

This is the peer reviewed version of the following article: Sjölund MJ, González-Díaz P, Moreno-Villena JJ, Jump AS. Brod, G., Lindenberger, U. and Shing, Y. L. (2017), Neural activation patterns during retrieval of schema-related memories: differences and commonalities between children and adults. *Dev Sci*, 20: n/a, e12475, which has been published in final form at <https://doi.org/10.1111/desc.12475>. This article may be used for non-commercial purposes in accordance With Wiley Terms and Conditions for self-archiving.

Running Head: DEVELOPMENT OF SCHEMA EFFECTS

Neural activation patterns during retrieval of schema-related memories: Differences and commonalities between children and adults

Garvin Brod^a, Ulman Lindenberger^{a,b}, & Yee Lee Shing^{a,c}

^aMax Planck Institute for Human Development, Center for Lifespan Psychology, Berlin, Germany ^bEuropean University Institute, San Domenico di Fiesole, Italy; ^cDivision of Psychology, School of Natural Sciences, University of Stirling, UK

Forthcoming in *Developmental Science*

Correspondence concerning this article should be addressed to Yee Lee Shing, Division of Psychology, School of Natural Sciences, University of Stirling, UK. Email: yee.shing@stir.ac.uk

Research Highlights

- To overcome the confound between knowledge level and maturation in developmental research on schema-related memory processing, we developed a paradigm in which a schema is experimentally induced to a similar degree in children aged 8–12 years and young adults.
- Despite comparable memory performance, children showed less medial prefrontal cortex (PFC) activity than adults for successfully retrieved schema-congruent events. Furthermore, medial PFC activity during successful retrieval correlated positively with children's age.
- For successfully retrieved schema-*incongruent* events, children showed stronger hippocampus activation as well as weaker connectivity between the striatum and the dorsolateral PFC than adults.
- The results indicate an age-related shift from a more hippocampus-bound memory processing to an increasing recruitment of prefrontal brain regions in the retrieval of schema-related events

Abstract

Schemas represent stable properties of individuals' experiences, and allow to classify new events as being congruent or incongruent with existing knowledge. Research with adults indicates that the prefrontal cortex (PFC) is involved in memory retrieval of schema-related information. However, developmental differences between children and adults in the neural correlates of schema-related memories are not well understood. One reason for this is the inherent confound between schema-relevant experience and maturation, as both are related to time. To overcome this limitation, we used a novel paradigm that experimentally induces, and then probes for task-relevant knowledge during encoding of new information. Thirty-one children aged 8–12 years and 26 young adults participated in the experiment. While successfully retrieving schema-*congruent* events, children showed less medial PFC activity than adults. In addition, medial PFC activity during successful retrieval correlated positively with children's age. While successfully retrieving schema-*incongruent* events, children showed stronger hippocampus (HC) activation as well as weaker connectivity between the striatum and the dorsolateral PFC than adults. These findings were corroborated by an exploratory full-factorial analysis investigating age differences in the retrieval of schema-congruent versus schema-incongruent events, comparing the two conditions directly. Consistent with the findings of the separate analyses, two clusters, one in the medial PFC, one in the HC, were identified that exhibited a memory x congruency x age group interaction. In line with the two-component model of episodic memory development, the present findings point to an age-related shift from a more HC-bound processing to an increasing recruitment of prefrontal brain regions in the retrieval of schema-related events.

Acknowledgments

The study was conducted within the ‘Cognitive and Neuronal Dynamics of Memory across the Lifespan’ (CONMEM) project at the Center for Lifespan Psychology, Max Planck Institute for Human Development, in partial fulfilment of the doctoral dissertation of GB. This work was funded by the Max Planck Society and the German Research Foundation (grant number DFG SH550/2-1 to YLS). GB was supported by a fellowship of the International Max Planck Research School on the Life Course (LIFE; www.imprs-life.mpg.de). UL was financially supported by a Gottfried Wilhelm Leibniz Award 2010 from the Deutsche Forschungsgemeinschaft, and a Fernand Braudel Senior Fellowship of the European University Institute. YLS was financially supported by a Minerva Research Group from the Max Planck Society. We thank Thad Polk, Markus Werkle-Bergner, and Mara Mather for their comments.

