 33 studies on the association between hippocampal volume and memory (Van Petten, 2004) favored a "smaller is better" rather than a "bigger is better" hypothesis in childhood, adolescence, and young adulthood. More recent studies that tested for associations between memory and hippocampal subfield volumes in children using high-resolution MRI (summarized in Table S1) paint a less straightforward picture. These studies detected associations between hippocampal subfield volumes and memory discrimination (Bouyeure et al., 2021; Canada et al., 2019; Keresztes et al., 2017; Lee et al., 2014), associative memory (Daugherty et al., 2017; Lee et al., 2014), source memory (Canada et al., 2021; Riggins et al., 2018), statistical learning and associative inference (Schlichting et al., 2017), as well as learning and delayed free recall (Tamnes et al., 2014). Both positive and negative associations were found, and in cross-sectional studies, age often moderated the direction of the association (Bouyeure et al., 2021; Canada et al., 2019; Riggins et al., 2018; Schlichting et al., 2017). Although such shifts in the direction of volume-memory associations may be partly driven by measurement variance across the tested age-window, they do point out that volume-memory associations may be the outcome of diverse underlying mechanisms potentially changing along different trajectories across childhood. This notion additionally underlies the need for longitudinal tests of volume-memory associations. In particular, cross-sectional studies do not provide information about heterogeneous trajectories of change (Lindenberger et al., 2011) and thus are unable to inform research on how change in brain is associated with change in cognition. Thus, to understand how hippocampal subfields development is coupled with memory development, longitudinal studies are needed.

To our knowledge, longitudinal data on hippocampal subfield volume–memory associations is only reported in two studies (see Table S1), and of these only one tested memory constructs that had been tested in any extant cross-sectional study. In a longitudinal study, Canada et al. (2021) found associations between change in CA1 and SUB and change in source memory, whereas in an earlier cross-sectional analysis of the same sample, Riggins et al. (2018) found associations between volumes of CA1 and CA2–4/DG and performance on the same source memory

task. In brief, cross-sectional data on associations between memory and hippocampal subfield volumes is sparse, and longitudinal data – necessary for inferences on brain-cognition couplings – is very limited.

Cross-sectional and longitudinal results are rarely compared within the same dataset also in the broader field of cognitive neuroscience (for examples that did such comparison, see Nyberg et al., 2010; Pfefferbaum and Sullivan, 2015; Raffington et al., 2019). One potential reason for this may be pressure on PhD students, postdocs and senior researchers who – due to limited term funding – cannot wait with publication of cross-sectional results until a longitudinal study is concluded. Thus, results of initial waves in a longitudinal study are usually already published when longitudinal results are written up. That said, a recap of cross-sectional results and their comparison to longitudinal results in all manuscripts publishing longitudinal data could prove critical in highlighting what cross-sectional data can tell researchers – if anything – about developmental mechanisms.

In sum, the current study aims at providing further longitudinal evidence on hippocampal subfield development as well as its associations to memory development. In addition, it aims to assess consistency between developmental trends inferred from cross-sectional and longitudinal data within the same sample.

1.3. The current study

Here we examine cross-sectional age-differences and longitudinal change in hippocampal subfield volumes as well as their relationship with 6–10-year-old children's memory performance. We assessed age-differences and change across two years in aggregated hippocampal subfield volumes, in memory measures assessing function of the hippocampus as a whole or of its specific subfields, as well as in hippocampal subfield volume–memory associations across time. Based on cross-sectional age–volume associations, we hypothesized that DG-CA3, CA1–2, and potentially SUB would exhibit longitudinal volumetric increases, but that we would observe no increase in EC volume. Further hypotheses were based on prior cross-sectional findings on age-volume-memory associations (reviewed in Keresztes et al., 2018; for a newer study see Canada et al., 2019). Based on the role of DG-CA3 in pattern separation on highly similar inputs to the hippocampus (Yassa and Stark, 2011), we hypothesized that volumetric change in DG-CA3 will be specifically associated with change in the behavioral ability to discriminate highly similar memories, i.e., with mnemonic discrimination (Bakker et al., 2008; Berron et al., 2016). Findings by Daugherty et al. (2017) additionally raised the possibility that change in DG-CA3 is potentially associated with change in associative memory. Finally, we expected that change in spatial memory, which seems to rely on overall function of the hippocampal circuitry (Burgess et al., 2002; Li and King, 2019) would be associated with change in total hippocampal volume, but not with change in specific subfield volumes.

2. Methods

2.1. Participants

Participants included 147 children (6.08–7.98 years, M_{age} = 7.19, SD_{age} = 0.46, 67 girls) from socioeconomically diverse families (parental income similar to city average; parental education and employment above city average) in Berlin, Germany who were enrolled in a longitudinal study (for details of the recruitment procedure, socioeconomic status, and additional demographics, see Raffington et al., 2018). Of them, 127 returned at wave 2, approximately two years later (8.3–10.15 years, M_{age} = 9.25, SD_{age} = 0.45, 59 girls), also described in detail elsewhere (Raffington et al., 2019). A subsample of children participated in one magnetic resonance imaging (MRI) session at both waves within three weeks of the behavioral sessions. In this study we included all children who had a high-resolution structural scan in at least one of the waves. The final sample therefore consisted of 109 children.

Hippocampal subfield data available at both waves was lower (n = 65, 24 girls; 6.08–8.00 years, M_{age} = 7.26, SD_{age} = 0.42 at wave 1, and 8.38–10.12 years, M_{age} = 9.27, SD_{age} = 0.42 at wave 2). Completeness of data for each variable, sex and age for complete cases, as well as causes of exclusion per variable, are reported in the supplement (Table S2). Informed consent was obtained from legal guardians of all participants: Parents provided written informed consent and children provided verbal assent. The study was approved by the 'Deutsche Gesellschaft für Psychologie' ethics committee (YLS_012015). Limited – mostly cross-sectional – results from wave 1 of this study have already been published (Keresztes et al., 2020; Raffington et al., 2020, 2019, 2018).

2.2. Procedure

At wave 1, all participants were invited for two sessions, and a subsample of participants was also invited for a third session. Each session lasted approximately two hours, and included various behavioral tests, physiological measurements, and questionnaires. Session 1 and 2 took place on consecutive days. Session 3 followed within three weeks of session 2 and included a ~35-minute-long MRI session. Of interest to the current study, participants completed an associative memory task followed by a spatial memory task at the beginning of session 1 and underwent structural MRI and completed a mnemonic similarity task at the end of session 3. At wave 2, all participants returning from wave 1 were invited to participate in two sessions, each two-hours-long. Session 1 comprised of behavioral tests and two ~ 20-minute-long MRI sessions separated by a break. Session 2, which followed session 1 within three weeks, comprised of only behavioral tests. Of interest, the associative memory, the spatial memory, and the mnemonic similarity tasks were all performed at session 2.

2.3. Volumetric measurement of hippocampal subfields

Acquisition of high-resolution MRI images of the hippocampus and surrounding mediotemporal areas, and the full process of subfields delineation on MRI images was identical across waves and is described in detail in Keresztes et al. (2020). In brief, we acquired partial field of view (FoV) proton density (PD)-weighted T2 images (FoV: 206 mm; repetition time (TR): 6500 ms; echo time (TE): 16 ms; number of slices: 30; voxel size: 0.4 mm × 0.4 mm × 2.0 mm) perpendicular to the longitudinal axis of the right hippocampus on a 3 T Siemens Magnetom TrioTim syngo MRI scanner. Subfield segmentation utilized the Automated Segmentation of Hippocampal Subfields (ASHS) (Yushkevich et al., 2015b) software using a customized atlas generated in ASHS from highly reliable manual segmentations measured in earlier studies (Bender et al., 2018). This approach has been shown to be highly reliable and valid in identifying hippocampal subfield boundaries in a lifespan sample, including 20 6–24 year-old and 20 62–79 year-old participants (Bender et al., 2018). Moreover, Homayouni et al. (2021) have shown that hippocampal subfield volumes delineated using the manual segmentation protocol described in Bender et al. (2018) had excellent test-retest reliability and the method had high sensitivity to detect two-year change in a longitudinal sample (n = 28) of children aged 7–17 years. The customized atlas included manual demarcations of three regions within the hippocampal body, bilaterally – SUB, a region including CA regions 1 and 2 (CA1–2), and a region including DG and CA3 (DG-CA3) – as well as the EC on six consecutive slices anterior to the hippocampal body (see Fig. 1 of Bender et al., 2018 for the illustration of the heuristic rules used for manual demarcation in the atlas sample). Importantly, this atlas includes additional demarcations in both anterior and posterior directions beyond the above described manually defined ranges. This extension of the atlas was necessary to ensure that ASHS always returns fully segmented output even for the most anterior and most posterior slices within the manually defined ranges (Bender et al., 2018).

For the current study, after delineating subfields using ASHS with the
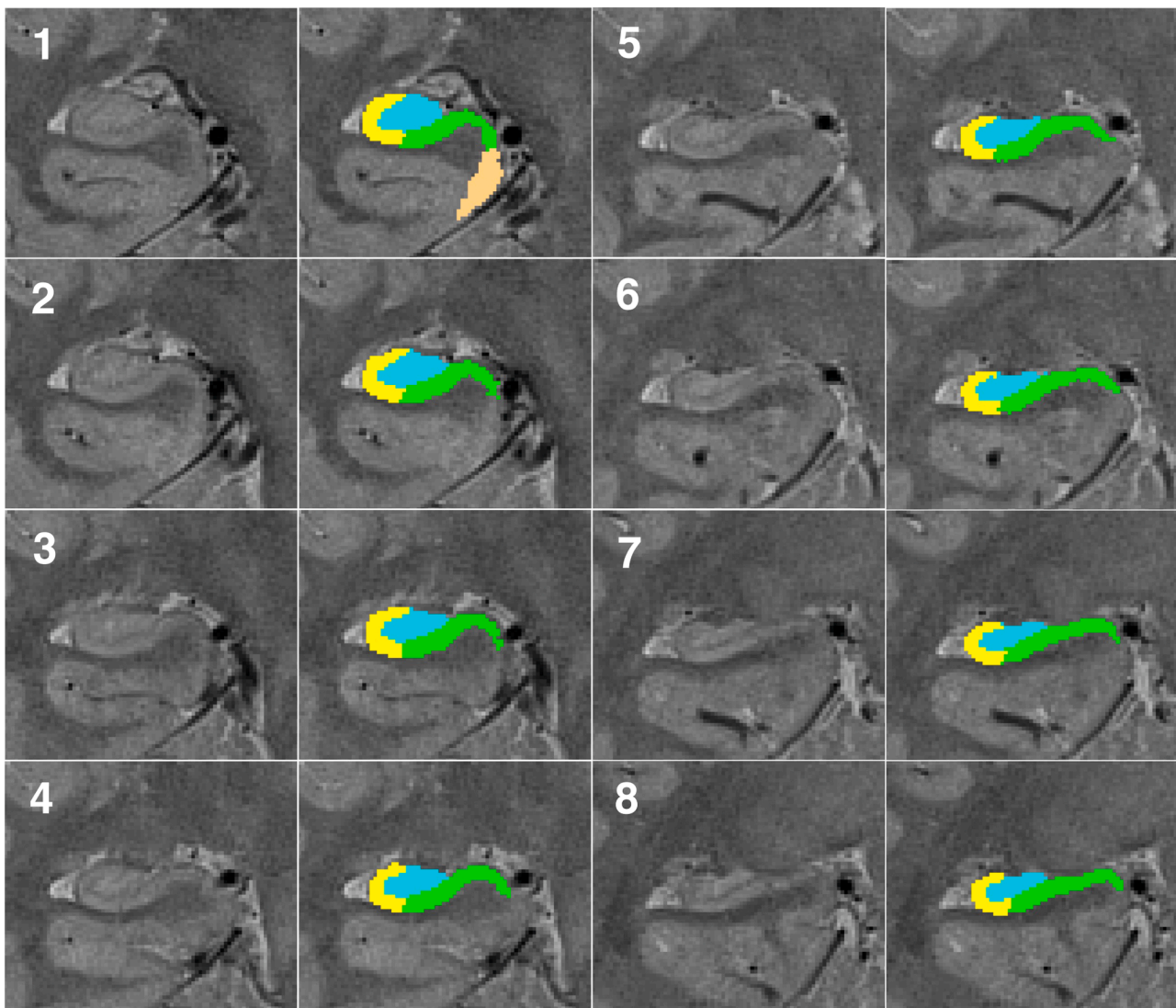
Fig. 1. Illustration of hippocampal subfields delineated. Subfields subiculum (in green), CA1–2, (in yellow), and DG-CA3 (in blue) were segmented in the hippocampal body shown here in order from most anterior slice (1) to the most posterior slice (8) in the right medial temporal lobe, with PD-T2 images unlabeled on the left and labeled on the right. Entorhinal cortex (in copper) was segmented on the most anterior slice of the body as well as 5 more consecutive slices (not shown here). (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