Introduction

Piaget (1929) coined the term *schema* to describe children's patterns of thought and behaviour, and postulated that schemas greatly influence children's interaction with the world. In particular, schemas have been shown to support memory for new, related information in both children (Chi, 1978) and young adults (Bransford & Johnson, 1972; Brod, Werkle-Bergner, & Shing, 2013). In addition, memory generalization and integration based on schemas is highly important to allow flexible behaviour in everyday life since we constantly face similar, but slightly different situations to which we have to react quickly based on our prior knowledge (Schlichting & Preston, 2015). Recent neuroscience studies in rodents (Navawongse & Eichenbaum, 2013; Tse et al., 2007, 2011) and humans (Ghosh, Moscovitch, Melo Colella, & Gilboa, 2014; van Kesteren et al., 2013; van Kesteren, Rijpkema, Ruiter, & Fernández, 2010) suggest that the medial prefrontal cortex (mPFC) and the hippocampus (HC) play important roles in schema-related memory processing. It is suggested that the mPFC detects overlaps between new experiences and existing schemas; when the overlap (congruency) is sufficiently large, it biases memory reactivation in the HC towards relevant, schema-consistent memories (Preston & Eichenbaum, 2013; van Kesteren, Ruiter, Fernández, & Henson, 2012).

In contrast to a situation in which new experiences fit in nicely into the existing knowledge network, there are also cases in which there is an apparent mismatch between existing knowledge and a novel experience. As an example, think of the last time your favourite football team lost against another team from a lower division. We recently provided evidence for a distinction between the neural mechanisms underlying memory retrieval of schema-congruent and schema-incongruent information in young adults (Brod et al., 2015). Whereas successfully retrieving schema-*congruent* events was related to enhanced activation in the mPFC, retrieving schema-*incongruent* events was related to enhanced activation in the dorsolateral PFC (DLPFC) and in the striatum as well as to enhanced connectivity between

these two areas. Brod et al. (2013) proposed earlier that retrieving information that is related but incongruent with a current schema requires the reinstatement of the encoding context and the suppression of the schema-congruent response. These processes have been associated with both the DLPFC and the striatum, and with enhanced connectivity between the two (Han, O'Connor, Eslick, & Dobbins, 2012; Scimeca & Badre, 2012). However, whether these regions are similarly involved in children's retrieval of schema-congruent and schema-incongruent events is thus far unknown.

Brain areas important for memory retrieval display both functional and structural changes during childhood and adolescence. The lateral PFC (IPFC), which is involved in strategic and control processes that act upon memory representations, is known to continue maturation at least until adolescence (Shaw et al., 2008; Sowell et al., 2003). The mPFC displays a maturational pattern that is similar to that of the IPFC, with the exception of its most ventral parts, which already mature during middle childhood (Shaw et al., 2008). In comparison to the PFC, the HC, which is involved in binding processes that integrate features into a coherent event, is characterized by a faster pace of structural maturation (Sowell et al., 2003). However, there is evidence for continued structural development in parts of the hippocampus as well as in hippocampal projections to cortical regions (DeMaster, Pathman, Lee, & Ghetti, 2013; Ghetti, DeMaster, Yonelinas, & Bunge, 2010; Ghetti & Bunge, 2012; Gogtay et al., 2006). Whether this ongoing structural development is reflected in age-related differences in hippocampal function, such as binding, is currently unclear. The indirect evidence available from developmental fMRI studies on memory encoding and retrieval has been inconsistent, showing both age-invariant hippocampal activation (Ofen et al., 2007, 2012) as well as age-dependent increases (Ghetti et al., 2010), decreases (Maril et al., 2010), and shifts in localization (DeMaster & Ghetti, 2013). In contrast, the protracted structural development of the PFC has been more consistently linked to an increase in PFC activation across middle childhood (see Ofen, 2012). Immature PFC processes are known to contribute

to children's difficulties in using strategies and in controlling memory retrieval (for a review, see Shing et al., 2010). So far, however, functional studies have not distinguished between the medial and the lateral PFC, which contribute differentially to the successful retrieval of congruent and incongruent events in young adults. Thus, a further goal of this study was to investigate whether medial and lateral PFC contributions to memory retrieval of schema-congruent and schema-incongruent information, respectively, can already be separated in middle childhood.

Assessing developmental influences on the neural correlates of schema-related memory retrieval is notoriously difficult. First and foremost, schema strength and age-graded maturation changes are often confounded, as both are related to time; some studies even have taken age as a proxy for schema strength (Stangor & McMillan, 1992). Most previous quasi-experimental work that used schemas of world knowledge to explore age differences in associations among prior knowledge, learning, and memory, confounds age-graded differences in pre-existing knowledge structures with age-graded differences in the maturation of relevant brain structures, and the interactions among these two sets of processes (Brod et al., 2013; Ofen, 2012). Second, different age groups may perceive congruency of new information with existing knowledge differently. Children and young adults might follow different criteria for deciding whether something fits their knowledge or not (Maril et al., 2011; Sloutsky & Fisher, 2004). Incongruity, in particular, is hard to ensure, because it can often be resolved by means of idiosyncratic processing, for example by using a link between two objects (e.g., an umbrella does not match a bathtub, but both are related to water). Third, age differences in general memory performance might overshadow schema effects of interest. The improvement of episodic memory during childhood and adolescence has been linked to more elaborate and strategic processing (Schneider & Pressley, 1999; Shing et al., 2008), which could explain an increasing recruitment of PFC areas for memory with age (Ofen, 2012; Ofen et al., 2007). Thus, to reveal age differences in neural mechanisms that underlie

schematic memory processing, differences in performance should be made as minimal as possible.

To overcome these obstacles, we recently introduced a novel paradigm in which schemas are experimentally induced (see Brod et al., 2015, for a study with young adults). The general structure of the paradigm is as follows. On the first day, participants acquire knowledge about a hierarchy of novel objects. On the next day, they encode events (competitions) for pairs of these objects, and are tested on their memory for the outcome of these competitions, which can be either congruent or incongruent to the hierarchy schema learned on the first day. In line with the literature, we found enhanced mPFC activity for successfully retrieved events that were congruent with the induced schema. Importantly, for successfully retrieved incongruent events, we observed enhanced activity in the posterior parietal cortex, dorsolateral PFC, and striatum, along with enhanced connectivity between the latter two (Brod et al., 2015).

In the present study, we performed the same paradigm in children, thereby allowing us to examine how newly acquired prior knowledge differently affects the neural correlates of memory processing in children compared to younger adults. The paradigm effectively unconfounds schema strength and age. The semantics of the induced knowledge network rendered it likely that schema-congruent and schema-incongruent events would be perceived similarly across age and could therefore be clearly distinguished by both children and adults. Moreover, extensive piloting allowed us to attenuate age differences in memory performance by using a lower number of trials in the children than in the young adults.

In sum, this study investigates developmental differences in the neural mechanisms supporting memory retrieval of schema-congruent and schema-incongruent events. To this end, we compared a group of children aged 8–12 with the young adult sample reported in Brod et al. (2015). Our analyses focused on brain areas that have been found to relate to schema-related memory processing in young adults (Brod et al., 2015). Specifically, we

predicted age-related increases in the engagement of mPFC in the successful retrieval of *schema-congruent* events, and of IPFC (particularly the DLPFC) in the successful retrieval of *schema-incongruent* events. Furthermore, we expected connectivity between striatum and DLPFC to be less pronounced in children than in young adults. In contrast, we hypothesized that the HC would show no age-related differences in activation. Finally, in an exploratory analysis complementing the separate analyses for *schema-congruent* and *schema-incongruent* events, we also examined the full factorial design (memory x age group x congruency interaction) to investigate age-related differences in the differential retrieval of *schema-congruent* versus *schema-incongruent* events.

Materials and Methods

Participants

The sample consisted of 31 right-handed children aged 8–12 (17 females, 103–144 months, mean age 125 months), and 26 right-handed young adults (13 females, aged 20–30, mean age 24.6 years). All participants were healthy and had normal or corrected-to-normal vision. The children were recruited via flyers that were handed out to their parents at schools and sports clubs. The young adults were recruited from Berlin universities. Participants were paid 37 Euros. In total, 6 children and 2 young adults had to be excluded for various reasons: two young adults did not acquire the hierarchy; three children moved excessively in the scanner (>1.5mm mean displacement in at least two runs and rating based on visual inspection); and further three children had below chance performance in the incongruent condition. Ethics approval was obtained from the ethics committee of the German Psychological Society (DGPs). Adult participants and the children's parents/legal guardians gave written, informed consent.