custom atlas by Bender et al. (2018), we used a custom shell script (Bender et al., 2018) to truncate CA1–2, DG-CA3, and SUB outputs to include only the delineations within a manually defined range of the hippocampal body, and to truncate EC outputs to the six consecutive slices anterior to the body (see Fig. 1 in Keresztes et al., 2020). Before manually defining hippocampal body ranges, to establish reliability of manually identified boundaries of the hippocampal body, first at wave 1, K.B. and A.K. both determined the starting and end slices, for left and right hemispheres, in a subset (n = 48) of the sample (for all, Cohen's kappa >0.82), then K.B. ranged the rest of the sample (for more details on this step, see Keresztes et al., 2020). Next, before determining hippocampal body boundaries in wave 2, K.B. established reliability of boundaries determined by herself, by determining boundaries again on a subset (n = 13) of the sample from wave 1 (Cohen's kappas >0.76 for left and right starting and end slices). Then, K.B. determined hippocampal body boundaries on the whole sample of wave 2. Final outputs were visually inspected by K.B. and A.K. and obvious errors were corrected (e.g., voxels labeled by ASHS as subfield outside the hippocampus were cleared of their label, and voxels not labeled by ASHS as subfield within a continuous cluster of subfield voxels were labeled as subfield.). In addition, for 3 cases in wave 1, errors in subfields delineated by ASHS

called for manual delineation of all subfields. A.K. performed these delineations following rules used for manual delineation of subfields for all atlas cases used, detailed in Bender and colleagues (Bender et al., 2018). An example delineation for the hippocampal body is shown in Fig. 1. Volumetric measures for all regions were adjusted for intracranial volume based on (Jack et al., 1989; Keresztes et al., 2017; Raz et al., 2005). In brief, first, we obtained ICV estimates using the brain extraction tool in FSL 5.0 (Smith, 2002) using procedures described in Bender et al. (2013). Then we regressed ICV on age to obtain an individual ICV estimate predicted by age only and adjusted each subfield volume by correcting for differences in subfield volumes due to differences in actual and predicted ICV (adjusted volume = raw volume – ß (ICV – predicted ICV). The resulting adjusted volumetric measures were summed across hemispheres and used for all analyses presented in this study. Note that running all analyses reported in this study using raw volumetric values yielded an identical pattern of results to those reported with ICV corrected values.

2.4. Memory measures

We administered three memory tasks to assess hippocampal

contributions to memory: a mnemonic similarity task, a spatial memory task, and an associative memory task. All three were computerized tasks administered on desktop computers with 21.5-inch screens.

Based on established methods (Kirwan and Stark, 2007) the mnemonic similarity task used a continuous recognition memory design to assess participants' ability to discriminate memories of highly similar stimuli. Participants saw 162 pictures depicting everyday objects, presented sequentially. Critically, 48 pictures were repeated after a delay of 2–14 intervening trials. Twenty-four of the pictures were repeated exactly, whereas 24 pictures were repeated with a slightly different lure picture of an identical object. For each trial the child was instructed to identify pictures as "old" (exact repetition), "similar" (lure repetitions), or "new" (new items). We calculated a lure discrimination index (i.e., a measure of participants' ability to discriminate between highly similar memories) as the proportion of "similar" responses to lure repetitions minus the proportion of "similar" responses to new items. More details on this task are provided in recent work (Keresztes et al., 2020).

The spatial memory task was based on methods originally reported by Kessels et al. (2007), and assessed memory for items and item-location associations. Briefly, participants encoded locations of 15 sequentially presented line drawings of everyday objects on a 6 × 6 grid. After a short delay, they performed a recognition task using the same pictures of studied objects randomly intermixed with 15 pictures of new objects. For each correctly recognized object, we asked them to point to the location of the given picture in the grid during encoding. As a measure of spatial memory, we used the percentage of correctly indicated locations for the 15 old items. For a detailed description of the task see Raffington et al. (2018).

Participants also completed a paired-associate memory task, with an incidental encoding phase followed by two retrieval phases, including one at a short (< 2 mins) delay and one after approximately 24 h, to assess associative memory performance across different time spans. At encoding, 34 pairs of German nouns (for the list of words see Table S3 in the Appendix) were presented to participants sequentially. Each word pair was presented simultaneously both in visual (on the computer screen) and auditory modalities (via loudspeaker). In each trial, following a fixation cross (500 ms), a cue word appeared in the middle of the screen for 2 s, then the target word appeared on the screen below the cue word; both cue and target words remained on screen for an additional 4 s. Participants then saw a question mark on the screen for 15 s, and their task was to decide whether the two words are related to each other or not. The two retrieval phases were identical, both came as a surprise to participants, and both included different halves of the word pairs encoded. The retrieval phase consisted of a cued recall block followed by a stem-cued recall block. Trials in cued recall consisted of a cue presented both visually and auditorily, with the visually presented cue remaining on screen for 15 s or until the participant pressed a response button on a keyboard. Then the experimenter asked the participant to report the answer verbally. The experimenter typed in the answer via the same keyboard and then advanced the experiment to the next trial. Trials in the stem cued recall block were identical to trials in the cued recall block but the cue was appended with the first two phoneme of the target (e.g. 'Gi' for Giraffe, and 'Kr' for "Kreide"/chalk). Because cued recall performance was close to floor on the second retrieval test performed a day after encoding, and because stem-cued recall is known to be less dependent on hippocampal processing, in this study we only used cued recall performance on the first retrieval test. As a measure of associative memory, we used percentage of targets correctly recalled.

2.5. Cross-sectional and longitudinal assessment

To assess cross-sectional age-differences, we linearly regressed variables of interest on age. We assessed bivariate associations between hippocampal subfield volumes and memory measures cross-sectionally using Pearson's zero-order correlations. As a measure of stability, we also report Pearson's zero order correlations between wave 1 and wave 2

data for all measures of interest. To assess longitudinal change across two waves, we applied latent change score (LCS) structural equation models (Kievit et al., 2018; McArdle and Nesselroade, 1994), using the lavaan package (Rosseel, 2012) in R (R Core Team, 2019). The LCS models produced three parameter estimates of particular interest: (1) mean change from wave 1 to wave 2, (2) variance in change, and (3) the covariance between the intercept (wave 1 values) and change. After fitting univariate LCS models to assess change in each of the variables (see Fig. 2a), we employed a bivariate LCS approach (Fig. 4a) to assess pairwise change–change associations between variables. Age at wave 1 was included as a covariate in all models, covarying with change, and predicting wave 1 intercept. For estimation of LCS model parameters, hippocampal subfield volume values were divided by 100 to match the scale of other variables in the analyses.

Volumetric measures acquired using MRI are not free of measurement error (Karch et al., 2019; Maclaren et al., 2014; Madan and Kensinger, 2017), therefore in addition to LCS models with single bilateral volumetric indicators, we also tested univariate LCS models of change in hippocampal subfields where left *and* right hemispheric volumes served as dual observed indicators for the volumetric latent variables at each wave (Figs. 2b and 4b). These models thus separate error variance from construct variance and establish measurement invariance over time. We used these models – henceforth referred to as bilateral indicator univariate LCS models (Kievit et al., 2018) of the hippocampal subfield volumes – to replicate findings of the univariate LCS models.

All models were computed using maximum likelihood estimation implemented in lavaan. To evaluate model fit, we used standard goodness-of-fit indices: root mean square error of approximation (RMSEA) and comparative fit index (CFI). Models were considered a good fit with RMSEA < 0.08 and CFI > 0.95 (Kline, 2015). The difference in χ^2 fit statistics was used to compare nested models, with the degrees of freedom being the difference in the number of free parameters. The threshold for statistical significance was *alpha* = 0.05. In addition, we calculated 95% confidence intervals for parameter estimates using 1000 bootstrapped resamples.

3. Results

3.1. Age-differences in hippocampal subfield volumes were stable across time points but did not match longitudinal change in hippocampal subfield volumes over time

Wave 1 and wave 2 vol estimates strongly correlated for all four subfields (*rs* between.76 and.82, all *ps* < 0.001). The pattern of cross-sectional associations in wave 2 was similar to that observed in wave 1 (Keresztes et al., 2020, see also Fig. S1). CA1–2, DG-CA3, and total HC volume correlated positively, and significantly with age, whereas neither SUB nor EC showed significant associations with age. However, these patterns were at odds with change observed within individuals. Latent change score models (see Fig. 2a for a graphical depiction of univariate LCS model specifications; and Table 1 for model fit and parameter estimates) indicated significant, positive mean volumetric changes for SUB and EC, as well as non-significant negative mean volumetric change in CA1–2, and non-significant positive mean change in DG-CA3 and in HC. Confidence intervals calculated from bootstrapped samples (see Table S5), provided support for the robustness of the significant associations. Importantly, bilateral indicator univariate LCS models of the hippocampal subfield volumes (see Fig. 2b and Table S4) also replicated these results. Discrepancies between cross-sectional and longitudinal slope estimates are visualized in Fig. 3.