Stimuli

Participants acquired a hierarchy of so-called fribbles (stimulus images courtesy of Michael J. Tarr, Center for the Neural Basis of Cognition and Department of Psychology,

Carnegie Mellon University), which are artificial stimuli developed for visual neuroscience experiments. The fribbles were designed in accordance with a species structure (12 species in total in the database), which is based on common bodies and differing types of four appendages. The different exemplars within a species can be distinguished by the shape of their four appendages, of which each can have three different occurrences (81 exemplars in total per species). For the current experiment, we chose the 36 exemplars that were most distinct from one another (4 exemplars for each of the 9 species). With these 36 exemplars we constructed a hierarchy that consisted of three levels, each level having 12 exemplars. Two versions of this hierarchy with different allocations of fribbles to levels were randomly used across participants for counterbalance. During the course of the experiment, the fribbles were always presented in pairs (see 'Task and Procedure' below). All pairs consisted of fribbles of different hierarchy levels and were unique in the sense that each specific pairing was only used once.

Tasks and Procedures

Day 1: Hierarchy learning phase

The tasks and procedures on the first day of the experiment were identical for children and young adults and took between 90 and 120 minutes. Participants acquired the fribbles hierarchy via a trial-and-error learning task (see Figure 1). The cover story for this part was that the fribbles were highly competitive and enjoyed racing against one another; the participants' task was to find out which fribbles were fast, medium, or slow. To do so, the participants had to predict the winner of a race of two fribbles. Corrective feedback was given following their prediction. The feedback was provided via a crown above the winning fribble and right/wrong feedback about the participants' decisions. The session consisted of two parts. In the first part, the 36 fribbles were assigned to three different blocks based on the colour of their body. In each block, the participants had to learn the hierarchy position of 12 fribbles (4 fast, 4 medium, 4 slow). Each block consisted of 36 races of two fribbles each and

was repeated until the participants reached a block performance of at least 90% correctly predicted outcomes. After completion of each block, participants were presented with all 12 fribbles again and were asked to arrange them according to the learnt hierarchy levels. This last test was to ensure explicit knowledge of the hierarchy and participants were again given corrective feedback.

In the second part of the first day, all 36 fribbles were presented within one block. This time, participants were presented with the race's outcome (indicated by a crown above the winner's head) and were asked to indicate whether or not they had expected that outcome based on their learnt hierarchy. Half of the presented outcomes were in accordance with the learnt hierarchy (congruent), the other half of the outcomes violated the learnt hierarchy (incongruent). Again, blocks of 36 pairs were presented until the participants reached a block performance of at least 90% correctly classified congruent/incongruent outcomes. Participants stated their answer on a 6-point confidence scale, ranging from completely expected to completely unexpected. These settings were similar to those used for the memory task on Day 2, and thus served to enhance participants' familiarity with the final task set-up. After reaching the performance criterion, participants were presented with all 36 exemplars again and had to arrange them in the according hierarchy levels.

Day 2: fMRI memory task

On day 2, the tasks and procedures were mostly the same for the children and young adults. The only two differences were that children had a practice phase before the memory task and that each block contained only 18 trials for the children, instead of 22 for the young adults. The lower trial number for the children was determined during pilot experiments to ensure comparable overall memory performance of the age groups.

The second day started with a repetition of the second part of Day 1. The performance criterion was again 90% correctly classified congruent/incongruent outcomes in a block. Then participants were again presented with all 36 fribbles and asked to arrange them by hierarchy

level. Prior to scanning, the participants were instructed about the memory task and the children carried out an additional short practice phase outside the scanner that resembled the later memory task; young adults entered the scanner immediately.

In the scanner, an initial 9-minute structural scan was made, during which the participants watched a child-friendly movie. This was followed by the memory task (see Figure 1), consisting of 9 encoding–retrieval blocks of 18 pairs each (162 unique pairs in total, 50% congruent outcomes) for the children and of 8 encoding–retrieval blocks of 22 pairs each (176 unique pairs in total, 50% congruent outcomes) for the young adults. Encoding and retrieval blocks were presented alternately. For the children, cycles of encoding and retrieval were divided into three runs that were separated by longer breaks of about 5 minutes, during which the children were allowed out of the scanner for a rest.