We observed significant variance in volumetric change in all hippocampal subfields (see Table 1 and Fig. 3). As shown in Fig. 3, the number of individuals showing volumetric increase (versus decrease) differed across subfields, with 27 (42%), 34 (52%), 46 (71%), and 42 (65%) of 65 participants showing volumetric increase in CA1–2, DG-CA3, SUB, and EC, respectively. Volumes at wave 1 were also
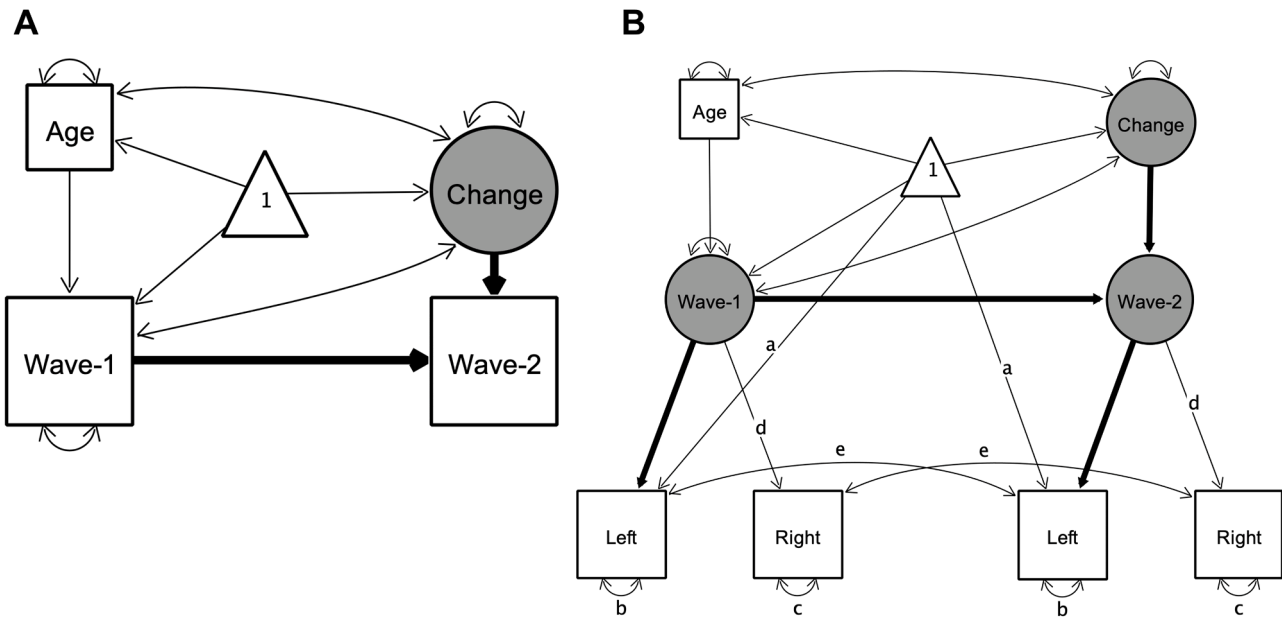
Fig. 2. Illustration of univariate latent change score models used to investigate change in variables of interest. Rectangles, circles, and triangles represent indicator variables, latent variables, and means, respectively. Estimated variances, covariances and regression paths are shown as thinner solid lines. Thicker solid lines represent path values fixed at 1. Age represents age at wave 1. In (A) used to model change in all variables of interest, wave-1 and wave-2 represent observed values of a given variable at each wave. In bilateral indicator univariate LCS models (B), used as an additional model of change in hippocampal subfield volumes, latent wave 1 and wave 2 estimates are modeled by observed left and right hemispheric volumetric measures. Identical path names denoted by (a–e) represent estimated parameters that are constrained to be equal to each other to ensure measurement invariance (Raz et al., 2005). Difference in χ^2 fit statistics indicated that lifting these constraints did not improve model fit, except for EC. For EC letting indicator variances of left (b) and right (c) differ across waves improved model fit, therefore we lifted these constraints in that model (see Bender and Raz, 2015). Parameter estimates of interest are shown separately in Table 1 for model (A) and Table S4 for model (B) for better readability of the figure.

Figures created by Onyx, version 1.0–1010 (von Oertzen et al., 2015).

significantly negatively associated with volumetric change for all subfields, except for EC. Removing age as a covariate from LCS models did not change the observed pattern of results.

3.2. Cross-sectional age-differences in memory measures were stable across time points and agreed with change in memory measures over time

Wave 1 and wave 2 performance moderately correlated for all three memory measures (*r*s between.36 and.45, *p*s < .005). Except for spatial memory at wave 1, memory measures and age were positively associated at both waves, although most of these associations did not significantly differ from zero (Fig. 4). LCS models on memory measures showed a significant linear mean change over time for all three memory measures (see Table 1 and Fig. 4). As shown in Fig. 4, we observed significant variance in change across the three tasks, with 61%, 74%, and 71% of participants showing an increase in spatial memory, lure discrimination and associative memory respectively. Wave 1 performance on all three tasks was negatively associated with change in performance across the two waves. Again, removing age as a covariate from LCS models did not change the observed pattern of results.

3.3. Consistency of cross-sectional associations between hippocampal subfields and memory measures at wave 1 and wave 2

Zero-order correlations between volumetric measures of hippocampal subfields and memory measures (Fig. S1) showed that SUB was the only hippocampal subfield significantly associated with memory at either wave. In wave 1, SUB correlated positively and significantly with associative memory, and in wave 2 it significantly correlated with all three memory measures. Thus, only one association – between associative memory and SUB volume – was consistent in both waves. In the model of total HC, we observed significant positive associations between

HC and lure discrimination and spatial memory at wave 2, as well as a positive trend at both waves between HC and associative memory.

3.4. Longitudinal associations between hippocampal subfields and memory measures

Next, we tested for potential longitudinal associations between performance on memory measures and volumetric measures in hippocampal subfields using bivariate LCS models. Altogether, we ran 15 LCS models, 4 subfield volumetric measures plus volume of total hippocampal body × 3 memory measures, and extracted parameter estimates for change–change covariances as well as covariances between memory measures at wave 1 and change in hippocampal subfield volumes, and between hippocampal subfield volumes at wave 1 and change in memory (see Fig. 5a for bivariate LCS model specification, and Table 2 for model fit and parameter estimates).

The only significant change-change association was a positive one between change in SUB volume and change in spatial memory performance (*p* = .036). However, a confidence interval calculated from bootstrapped samples [− 0.001,0.065] did not provide support for the robustness of this association. In addition, we observed a significant positive association between performance on the spatial memory task at wave 1 and change in EC (*p* = .029). A confidence interval from bootstrapped samples [0.001,0.03] provided additional support for the robustness of this association. No other longitudinal parameter estimates of interest were significant. Importantly, bivariate LCS models including bilateral indicator univariate LCS models for the hippocampal subfield volumes (see Fig. 5b), replicated the observed pattern of results (see Table S6), that was also unaltered when age was removed as a covariate from the models.
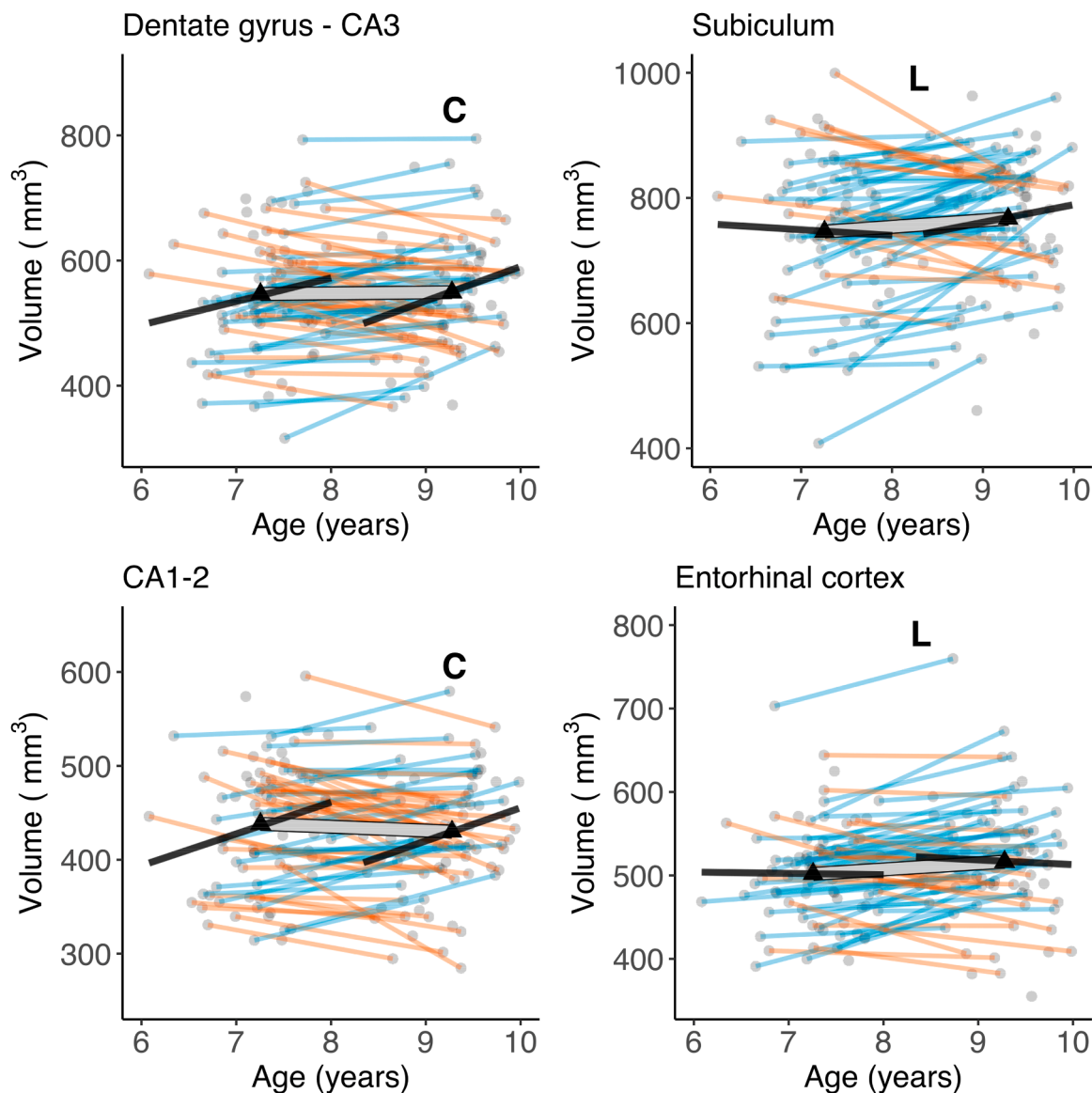
Fig. 3. Cross-sectional age-differences and longitudinal age-trends in hippocampal subfield volumes. Dots represent individuals, triangles represent mean volume at each wave. Black lines show regressions of volume on age cross-sectionally at each wave. The thick transparent white line represents mean change, thinner colored lines represent individual change with orange and blue lines showing a decrease or increase in volume, respectively. Bold letters 'C' and 'L' represent significant cross-sectional age-difference at a given wave, and significant mean change across waves, respectively. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

3.5. Sex effects

To control for effects of sex, we tested for sex differences as well as for age × sex interactions in all variables of interest, using independent samples t-tests and multiple linear regressions, respectively. We found a significant difference in EC volume – with girls having larger volumes – in both wave 1 ($t(82) = 2.1$, $p = .04$), and wave 2 ($t(83) = 4.06$, $p <$ [TS82 0.001 other sex effects were significant, nor any sex × age interactions. Including Sex in LCS models of EC did not change the pattern of results reported in 3.1 and 3.4.

3.6. Additional control and power analyses

We performed additional independent sample t-tests to rule out the possibility that selective attrition between waves may underlie, at least in part, the discrepancies between estimates of cross-sectional age-differences and longitudinal change. For instance, if dropouts have lower or higher values compared to ongoing participants, we could observe

longitudinal changes even if there is no actual change or miss out on detecting longitudinal change. Because the sample also included participants who participated in both waves but were scanned only at wave 2, we defined dropouts for analyses both forward (wave 1 data present and wave 2 data missing) and backward (wave 2 data present and wave 1 data missing) in time. For all volume and memory variables, for both forward and backward dropouts, we performed Welch t-tests comparing dropouts with non-dropouts. Neither of these analyses yielded significant effects (all $ts < .97$, all $ps > .34$), suggesting that potential differences between dropouts and non-dropouts were not affecting longitudinal estimates of change in any variables of interest.

In addition, we also tested whether dropouts differed from non-dropouts on sex or age. Forward dropouts were significantly older than non-dropouts, $t(8.7) = 4.00$, $p = .003$. We found no other age or sex differences between dropouts and non-dropouts (See section S1.1 of the Supplement for the same analyses performed separately for each variable of interest separately).

To rule out a limit on power to detect otherwise existing associations
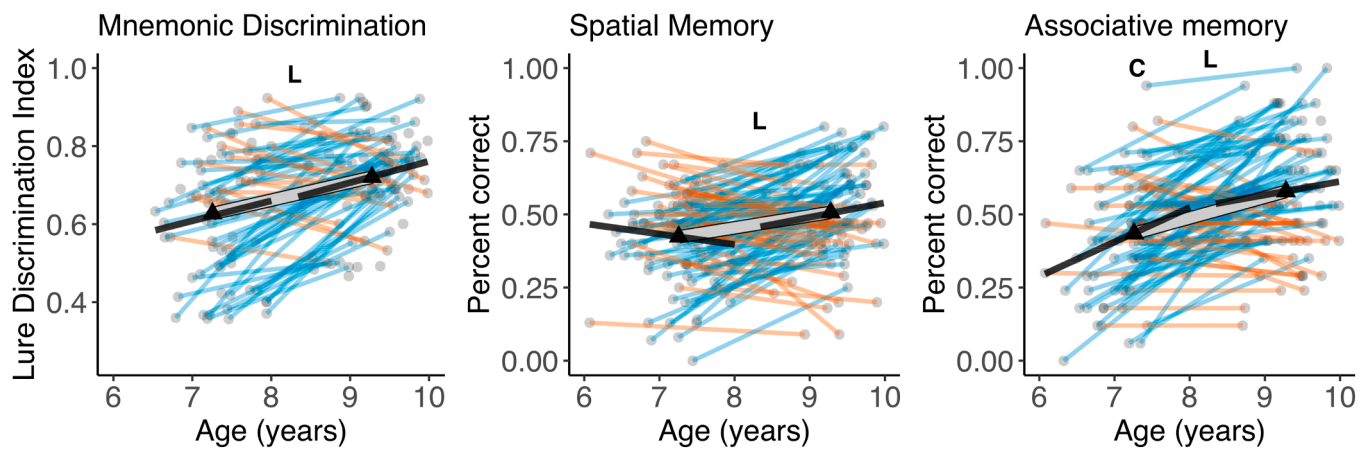
Fig. 4. Cross-sectional age-differences and longitudinal age-trends in memory measures. Dots represent individuals, triangles represent mean volume at each wave. Black lines show regressions of volume on age cross-sectionally at each wave. The thick transparent white line represents mean change, thinner colored lines represent individual change with red and green lines showing a decrease or increase in volume, respectively. Bold letters 'C' and 'L' represent significant cross-sectional age-difference at a given wave, and significant mean change across waves, respectively. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Table 1
Key parameter estimates in univariate latent change score models for all variables of interest.

	Parameter estimates							
	M_{change}		Var_{change}		$\beta_{Age\text{-}at\text{-}wave\ 1 \rightarrow wave\ 1}$		$Cov_{change\text{--}wave\ 1}$	
	PE (SE)	$\Delta\chi^2$ (1)	PE (SE)	Wald statistic[a]	PE (SE)	$\Delta\chi^2$ (1)	PE (SE)	$\Delta\chi^2$ (1)
Hippocampal subfields								
CA1–2	-0.08 (0.05)	2.45	0.172 (0.031)	5.6 * **	0.307 (0.153)	3.94 *	-0.094 (0.035)	9.20 * *
DG-CA3	0.015 (0.065)	0.06	0.288 (0.051)	5.59 * **	0.294 (0.222)	1.74	-0.204 (0.066)	12.70 * **
SUB	0.25 (0.088)	7.66 * *	0.533 (0.094)	5.67 * **	-0.01 (0.277)	0.001	-0.417 (0.109)	22.22 * ***
EC	0.143 (0.047)	8.55 * *	0.154 (0.027)	5.73 * **	-0.007 (0.14)	0.003	-0.033 (0.027)	1.55
Total HC	0.192 (0.189)	1.02	2.5 (0.451)	5.55 * **	0.577 (0.572)	1.01	-1.892 (0.498)	21.63 * **
Memory measures								
LDI	0.101 (0.019)	24.11 * ***	0.023 (0.004)	5.78 * **	0.044 (0.050)	0.76	-0.018 (0.003)	73.23 * ***
SM	0.084 (0.017)	23.02 * ***	0.029 (0.004)	7.23 * **	-0.036 (0.033)	1.17	-0.013 (0.003)	8.375 * *
Cued recall	0.139 (0.019)	43.36 * ***	0.036 (0.005)	7.03 * **	0.116 (0.037)	9.35 * *	-0.013 (0.003)	22.67 * ***

Note. M: mean, Var: variance, PE (SE): parameter estimate (standard error), DG: dentate gyrus, SUB: subiculum, EC: entorhinal cortex, HC: hippocampus. SM: Spatial memory, LDI: Lure discrimination index. Model fit was perfect for all models, $\chi^2 = 0$, RMSEA= 0, CFI= 1. Estimates for error variances, as well as for indicator variable means are not presented. `: $p < .1$, * :$p < .05$, * *:$p < .01$, * ** :$p < .001$, uncorrected for multiple comparisons. Confidence intervals calculated from bootstrapped samples provided support for the robustness of all associations significant at $p < .05$. [a]: For Var_{change} estimates, the difference in χ^2 could not be obtained because the models did not converge without this parameter. Therefore, the Wald statistic was used instead. Units are $mm^3/100$ for hippocampal subfields and % for memory measures.

in our relatively small sample, we performed post-hoc power analyses for univariate LCS models using the RAMpath R package (Zhang et al., 2015; Zhang and Liu, 2018), based on the original RAMpath software (Boker et al., 2002), relying on lavaan (Rosseel, 2012). This analysis indicated that LCSs in our longitudinal sample, given an $\alpha = 0.05$, had a power of.99 to detect a mean volumetric change in DG-CA3 and CA1–2; this was equivalent to the smallest cross-sectional slope for the same regions across the two waves (slope of CA1–2 on Age at wave 1 = 0.34). Because longitudinal slopes may indeed be smaller than cross-sectional slopes, we drew power curves plotting power against sample size as a function of longitudinal slope estimates. This analysis (see Fig. S2) indicated that our LCS model's power decreased below the conventionally accepted.8 value when the estimated slope was below.19, meaning that we may have missed existing effects of change if these were below an annual change of 19 mm^3.

4. Discussion

The two most important highlights from our study are (1) we found longitudinal changes that had not been predicted based on cross-sectional data, and (2) most of our hypotheses based on cross-

sectional age associations were not supported by longitudinal data. Crucially, for hippocampal subfields, we found a cross-sectional trend for age-related differences in CA1–2 at wave 1 and significant age-related differences in CA1–2 and DG-CA3 at wave 2, coupled with no differences in SUB and EC. This pattern of results is partly in line with the available evidence for cross-sectional age-differences in childhood (Canada et al., 2019; Keresztes et al., 2018; for a review, see: Lee et al., 2017; Riggins et al., 2018). Together, these findings would suggest an ongoing development of hippocampal subfields potentially lasting into adolescence.

In sharp contrast, our hypothesis that DG-CA3 and CA1–2 volumes increase over time was not supported by our longitudinal analysis. Rather, we observed significant positive changes for SUB and EC. We observed substantial individual variation in change over time across all variables except for EC. Some of these variations were due to a negative association between initial values and change. Despite high variability across individuals, we did not find strong change-change associations between subfield volume and memory. Importantly, we found no support for the hypothesis that change in DG-CA3 volume is associated with change in the ability to discriminate highly similar memories. Instead, our data showed an association between change in SUB volume and
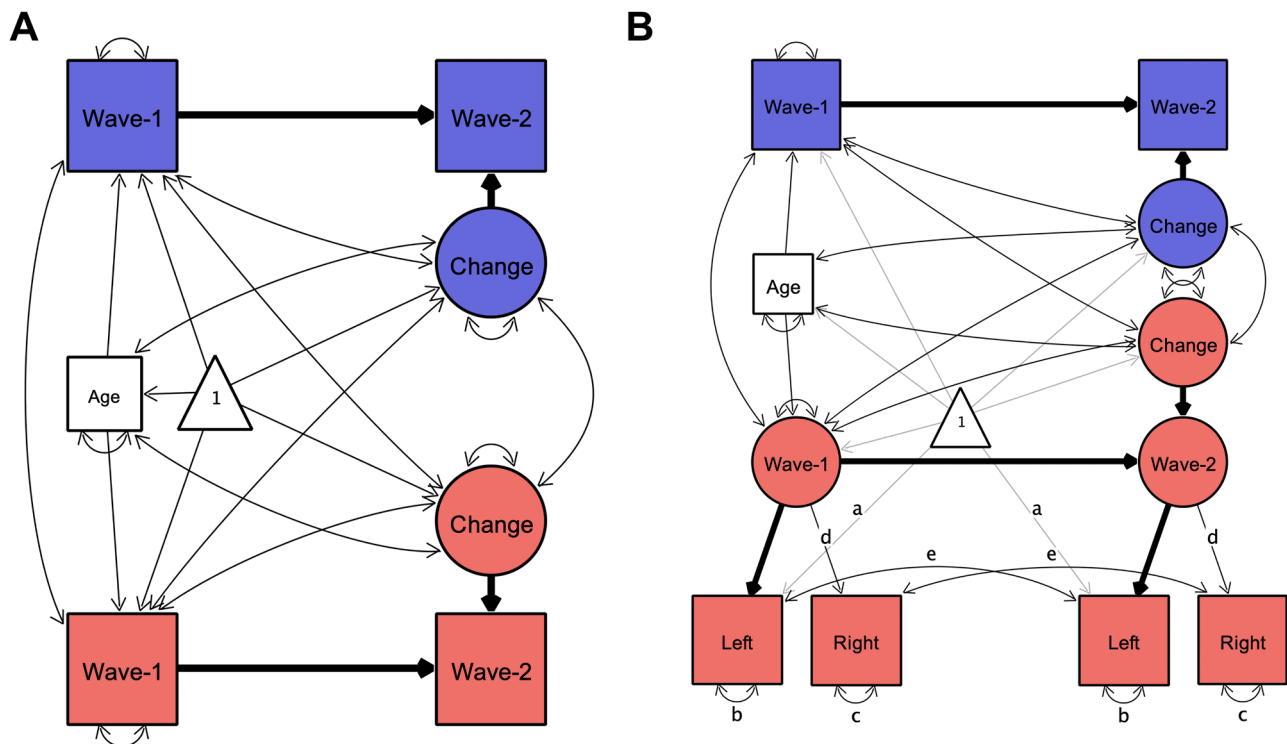
Fig. 5. Illustration of bivariate latent change score models used to assess change-change association between hippocampal subfield volumes and memory measures. Rectangles, circles, and triangles represent indicator variables, latent variables, and means, respectively. Estimated variances, covariances and regression paths are shown as thinner solid lines. Thicker solid lines represent path values fixed at 1. Age represents age at wave 1. Blue and red colors used to distinguish between memory and hippocampal subfield variables respectively. In (A) wave-1 and wave-2 represent observed values of both variables at each wave. In (B), latent wave 1 and wave 2 estimates of hippocampal subfields are modeled by observed left and right hemispheric volumetric measures. Identical path names denoted by (a–e) represent estimated parameters that are constrained to be equal to each other to ensure measurement invariance (Raz et al., 2005). Paths representing means are shaded for better visibility. Parameter estimates of the change-change covariance between hippocampal subfields and memory measures are shown separately in Table 2 for model (A) and in Table S6 for model (B) for better readability of the figure.(For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Figures created by Onyx, version 1.0–1010 (von Oertzen et al., 2015).

change in performance on a spatial memory task, although this weak effect should be dealt with caution. That said, it is worth noting that this effect was not predicted based on the structure–memory associations observed at wave 1, but is somewhat consistent with associations between SUB volume and memory found at wave 2. In addition, we found a positive association between spatial memory performance at wave 1 and change in EC volume. Because this was also an unpredicted and weak association, we refrain from interpreting this specific association until further replication.

4.1. Potential factors driving the disagreement between cross-sectional age-associations and longitudinal trends

Below, we first discuss how these critical discrepancies between cross-sectional and longitudinal data may emerge, and their implication for research on neural and cognitive development. Then, we consider the implications of our specific findings for research on the neuro-cognitive development of memory, and in particular hippocampal sub-field development.