During encoding, participants were presented with new pairs of fribbles from different hierarchy levels and schema-congruent/schema-incongruent outcomes of “competitions” between them that were indicated by a crown above the winner’s head. After seeing the pair and the outcome for 4 seconds, participants were asked to indicate whether the outcome was expected (congruent condition) or not expected (incongruent condition) based on the hierarchy. In addition, they had to state their confidence on a 6-point confidence scale, with which participants had been familiarized during the second learning phase on Day 1. They were given 2 seconds for their decision. The purpose of this task was to prompt participants of the hierarchy, and to prohibit the use of purely perceptual strategies. Congruent and incongruent outcomes were equally frequent and participants were informed of this before they entered the scanner. There was a break of 8 seconds between encoding and retrieval blocks. During the retrieval blocks, all pairs (18 pairs for children, 22 pairs for young adults) were presented again in a pseudo-randomized order, with the constraint that the final encoding pairs were not among the first ones presented during retrieval. This ensured that the encoding event was cleared from short-term memory. The task of the participants was to

indicate whether the left or right fribble won during the encoding phase. In addition, they were asked to indicate confidence in their judgment on a confidence scale (uncertain, rather certain, certain). Before scanning, the participants were explicitly told that only the outcome of the immediately preceding encoding block was relevant for the retrieval phase; they were also told that responses based solely on the hierarchy would lead to chance performance and were motivated to perform as highly as possible. Performance feedback was provided on the screen after each block. After finishing the memory task, the participants were taken out of the scanner and were again asked to arrange all 36 exemplars by hierarchy level. This last procedure was to examine whether the participants had modified their knowledge of the hierarchy during the memory task. Day 2 took between 90 and 120 minutes.

fMRI Data Acquisition

Whole-brain MRI data were collected with a Siemens 3T Trio Magnetom. Functional data were acquired using a single shot echo-planar imaging (EPI) gradient echo sequence that is sensitive to BOLD contrast (direction = interleaved, FOV = 216 mm, TR = 2500 ms, TE = 30 ms, number of slices = 45, slice thickness = 2.5mm, matrix = 72 x 72, voxel size = 3 x 3 x 2.5, distance factor = 20%). After being aligned to the genu and splenium of the corpus callosum, the scanning window was tilted vertically by 15 degrees to help attenuate signal dropout in orbitofrontal regions (Deichmann, Gottfried, Hutton, & Turner, 2003; Weiskopf, Hutton, Josephs, & Deichmann, 2006). To ensure that a steady-state condition was reached, the first three scans of each run were discarded. To optimize statistical efficiency of our rapid event-related design, jittered fixation periods were used, which were optimized with Optseq 2 (Dale, 1999). To estimate and correct for geometric distortion and signal loss in the EPI, a 53-second fieldmap sequence was also measured. Structural data were acquired using a T1-weighted 3D magnetization-prepared rapid gradient echo sequence (TR 2500 ms, TE 2500 ms, sagittal orientation, spatial resolution 1 x 1 x 1 mm).

Statistical Analysis: Behavioural Data

Learning and memory performance was analyzed using R (R Core Team, 2014). For the two learning phases on Day 1 and for the repetition of the second learning phase at the beginning of Day 2, average number of attempts needed to reach performance criterion ($\geq 90\%$ correct within a block) were calculated separately for the two age groups. In addition, mean performance was calculated for the last block of the repeated learning phase on Day 2 to assess potential differences in schema strength before the memory task. Two-sample t-tests were conducted to determine differences in learning performance between the two age groups. For the data of the memory task on Day 2, mean classification performance was calculated for the encoding phase, and percentage of correctly remembered winners (hits) was calculated for the retrieval phase; this was done separately for the congruent and incongruent condition. For the retrieval phase, two one-sample t-tests were conducted for both age groups to determine whether memory performance was above chance (50%) in both conditions. Most importantly, mixed ANOVAs were conducted to determine whether memory performance (accuracy and reaction time) significantly differed between age groups, conditions, and their interactions. The α level was set at 0.05 throughout the behavioural analyses.

Statistical Analysis: fMRI Data

The fMRI data were preprocessed and analyzed using FEAT in FSL (FMRIB's Software Library, <http://www.fmrib.ox.ac.uk/fsl>; Smith, Jenkinson, & Woolrich, 2004). First, non-brain tissue was removed. Second, slice time (interleaved) and motion correction (using MCFLIRT) were applied. Third, the data were spatially smoothed using a 5-mm full-width half-maximum Gaussian filter. The FILM pre-whitening technique was used to account for the intrinsic temporal autocorrelation of BOLD imaging. Using FUGUE in FSL, distortions in the EPI sequences due to magnetic field inhomogeneities, which were measured with a fieldmap image, were compensated by geometrically unwarping the EPI images and ignoring areas with signal loss greater than 10%. Low-frequency artifacts were removed by applying a high-pass temporal filter (Gaussian-weighted straight-line fitting, $\sigma = 50$ s). Registration

to the subject-specific high-resolution T1 image and subsequently to the standard image (Montreal Neurological Institute, MNI, 2 mm isotropic voxels) was carried out using FLIRT (Jenkinson & Smith, 2001).