Differences between cross-sectional and longitudinal results are commonly attributed to a multitude of factors, as they stem from different studies with varying methodologies. One strength in our approach here is that we used identical methods across waves, hence effects of discrepant methodological details including delineation of subfields, the choice of volumetric measure, and delays between waves (Canada et al., 2020) were attenuated. However, at least for total hip-pocampal volume, cross-sectionally agreeing automatic segmentation methods have been shown to provide heterogenous estimates for

longitudinal change (Sankar et al., 2017), suggesting that change esti-mation may be noisy even with identical automatic segmentation methods. Although this likely applies to our method, we believe that our analytic approach –modeling hippocampal subfield volumes as latent variables expressed by separate hemispheric indicators– reduced this concern.

Moreover, the cross-sectional age span and the interval between the two waves of the study were both approximately 2 years, allowing us to assess longitudinal and cross-sectional slopes on equal timescales, which rarely is the case in a longitudinal study. Second, given the small age-range of the samples collected at each wave, it is unlikely that cohort effects (Raz and Lindenberger, 2011) drive the observed age-differences in DG-CA3 and CA1–2. Third, there were no selective dropouts that could have led to disagreement between cross-sectional age-differences and longitudinal estimates of change (Nyberg et al., 2010). Although we did not find differences in memory or volumetric measures between dropouts and non-dropouts, we did find that dropouts after wave 1 were older than ongoing participants, and for some variables (see Supplement S1.1) predominantly boys. However, these differences can hardly explain the effects found: We only found sex differences in volume of EC, with boys having smaller volumes than girls. Given that wave 1 volumes were negatively associated with change for all subfields, if sex differences in dropouts had an effect, if any, on observed change in EC, this should have been an attenuation of the observed increase. In addition, because age was included as a covariate in our LCS models estimating change, age differences between dropouts and non-dropouts cannot drive the observed effect of change. That said, based on these post-hoc tests, we cannot fully rule out that we missed out on any effects, had

Table 2

Model fit, and parameter estimates for covariance between change in hippocampal subfields and change in memory measures in bivariate latent change score models.

	LDI		Spatial memory		Cued recall	
	PE (SE)	Δχ2 (1)	PE (SE)	Δχ2 (1)	PE (SE)	Δχ2 (1)
Cov$_{change-change}$						
CA1–2	0 (0.008)	0.002	0.005 (0.008)	0.395	0.005 (0.011)	0.225
DG-CA3	-0.004 (0.012)	0.119	0.012 (0.011)	1.225	-0.003 (0.014)	0.043
SUB	0.002 (0.015)	0.027	0.03 (0.015)	4.395 *	0.006 (0.02)	0.096
EC	-0.001 (0.009)	0.005	-0.002 (0.008)	0.057	-0.006 (0.01)	0.313
HC	-0.004 (0.033)	0.012	0.047 (0.032)	2.199	0.01 (0.041)	0.055
Cov$_{volume\ at\ wave\ 1\ -\ change\ in\ memory}$						
CA1–2	-0.005 (0.011)	0.181	0.006 (0.011)	0.301	0.001 (0.013)	0.008
DG-CA3	-0.009 (0.017)	0.283	0.002 (0.015)	0.026	-0.005 (0.019)	0.083
SUB	-0.016 (0.020)	0.668	-0.008 (0.019)	0.197	-0.005 (0.023)	0.038
EC	-0.005 (0.011)	0.193	0.001 (0.01)	0.004	0.017 (0.012)	2.181
HC	-0.032 (0.042)	0.588	0.001 (0.039)	0.000	-0.011 (0.048)	0.057
Cov$_{memory\ at\ wave\ 1\ -\ change\ in\ volume}$						
CA1–2	0.005 (0.009)	0.283	0.01 (0.008)	1.602	-0.003 (0.009)	0.085
DG-CA3	0.012 (0.012)	1.028	0.01 (0.01)	0.994	0.009 (0.012)	0.531
SUB	0.011 (0.015)	0.485	0.011 (0.014)	0.629	0.001 (0.017)	0.006
EC	0.002 (0.01)	0.034	0.016 (0.007)	4.788[a]	0.001 (0.008)	0.004
HC	0.031 (0.034)	0.846	0.034 (0.03)	1.353	0.006 (0.035)	0.034

Note. M: mean, Var: variance, PE (SE): parameter estimate (standard error), DG: dentate gyrus, SUB: subiculum, EC: entorhinal cortex, HC: hippocampus. LDI: Lure discrimination index. Model fit was perfect for all models, χ2 = 0, RMSEA= 0, CFI= 1. Estimates for error variances, as well as for indicator variable means are not presented.

[a] p < .05, uncorrected for multiple comparisons.

we no dropouts at all. Fourth, using post-hoc power calculations we have also shown that despite the relatively low sample size, our study was well powered to detect the magnitude of change in DG-CA3 and CA1–2 hypothesized based on cross-sectional data. Our null finding of change in CA1–2 is even more unlikely to result from low power as the observed effect was negative. That said, we may still be missing brain-cognition associations due to limited power. Fifth, low sample size may also lead to unreliable spurious correlations, and such correlations have been pinpointed as potential causes behind disagreeing longitudinal and cross-sectional results from the same sample (Louis et al., 1986). That said, cross-sectional age-associations and longitudinal changes reported here are unlikely to be spurious; the cross-sectional results fit well to a consistent line of previous findings.

One additional possibility is that change in hippocampal subfields is nonlinear. The use of only two measurement occasions in our study precluded the detection of any non-linear effects. Louis et al. (1986) have mathematically shown that linearly modeled cross-sectional and longitudinal slopes from the same sample should agree if the age distributions in cross-sectional samples are Gaussian and the change is linear or quadratic nonlinear. Shapiro-Wilk's test of normality indicated that our age distributions were normal (for both waves p > .1). This leaves open the possibility that the disagreement between cross-sectional and longitudinal slopes for hippocampal subfield volumes in our study are due to non-quadratic (e.g., cubic) nonlinearity of actual change. This nonlinearity is supported by both existing longitudinal studies in the field that have more than two measurement occasions. Canada et al. (2021) used an accelerated longitudinal design with a 4- and a 6-year old cohort followed for up to three years. Using this design, they could assess longitudinal change in hippocampal subfield volumes in a sample of children aged 4–8. Intriguingly, the authors found a positive change in all subfields investigated (CA2–4/DG, CA1, and SUB) but only between 5 and 6 years of age. In another study using an accelerated longitudinal design, Tamnes et al. (2018) assessed a sample of 8–26 year-old participants at 3 measurement occasions two year apart from one another. These authors showed quadratic effects for the CA1 and cubic effects for the SUB. Both of these subfields started off with an initial increase until 13–15 years of age, whereas all other regions showed a linear decrease.

Nonlinearity of hippocampal development as well as of its association to memory is also supported by recent longitudinal investigations of developmental change in total hippocampal volume (Herting et al., 2018; Langnes et al., 2020; Lee et al., 2020). Herting et al. (2018) pooled multisite data (n = 216) from three longitudinal samples of 8–22-year-old participants, and found significant nonlinear, including cubic, patterns of change. These nonlinear effects may also partly explain inconsistent results of investigations of developmental change in total hippocampal volume, with studies finding increase (Raffington et al., 2019; Swagerman et al., 2014), decrease (Tamnes et al., 2013), and null-effect (Barnea-Goraly et al., 2014; Gogtay et al., 2006; Sullivan et al., 2011; Yurgelun-Todd et al., 2003). In addition to these data, some cross-sectional studies have also found non-linear age-effects in particular in the case of the SUB (Keresztes et al., 2017; Krogsrud et al., 2014; Lee et al., 2014; Tamnes et al., 2014). Nonetheless, given that cubic nonlinear effects have been observed for the SUB only, this explanation falls short of accommodating our discrepant findings for the other three subfields investigated.

We should note that discrepancies between cross-sectional and longitudinal results may also emerge because sources of within-person variation, i.e., change, and sources of between-person variation, i.e., differences as a function of factors such as age, can largely differ both in scope and magnitude (Ritchie et al., 2016; Schmiedek et al., 2020). As a striking example, in a meta-analysis of 92 studies Seblova et al. (2020) showed that between-person variation in cognitive abilities was in part driven by levels of education, but the latter was unrelated to within-person change in cognitive performance. The same study revealed a large heterogeneity in change in cognition that was unexplained by several factors (e.g., age, GDP) that explained between-person variance in cognition. This underscores the notion that research trying to identify key factors affecting change may not be well informed by cross-sectional associations. A full discussion on how to bridge between-person and within-person approaches is beyond the scope of this paper, but we want to point towards important conceptual and empirical work on this topic (Voelkle et al., 2014).

4.2. Developmental change in the subiculum and its implications for memory development

Our results clearly suggest that the SUB undergoes volumetric increase between ages 6 and 10. This notable finding merits particular

emphasis for two reasons. First, it is partly in line with previous longitudinal as well as cross-sectional studies, and so far provides the only consistent observation across extant longitudinal studies. Second, in investigations of hippocampal subfield development, the developmental trajectory of the SUB and its association to memory has received little attention as compared to the DG, and the CA regions. As pointed out earlier, the association between change in SUB volume and change in spatial memory performance we observed in this study was weak and should be dealt with care. However, the only change–change association between memory measures and hippocampal subfields in studies with children has been reported for the association of SUB and source memory (Canada et al., 2021). Thus, although we caution against overinterpretation, we do suggest that longitudinal associations between SUB and memory development warrant more attention.

The role of the SUB in memory development has been under-investigated. The limited evidence available for humans have linked SUB structure and function to delayed recall in adolescents (Jeon et al., 2019), to mnemonic specificity in young and older adults (Nash et al., 2021; Stark and Stark, 2017), and to spatial learning in a lifespan sample (Daugherty et al., 2016). In addition, cross-sectional studies of hippocampal subfield in children have reported associations of SUB volume with context (Lee et al., 2014), statistical learning (Schlichting et al., 2017), and mnemonic discrimination (Bouyeure et al., 2021), whereas other studies testing for similar associations in children provided null results (see Table S1).

Even in animal studies, the role of the subicular complex – comprising the presubiculum, the parasubiculum, and the SUB – has received little attention, and its functions remain elusive. Although several studies provided evidence for the involvement of the subicular complex in spatial memory and episodic memory, these studies also have highlighted the large heterogeneity of its cytoarchitectonic properties, cognitive functions, and maturational profile (Aggleton, 2012; Brotons-Mas et al., 2017; Ku et al., 2017; Lavenex and Lavenex, 2013; O'Mara et al., 2009, 2001). For instance, based on cross-sectional histological examination of hippocampi of rhesus monkeys, Lavenex and Lavenex (2013) suggested an early maturing network connecting the SUB with the EC, and a later maturing network connecting the SUB to the CA1. The same data also suggested differential developmental trajectories of presubiculum, parasubiculum, and SUB. In addition, given that the subicular regions and neuronal layers are part of distinct hippocampal networks (Aggleton, 2012), deducing their specific functions is difficult from either lesion or activation studies (O'Mara et al., 2009). Because these regions are – by constraints of technological limits – lumped together in high-resolution MR imaging of human hippocampal subfields, it is highly likely that such heterogeneity will contribute to the available observations (Canada et al., 2021; Tamnes et al., 2018) of non-linear development of the SUB. Moreover, SUB measures in human high-resolution volumetry often include some transitional zone in which both SUB and CA1 cells are present; it is unclear whether this poor specificity may contribute to the observed effects or lack thereof.

4.3. Implications for future research on hippocampal contributions to memory development

Based on the results and theoretical consideration presented in this article, we formulate some suggestions for future studies assessing developmental trajectories of hippocampal subfields. First, inclusion of more than two measurement points per individual will allow researchers to detect non-linear change if these exist. More than three measurement time points per individual will further allow researchers to detect non quadratic non-linearity in change if these exist (Ghisletta et al., 2020). Second, a priori power calculations based on the growing number of related studies may help to decide on necessary sample size to detect both linear and non-linear effects of change (for available tools, see Brandmaier et al., 2015; Zhang and Liu, 2018). Importantly, using measures with established test-retest reliability and good sensitivity to

detect change (not only) in hippocampal subfield volumes may help increase power to detect change (Homayouni et al., 2021). Third, and related, longitudinal sample sizes are necessarily constrained by resources available for recruiting and testing participants, MRI hours, as well as manual or semi-automated hippocampal segmentations. Therefore, a viable route to achieve large enough sample sizes may be to combine data in consortia (Herting et al., 2018; cf., Walhovd et al., 2018). Complementing these efforts with non-verbal behavioral measures assessing specific hippocampal functions (e.g., variations of the mnemonic similarity task; Stark et al., 2019) and spatial tasks should enhance such efforts across different nations and regions. Fourth, when reporting longitudinal results, reporting cross-sectional results from the same studies may help tease apart cohort and period effects and actual change. Comparing cross-sectional effects in openly available longitudinal datasets to longitudinal change may be a fruitful direction in this regard.

Longitudinal studies are not without methodological challenges (Louis et al., 1986; Raz and Lindenberger, 2011). Beyond their resource needs, they also amplify practice and test-retest effects in performance (Telzer et al., 2018), which have been related to hippocampal subfield volumes in older adults (Bender et al., 2013). Although practice effects are intuitive in the case of behavioral studies, longitudinal MRI measurement are also subject to them. For instance, because motion-artefacts affect both structural (Reuter et al., 2015) and functional (Satterthwaite et al., 2012) MRI measures, and motion is in turn affected by initial exposure to the scanning environment as well as age, estimations of changes in MRI measures are likely to be confounded by both age and number of repeated measurement occasions (Tamnes et al., 2017; see Satterthwaite et al., 2012 for a demonstration of this effect on functional connectivity). Thus, longitudinal studies need careful planning to incorporate – if possible – measures of practice effects on behavioral and on MRI measures (Maclaren et al., 2012; Telzer et al., 2018).

Finally, longitudinal and cross-sectional results may provide information about distinct mechanisms of change. Complex and sudden changes in one's environment during development, e.g., schooling, may trigger neural changes distinct from ongoing neurodevelopmental change (Brod et al., 2017). For instance, intense new types of learning in the first school year may lead to synaptogenesis in DG/CA3 and CA1–2 (Shors, 2004), and perhaps neurogenesis in DG/CA3 (Boldrini et al., 2018; Sorrells et al., 2018), leading to increase in volume, while at the same time, ongoing developmental reorganization of the hippocampal circuit is accompanied by pruning in other hippocampal subfields (Bagri et al., 2003), leading to a decrease in volume. Combining cross-sectional and longitudinal analysis of the same sample may help tease apart such parallel but opposing changes in indirect measures of neural change, such as volumetry.

4.4. Limitations

Our results need to be interpreted in light of some limitations: Following extant conventions, we followed a protocol for delineating the EC on six consecutive slices anterior to the hippocampal body (see Section 2.2). Using a fix number of slices to extract volumetric measures from at both waves may have biased any estimates of change for EC volume. This makes the observed association between change in EC volume and wave 1 performance on spatial memory challenging. Related to this, due to the low validity of available hippocampal head segmentation methods, our volumetric measures for DG-CA3, CA1–2 and SUB were constrained to the body of HC, therefore we could not capture potential changes in HC head during the observed period. Another limitation of our study is the lack of data collected to allow us to calculate cohort effects, e.g., in which school year participants were. Unfortunately, birth date in this sample is not enough to determine when a child started school. Although cohort effects are unlikely given the small age-range of the sample, schooling at the age of 5–7 has been

shown to affect brain function and cognition independent of age (Brod et al., 2017). This result raises the possibility that brain structure may be affected by schooling, even in such a limited age-range. In addition, although our study was well-powered to detect univariate change, it was underpowered to assess fine-grained effects of covariates on change. For instance, studies on total hippocampal development using larger samples found sex interactions (Herting et al., 2018), and thus warrant exploring sex effects in future studies of hippocampal subfield development.

Further, additional age-related covariates not assessed in this study may provide a more fine-grained picture of neural changes in the hippocampus. Puberty status may be one such variable of interest as it has been shown to have both a main effect and an interaction effect with sex on hippocampal morphology (Goddings et al., 2014). Lastly, we used identical task versions across wave 1 and wave 2, a design feature that did not allow us to separate memory gains from practice effects. For instance, our associative memory task included an unexpected memory test which participants at wave 2 may have anticipated. This in turn may explain gains in performance beyond actual memory development. Similarly, we did not use online motion detection during structural MRI, thus were not able to covary practice effects out of volumetric measures in our study.

5. Conclusion

This study highlights striking inconsistencies between cross-sectional age-associations and longitudinal change. Our results specifically question the hypothesis that the DG-CA3, as well as CA1–2 subfields of the hippocampus undergo protracted volumetric increase in middle childhood and that volumetric change in these regions is related to change in memory performance – at least between 6 and 10 years of age. The inconsistency between cross-sectional and longitudinal estimates of volume–behavior associations has direct implications not only for studies of hippocampal subfield volumes–memory associations, but also more broadly for studies investigating effects of various mediators (e.g., life experiences) of hippocampal subfield development (Phillips et al., 2021), and studies of developmental brain–behavior couplings in general. We hope that emphasizing the observed discrepancies, as well as the outlined mechanisms potentially driving them will prove useful to studies of neurocognitive change.

CRediT authorship contribution statement

Attila Keresztes: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Laurel Raffington**: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – review & editing, Project administration. **Andrew R. Bender**: Software, Formal analysis, Investigation, Writing – review & editing. **Katharina Bögl**: Formal analysis, Investigation, Writing – review & editing. **Christine Heim**: Conceptualization, Resources, Writing – review & editing, Funding acquisition. **Yee Lee Shing**: Conceptualization, Methodology, Formal analysis, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.dcn.2022.101085.

References

Aggleton, J.P., 2012. Multiple anatomical systems embedded within the primate medial temporal lobe: implications for hippocampal function. Neurosci. Biobehav. Rev. Mem. Form. 36, 1579–1596. https://doi.org/10.1016/j.neubiorev.2011.09.005.

Bagri, A., Cheng, H.-J., Yaron, A., Pleasure, S.J., Tessier-Lavigne, M., 2003. Stereotyped pruning of long hippocampal axon branches triggered by retraction inducers of the semaphorin family. Cell 113, 285–299. https://doi.org/10.1016/S0092-8674(03)00267-8.

Bakker, A., Kirwan, C.B., Miller, M., Stark, C.E.L., 2008. Pattern separation in the human hippocampal CA3 and dentate gyrus. Science 319, 1640–1642. https://doi.org/10.1126/science.1152882.

Barnea-Goraly, N., Frazier, T.W., Piacenza, L., Minshew, N.J., Keshavan, M.S., Reiss, A.L., Hardan, A.Y., 2014. A preliminary longitudinal volumetric MRI study of amygdala and hippocampal volumes in autism. Prog. Neuro Psychopharmacol. Biol. Psychiatry 48, 124–128.

Bender, A.R., Daugherty, A.M., Raz, N., 2013. Vascular risk moderates associations between hippocampal subfield volumes and memory. J. Cogn. Neurosci. 25, 1851–1862.

Bender, A.R., Keresztes, A., Bodammer, N.C., Shing, Y.L., Werkle-Bergner, M., Daugherty, A.M., Yu, Q., Kühn, S., Lindenberger, U., Raz, N., 2018. Optimization and validation of automated hippocampal subfield segmentation across the lifespan. Hum. Brain Mapp. 39, 916–931. https://doi.org/10.1002/hbm.23891.

Bender, A.R., Raz, N., 2015. Normal-appearing cerebral white matter in healthy adults: mean change over 2 years and individual differences in change. Neurobiol. Aging 36, 1834–1848. https://doi.org/10.1016/j.neurobiolaging.2015.02.001.

Berron, D., Schütze, H., Maass, A., Cardenas-Blanco, A., Kuijf, H.J., Kumaran, D., Düzel, E., 2016. Strong evidence for pattern separation in human dentate gyrus. J. Neurosci. 36, 7569–7579. https://doi.org/10.1523/JNEUROSCI.0518-16.2016.

Boker, S.M., McArdle, J.J., Neale, M., 2002. An algorithm for the hierarchical organization of path diagrams and calculation of components of expected covariance. Struct. Equ. Model. 9, 174–194.

Boldrini, M., Fulmore, C.A., Tartt, A.N., Simeon, L.R., Pavlova, I., Poposka, V., Rosoklija, G.B., Stankov, A., Arango, V., Dwork, A.J., Hen, R., Mann, J.J., 2018. Human hippocampal neurogenesis persists throughout aging. Cell Stem Cell 22, 589–599.e5. https://doi.org/10.1016/j.stem.2018.03.015.

Bouyeure, A., Patil, S., Mauconduit, F., Poiret, C., Isai, D., Noulhiane, M., 2021. Hippocampal subfield volumes and memory discrimination in the developing brain. Hippocampus 1–13. https://doi.org/10.1002/hipo.23385.

Brandmaier, A.M., von Oertzen, T., Ghisletta, P., Hertzog, C., Lindenberger, U., 2015. LIFESPAN: A tool for the computer-aided design of longitudinal studies. Front. Psychol. 6. https://doi.org/10.3389/fpsyg.2015.00272.

Brod, G., Bunge, S.A., Shing, Y.L., 2017. Does one year of schooling improve children's cognitive control and alter associated brain activation? Psychol. Sci. 28, 967–978. https://doi.org/10.1177/0956797617699838.

Brotons-Mas, J.R., Schaffelhofer, S., Guger, C., O'Mara, S.M., Sanchez-Vives, M.V., 2017. Heterogeneous spatial representation by different subpopulations of neurons in the subiculum. Neuroscience 343, 174–189. https://doi.org/10.1016/j.neuroscience.2016.11.042.

Burgess, N., Maguire, E.A., O'Keefe, J., 2002. The human hippocampus and spatial and episodic memory. Neuron 35, 625–641. https://doi.org/10.1016/S0896-6273(02)00830-9.

Canada, K.L., Botdorf, M., Riggins, T., 2020. Longitudinal development of hippocampal subregions from early- to mid-childhood. Hippocampus 30, 1098–1111. https://doi.org/10.1002/hipo.23218.

Canada, K.L., Hancock, G.R., Riggins, T., 2021. Modeling longitudinal changes in hippocampal subfields and relations with memory from early- to mid-childhood. Dev. Cogn. Neurosci. 48, 100947 https://doi.org/10.1016/j.dcn.2021.100947.

Canada, K.L., Ngo, C.T., Newcombe, N.S., Geng, F., Riggins, T., 2019. It's all in the details: relations between young children's developing pattern separation abilities and hippocampal subfield volumes. Cereb. Cortex 29, 3427–3433. https://doi.org/10.1093/cercor/bhy211.

Daugherty, A.M., Bender, A.R., Yuan, P., Raz, N., 2016. Changes in search path complexity and length during learning of a virtual water maze: age differences and differential associations with hippocampal subfield volumes. Cereb. Cortex 26, 2391–2401. https://doi.org/10.1093/cercor/bhv061.

Daugherty, A.M., Flinn, R., Ofen, N., 2017. Hippocampal CA3-dentate gyrus volume uniquely linked to improvement in associative memory from childhood to adulthood. NeuroImage 153, 75–85. https://doi.org/10.1016/j.neuroimage.2017.03.047.

Eldridge, L.L., Engel, S.A., Zeineh, M.M., Bookheimer, S.Y., Knowlton, B.J., 2005. A dissociation of encoding and retrieval processes in the human hippocampus. J. Neurosci. 25, 3280–3286. https://doi.org/10.1523/JNEUROSCI.3420-04.2005.

Frotscher, M., Seress, L., 2007. Morphological development of the hippocampus. In: The Hippocampus Book. Oxford University Press, pp. 115–131.

Ghisletta, P., Mason, F., von Oertzen, T., Hertzog, C., Nilsson, L.-G., Lindenberger, U., 2020. On the use of growth models to study normal cognitive aging. Int. J. Behav. Dev. 44, 88–96. https://doi.org/10.1177/0165025419851576.

Goddings, A.-L., Mills, K.L., Clasen, L.S., Giedd, J.N., Viner, R.M., Blakemore, S.-J., 2014. The influence of puberty on subcortical brain development. NeuroImage 88, 242–251. https://doi.org/10.1016/j.neuroimage.2013.09.073.

Gogtay, N., Nugent, T.F., Herman, D.H., Ordonez, A., Greenstein, D., Hayashi, K.M., Clasen, L., Toga, A.W., Giedd, J.N., Rapoport, J.L., Thompson, P.M., 2006. Dynamic mapping of normal human hippocampal development. Hippocampus 16, 664–672. https://doi.org/10.1002/hipo.20193.

Herting, M.M., Johnson, C., Mills, K.L., Vijayakumar, N., Dennison, M., Liu, C., Goddings, A.-L., Dahl, R.E., Sowell, E.R., Whittle, S., Allen, N.B., Tamnes, C.K., 2018. Development of subcortical volumes across adolescence in males and females: a multisample study of longitudinal changes. Neuroimage 172, 194–205. https://doi.org/10.1016/j.neuroimage.2018.01.020.

Homayouni, R., Yu, Q., Ramesh, S., Tang, L., Daugherty, A.M., Ofen, N., 2021. Test–retest reliability of hippocampal subfield volumes in a developmental sample: implications for longitudinal developmental studies. J. Neurosci. Res. 99, 2327–2339. https://doi.org/10.1002/jnr.24831.

Jack, C.R., Twomey, C.K., Zinsmeister, A.R., Sharbrough, F.W., Petersen, R.C., Cascino, G.D., 1989. Anterior temporal lobes and hippocampal formations: normative volumetric measurements from MR images in young adults. Radiology 172, 549–554. https://doi.org/10.1148/radiology.172.2.2748838.

Jeon, S., Hwang, S.-I., Son, Y.D., Kim, Y.-B., Lee, Y.J., Kim, S.J., 2019. Association between delayed recall and T2* relaxation time of the subiculum in adolescents: implications for ultra-high-field magnetic resonance imaging. Psychiatry Clin. Neurosci. 73, 340–346. https://doi.org/10.1111/pcn.12843.

Karch, J.D., Filevich, E., Wenger, E., Lisofsky, N., Becker, M., Butler, O., Mårtensson, J., Lindenberger, U., Brandmaier, A.M., Kühn, S., 2019. Identifying predictors of within-person variance in MRI-based brain volume estimates. NeuroImage 200, 575–589. https://doi.org/10.1016/j.neuroimage.2019.05.030.

Keresztes, A., Bender, A.R., Bodammer, N.C., Lindenberger, U., Shing, Y.L., Werkle-Bergner, M., 2017. Hippocampal maturity promotes memory distinctiveness in childhood and adolescence. PNAS 114, 9212–9217. https://doi.org/10.1073/pnas.1710654114.

Keresztes, A., Ngo, C.T., Lindenberger, U., Werkle-Bergner, M., Newcombe, N.S., 2018. Hippocampal maturation drives memory from generalization to specificity. Trends Cogn. Sci. 22, 676–686. https://doi.org/10.1016/j.tics.2018.05.004.

Keresztes, A., Raffington, L., Bender, A.R., Bögl, K., Heim, C., Shing, Y.L., 2020. Hair cortisol concentrations are associated with hippocampal subregional volumes in children. Sci. Rep. 10, 4865. https://doi.org/10.1038/s41598-020-61131-x.

Kessels, R.P., Hobbel, D., Postma, A., 2007. Aging, context memory and binding: a comparison of "what, where and when" in young and older adults. Int. J. Neurosci. 117, 795–810. https://doi.org/10.1080/00207450600910218.

Kievit, R.A., Brandmaier, A.M., Ziegler, G., van Harmelen, A.-L., de Mooij, S.M.M., Moutoussis, M., Goodyer, I.M., Bullmore, E., Jones, P.B., Fonagy, P., Lindenberger, U., Dolan, R.J., 2018. Developmental cognitive neuroscience using latent change score models: a tutorial and applications. Dev. Cogn. Neurosci. 33, 99–117. https://doi.org/10.1016/j.dcn.2017.11.007.

Kirwan, C.B., Stark, C.E.L., 2007. Overcoming interference: an fMRI investigation of pattern separation in the medial temporal lobe. Learn. Mem. 14, 625–633. https://doi.org/10.1101/lm.663507.

Kline, R.B., 2015. Principles and Practice of Structural Equation Modeling. Guilford publications.

Krogsrud, S.K., Tamnes, C.K., Fjell, A.M., Amlien, I., Grydeland, H., Sulutvedt, U., Due-Tønnessen, P., Bjørnerud, A., Sølsnes, A.E., Håberg, A.K., Skrane, J., Walhovd, K.B., 2014. Development of hippocampal subfield volumes from 4 to 22 years. Hum. Brain Mapp. 35, 5646–5657. https://doi.org/10.1002/hbm.22576.

Ku, S., Nakamura, N.H., Maingret, N., Mahnke, L., Yoshida, M., Sauvage, M.M., 2017. Regional specific evidence for memory-load dependent activity in the dorsal subiculum and the lateral entorhinal cortex. Front. Syst. Neurosci. 11. https://doi.org/10.3389/fnsys.2017.00051.

Langnes, E., Sneve, M.H., Sederevicius, D., Amlien, I.K., Walhovd, K.B., Fjell, A.M., 2020. Anterior and posterior hippocampus macro- and microstructure across the lifespan in relation to memory—a longitudinal study. Hippocampus 30, 678–692. https://doi.org/10.1002/hipo.23189.

Lavenex, P., Lavenex, P.B., 2013. Building hippocampal circuits to learn and remember: insights into the development of human memory. Behav. Brain Res., SI:Medial Tempo Lobe Mem. Netw. 254, 8–21. https://doi.org/10.1016/j.bbr.2013.02.007.

Lee, J.K., Ekstrom, A.D., Ghetti, S., 2014. Volume of hippocampal subfields and episodic memory in childhood and adolescence. NeuroImage 94, 162–171. https://doi.org/10.1016/j.neuroimage.2014.03.019.

Lee, J.K., Fandakova, Y., Johnson, E.G., Cohen, N.J., Bunge, S.A., Ghetti, S., 2020. Changes in anterior and posterior hippocampus differentially predict item-space, item-time, and item-item memory improvement. Dev. Cogn. Neurosci. 41, 100741 https://doi.org/10.1016/j.dcn.2019.100741.

Lee, J.K., Johnson, E.G., Ghetti, S., 2017. Hippocampal development: structure, function and implications. In: The Hippocampus from Cells to Systems. Springer, Cham, pp. 141–166. https://doi.org/10.1007/978-3-319-50406-3_6.

Li, A.W.Y., King, J., 2019. Spatial memory and navigation in ageing: a systematic review of MRI and fMRI studies in healthy participants. Neurosci. Biobehav. Rev. 103, 33–49. https://doi.org/10.1016/j.neubiorev.2019.05.005.

Lindenberger, U., Von Oertzen, T., Ghisletta, P., Hertzog, C., 2011. Cross-sectional age variance extraction: what's change got to do with it? Psychol. Aging 26, 34–47. https://doi.org/10.1037/a0020525.

Louis, T.A., Robins, J., Dockery, D.W., Spiro, A., Ware, J.H., 1986. Explaining discrepancies between longitudinal and cross-sectional models. J. Chronic Dis. 39, 831–839. https://doi.org/10.1016/0021-9681(86)90085-8.

Maclaren, J., Armstrong, B.S.R., Barrows, R.T., Danishad, K.A., Ernst, T., Foster, C.L., Gumus, K., Herbst, M., Kadashevich, I.Y., Kusik, T.P., Li, Q., Lovell-Smith, C., Prieto, T., Schulze, P., Speck, O., Stucht, D., Zaitsev, M., 2012. Measurement and correction of microscopic head motion during magnetic resonance imaging of the brain. PLOS ONE 7, e48088. https://doi.org/10.1371/journal.pone.0048088.

Maclaren, J., Han, Z., Vos, S.B., Fischbein, N., Bammer, R., 2014. Reliability of brain volume measurements: a test-retest dataset. Sci. Data 1, 140037. https://doi.org/10.1038/sdata.2014.37.

Madan, C.R., Kensinger, E.A., 2017. Test–retest reliability of brain morphology estimates. Brain Inf. 4, 107–121. https://doi.org/10.1007/s40708-016-0060-4.

McArdle, J.J., Nesselroade, J.R., 1994. Using multivariate data to structure developmental change.

Nash, M.I., Hodges, C.B., Muncy, N.M., Kirwan, C.B., 2021. Pattern separation beyond the hippocampus: a high-resolution whole-brain investigation of mnemonic discrimination in healthy adults. Hippocampus 1–14. https://doi.org/10.1002/hipo.23299.

Nyberg, L., Salami, A., Andersson, M., Eriksson, J., Kalpouzos, G., Kauppi, K., Lind, J., Pudas, S., Persson, J., Nilsson, L.-G., 2010. Longitudinal evidence for diminished frontal cortex function in aging. PNAS 107, 22682–22686. https://doi.org/10.1073/pnas.1012651108.

O'Mara, S.M., Commins, S., Anderson, M., Gigg, J., 2001. The subiculum: a review of form, physiology and function. Prog. Neurobiol. 64, 129–155. https://doi.org/10.1016/S0301-0082(00)00054-X.

O'Mara, S.M., Sanchez-Vives, M.V., Brotons-Mas, J.R., O'Hare, E., 2009. Roles for the subiculum in spatial information processing, memory, motivation and the temporal control of behaviour. Progress in neuro-psychopharmacology and biological psychiatry. Spec. SECTION Hippocampal Funct. Spat. Mem. 33, 782–790. https://doi.org/10.1016/j.pnpbp.2009.03.040.

Van Petten, C., 2004. Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: review and meta-analysis. Neuropsychologia 42, 1394–1413. https://doi.org/10.1016/j.neuropsychologia.2004.04.006.

Pfefferbaum, A., Sullivan, E.V., 2015. Cross-sectional versus longitudinal estimates of age-related changes in the adult brain: overlaps and discrepancies. Neurobiol. Aging 36, 2563–2567. https://doi.org/10.1016/j.neurobiolaging.2015.05.005.

Phillips, R.D., De Bellis, M.D., Brumback, T., Clausen, A.N., Clarke-Rubright, E.K., Haswell, C.C., Morey, R.A., 2021. Volumetric trajectories of hippocampal subfields and amygdala nuclei influenced by adolescent alcohol use and lifetime trauma. Transl. Psychiatry 11, 1–13. https://doi.org/10.1038/s41398-021-01275-0.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, VIenna, Austria.

Raffington, L., Czamara, D., Mohn, J.J., Falck, J., Schmoll, V., Heim, C., Binder, E.B., Shing, Y.L., 2019. Stable longitudinal associations of family income with children's hippocampal volume and memory persist after controlling for polygenic scores of educational attainment. Dev. Cogn. Neurosci. 40, 100720 https://doi.org/10.1016/j.dcn.2019.100720.

Raffington, L., Falck, J., Heim, C., Mather, M., Shing, Y.L., 2020. Effects of stress on 6- and 7-year-old children's emotional memory differs by gender. J. Exp. Child Psychol. 199, 104924 https://doi.org/10.1016/j.jecp.2020.104924.

Raffington, L., Prindle, J., Keresztes, A., Binder, J., Heim, C., Shing, Y.L., 2018. Blunted cortisol stress reactivity in low–income children relates to lower memory function. Psychoneuroendocrinology 90, 110–121. https://doi.org/10.1016/j.psyneuen.2018.02.002.

Raz, N., Lindenberger, U., 2011. Only time will tell: cross-sectional studies offer no solution to the age-brain-cognition triangle—comment on. Psychol. Bull. 137, 790–795. https://doi.org/10.1037/a0024503.

Raz, N., Lindenberger, U., Rodrigue, K.M., Kennedy, K.M., Head, D., Williamson, A., Dahle, C., Gerstorf, D., Acker, J.D., 2005. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. Cereb. Cortex 15, 1676–1689. https://doi.org/10.1093/cercor/bhi044.

Reuter, M., Tisdall, M.D., Qureshi, A., Buckner, R.L., van der Kouwe, A.J.W., Fischl, B., 2015. Head motion during MRI acquisition reduces gray matter volume and thickness estimates. NeuroImage 107, 107–115. https://doi.org/10.1016/j.neuroimage.2014.12.006.

Riggins, T., Geng, F., Botdorf, M., Canada, K.L., Cox, L., Hancock, G.R., 2018. Protracted hippocampal development is associated with age-related improvements in memory during early childhood. NeuroImage 174, 127–137. https://doi.org/10.1016/j.neuroimage.2018.03.009.

Ritchie, S.J., Tucker-Drob, E.M., Cox, S.R., Corley, J., Dykiert, D., Redmond, P., Pattie, A., Taylor, A.M., Sibbett, R., Starr, J.M., Deary, I.J., 2016. Predictors of ageing-related decline across multiple cognitive functions. Intelligence 59, 115–126. https://doi.org/10.1016/j.intell.2016.08.007.

Rosseel, Y., 2012. lavaan: an R package for structural equation modeling. J. Stat. Softw. 48, 1–36. https://doi.org/10.18637/jss.v048.i02.

Sankar, T., Park, M.T.M., Jawa, T., Patel, R., Bhagwat, N., Voineskos, A.N., Lozano, A.M., Chakravarty, M.M., 2017. Your algorithm might think the hippocampus grows in Alzheimer's disease: caveats of longitudinal automated hippocampal volumetry. Hum. Brain Mapp. 38, 2875–2896. https://doi.org/10.1002/hbm.23559.

Satterthwaite, T.D., Wolf, D.H., Loughead, J., Ruparel, K., Elliott, M.A., Hakonarson, H., Gur, R.C., Gur, R.E., 2012. Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. NeuroImage 60, 623–632. https://doi.org/10.1016/j.neuroimage.2011.12.063.

Schlichting, M.L., Guarino, K.F., Schapiro, A.C., Turk-Browne, N.B., Preston, A.R., 2017. Hippocampal structure predicts statistical learning and associative inference abilities during development. J. Cogn. Neurosci. 29, 37–51. https://doi.org/10.1162/jocn_a_01028.

Schmiedek, F., Lövdén, M., von Oertzen, T., Lindenberger, U., 2020. Within-person structures of daily cognitive performance differ from between-person structures of cognitive abilities. PeerJ 8, e9290. https://doi.org/10.7717/peerj.9290.

Scoville, W.B., Milner, B., 1957. Loss of recent memory after bilateral hippocampal lesions. J. Neurol. Neurosurg. Psychiatry 20, 11–21.

Seblova, D., Berggren, R., Lövdén, M., 2020. Education and age-related decline in cognitive performance: systematic review and meta-analysis of longitudinal cohort studies. Ageing Res. Rev. 58, 101005 https://doi.org/10.1016/j.arr.2019.101005.

Shors, T.J., 2004. Memory traces of trace memories: neurogenesis, synaptogenesis and awareness. Trends Neurosci. 27, 250–256. https://doi.org/10.1016/j.tins.2004.03.007.

Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17, 143–155. https://doi.org/10.1002/hbm.10062.

Sorrells, S.F., Paredes, M.F., Cebrian-Silla, A., Sandoval, K., Qi, D., Kelley, K.W., James, D., Mayer, S., Chang, J., Auguste, K.I., Chang, E.F., Gutierrez, A.J., Kriegstein, A.R., Mathern, G.W., Oldham, M.C., Huang, E.J., Garcia-Verdugo, J.M., Yang, Z., Alvarez-Buylla, A., 2018. Human hippocampal neurogenesis drops sharply in children to undetectable levels in adults. Nature 555, 377–381. https://doi.org/10.1038/nature25975.

Stark, S.M., Kirwan, C.B., Stark, C.E.L., 2019. Mnemonic similarity task: a tool for assessing hippocampal integrity. Trends Cogn. Sci. 23, 938–951. https://doi.org/10.1016/j.tics.2019.08.003.

Stark, S.M., Stark, C.E.L., 2017. Age-related deficits in the mnemonic similarity task for objects and scenes. Behav. Brain Res. 333, 109–117. https://doi.org/10.1016/j.bbr.2017.06.049.

Sullivan, E.V., Pfefferbaum, A., Rohlfing, T., Baker, F.C., Padilla, M.L., Colrain, I.M., 2011. Developmental change in regional brain structure over 7 months in early adolescence: Comparison of approaches for longitudinal atlas-based parcellation. NeuroImage 57, 214–224. https://doi.org/10.1016/j.neuroimage.2011.04.003.

Swagerman, S.C., Brouwer, R.M., de Geus, E.J.C., Hulshoff Pol, H.E., Boomsma, D.I., 2014. Development and heritability of subcortical brain volumes at ages 9 and 12. Genes, Brain Behav. 13, 733–742.

Tamnes, C.K., Bos, M.G.N., van de Kamp, F.C., Peters, S., Crone, E.A., 2018. Longitudinal development of hippocampal subregions from childhood to adulthood. Dev. Cogn. Neurosci. 30, 212–222. https://doi.org/10.1016/j.dcn.2018.03.009.

Tamnes, C.K., Herting, M.M., Goddings, A.-L., Meuwese, R., Blakemore, S.-J., Dahl, R.E., Güroğlu, B., Raznahan, A., Sowell, E.R., Crone, E.A., Mills, K.L., 2017. Development of the cerebral cortex across adolescence: a multisample study of inter-related longitudinal changes in cortical volume, surface area, and thickness. J. Neurosci. 37, 3402–3412. https://doi.org/10.1523/JNEUROSCI.3302-16.2017.

Tamnes, C.K., Walhovd, K.B., Dale, A.M., Østby, Y., Grydeland, H., Richardson, G., Westlye, L.T., Roddey, J.C., Hagler, D.J., Due-Tønnessen, P., Holland, D., Fjell, A.M., 2013. Brain development and aging: overlapping and unique patterns of change. NeuroImage 68, 63–74. https://doi.org/10.1016/j.neuroimage.2012.11.039.

Tamnes, C.K., Walhovd, K.B., Engvig, A., Grydeland, H., Krogsrud, S.K., Østby, Y., Holland, D., Dale, A.M., Fjell, A.M., 2014. Regional hippocampal volumes and development predict learning and memory. Dev. Neurosci. 36, 161–174. https://doi.org/10.1159/000362445.

Telzer, E.H., McCormick, E.M., Peters, S., Cosme, D., Pfeifer, J.H., van Duijvenvoorde, A.C.K., 2018. Methodological considerations for developmental longitudinal fMRI research. Dev. Cogn. Neurosci., Methodol. Chall. Dev. Neuroimaging Contemp. Approaches Solut. 33, 149–160. https://doi.org/10.1016/j.dcn.2018.02.004.

Utsunomiya, H., Takano, K., Okazaki, M., Mitsudome, A., 1999. Development of the temporal lobe in infants and children: analysis by mr-based volumetry. Am. J. Neuroradiol. 20, 717–723.

Voelkle, M.C., Brose, A., Schmiedek, F., Lindenberger, U., 2014. Toward a unified framework for the study of between-person and within-person structures: building a bridge between two research paradigms. Multivar. Behav. Res 49, 193–213. https://doi.org/10.1080/00273171.2014.889593.

von Oertzen, T., Brandmaier, A.M., Tsang, S., 2015. Structural equation modeling with Ωnyx. Struct. Equ. Model. A Multidiscip. J. 22, 148–161. https://doi.org/10.1080/10705511.2014.935842.

Walhovd, K.B., Fjell, A.M., Westerhausen, R., Nyberg, L., Ebmeier, K.P., Lindenberger, U., Bartrés-Faz, D., Baaré, W.F.C., Siebner, H.R., Henson, R., Drevon, C.A., Knudsen, G.P.S., Ljøsne, I.B., Penninx, B.W.J.H., Ghisletta, P., Rogeberg, O., Tyler, L., Bertram, L., Consortium, L., 2018. Healthy minds 0–100 years: optimising the use of European brain imaging cohorts ("Lifebrain"). Eur. Psychiatry 50, 47–56. https://doi.org/10.1016/j.eurpsy.2017.12.006.

Yassa, M.A., Stark, C.E., 2011. Pattern separation in the hippocampus. Trends Neurosci. 34, 515–525.

Yurgelun-Todd, D.A., Killgore, W.D., Cintron, C.B., 2003. Cognitive correlates of medial temporal lobe development across adolescence: a magnetic resonance imaging study. Percept. Mot. Skills 96, 3–17.

Yushkevich, P.A., Amaral, R.S.C., Augustinack, J.C., Bender, A.R., Bernstein, J.D., Boccardi, M., Bocchetta, M., Burggren, A.C., Carr, V.A., Chakravarty, M.M., Chételat, G., Daugherty, A.M., Davachi, L., Ding, S.L., Ekstrom, A., Geerlings, M.I., Hassan, A., Huang, Y., Iglesias, J.E., La Joie, R., Kerchner, G.A., LaRocque, K.F., Libby, L.A., Malykhin, N., Mueller, S.G., Olsen, R.K., Palombo, D.J., Parekh, M.B., Pluta, J.B., Preston, A.R., Pruessner, J.C., Ranganath, C., Raz, N., Schlichting, M.L., Schoemaker, D., Singh, S., Stark, C.E.L., Suthana, N., Tompary, A., Turowski, M.M., Van Leemput, K., Wagner, A.D., Wang, L., Winterburn, J.L., Wisse, L.E.M., Yassa, M.A., Zeineh, M.M., 2015a. Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in vivo MRI: towards a harmonized segmentation protocol. NeuroImage 111, 526–541.

Yushkevich, P.A., Pluta, J.B., Wang, H., Xie, L., Ding, S.-L., Gertje, E.C., Mancuso, L., Kliot, D., Das, S.R., Wolk, D.A., 2015b. Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. Hum. Brain Mapp. 36, 258–287. https://doi.org/10.1002/hbm.22627.

Zhang, Z., Hamagami, F., Grimm, K.J., McArdle, J.J., 2015. Using R package RAMpath for tracing SEM path diagrams and conducting complex longitudinal data analysis. Struct. Equ. Model. A Multidiscip. J. 22, 132–147.

Zhang, Z., Liu, H., 2018. Sample Size and Measurement Occasion Planning for Latent Change Score Models Through Monte Carlo Simulation. Advances in Longitudinal Models for Multivariate Psychology: A Festschrift for Jack McArdle. Taylor & Francis.