After preprocessing, individual time series were modeled separately for the different runs with different regressors for each of the four events (successfully retrieved congruent and incongruent events, unsuccessfully retrieved congruent and incongruent events), separately for the encoding and retrieval phases. We focused exclusively on the retrieval phase for the following fMRI analyses. Encoding classification was not taken into account for sorting of trials in the retrieval analyses. Since the focus of this study was to examine the effects of schema congruency on memory retrieval, our paradigm, with its rapidly alternating encoding and retrieval blocks, is suboptimal for the identification of subsequent memory effects at encoding. Those are typically examined in experimental designs with more trials and substantially longer study–test delays (cf. Stark & Squire, 2000; Uncapher & Rugg, 2005). Therefore, the fMRI analyses reported here are restricted to the retrieval phase. Event length for retrieval was 4.75 seconds, with the onset being linked to the presentation of the fribbles. The regressors were generated by convolving the impulse function related to the onsets and lengths of the events of interest with a gamma hemodynamic response function (HRF). Contrast images were computed for each run per subject. In a second step, the different runs within each participant were spatially normalized, transformed into MNI space, and combined using a higher-level within-subjects fixed effects analysis. In a third step, an across-subject analysis was carried out using a mixed-effects model (FLAME Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004).

Based on our predictions and findings in young adults (Brod et al., 2015), our main analyses focused on *a priori* defined anatomical regions of interest (ROI) in the bilateral mPFC, IPFC, and HC. Anatomical ROIs were defined based on FSL’s Harvard Oxford Structural Atlas. The mPFC ROI comprised the bilateral frontal medial cortex and the

paracingulate gyrus area. The IPFC ROI encompassed the bilateral superior, middle, and inferior frontal gyri. The HC ROI consisted of the bilateral HC. In complementary whole-brain analyses, we explored whether additional brain regions were involved in successfully retrieving schema-related episodes in the children and in the young adults. Across both whole-brain and ROI analyses, Z-statistic images were thresholded voxel-wise at a threshold of $z > 2.3$ (corresponding to an α level of 0.01). Multiple comparison correction was performed using the 3DClustSim program of the AFNI software package (http://afni.nimh.nih.gov/pub/dist/doc/program_help/3dClustSim.html). Monte Carlo simulations were conducted using the three ROIs and a whole-brain mask. First, smoothness was estimated on the residual time series image using AFNI's 3dFWHMx (http://afni.nimh.nih.gov/pub/dist/doc/program_help/3dFWHMx.html). Using the estimated smoothness, we conducted a simulation with 10 000 runs that yielded minimum cluster extents of 48 (mPFC), 76 (IPFC), 18 (HC), and 113 (whole-brain) with a family-wise error rate of $p < .05$. The mPFC mask contained 11357 voxels (2 x 2 x 2 mm), the IPFC mask contained 37663 voxels, and the HC mask contained 1173 voxels.

To determine age differences in the recruitment of brain areas that are important for the successful retrieval of congruent and incongruent events respectively, three sets of analyses were conducted for each of the two events using mixed-effects models (Woolrich et al., 2004): first, to characterize brain areas important for successful memory retrieval in the child and the young adult groups separately, successful retrieval (SR) > unsuccessful retrieval (UR) contrasts were computed. Second, to confirm the age-differential patterns of activation in those brain areas, we performed an unpaired two-group test that yields brain regions where differences between SR and UR trials were larger in the young adults than in the children, and vice versa ($[(SR > UR)_{\text{young adults}}] > [(SR > UR)_{\text{children}}]$, and vice versa). To constrain our age-difference analyses to task-positive regions, we used the combined above-threshold activation of children and young adults from the successful retrieval contrasts as a mask for this step.

Third, given that the child group covered an age range of five years, we also examined whether there were age-related increases in the recruitment of mPFC and IPFC areas within this group. We did so in separate mixed effects models consisting of children only, where, in addition to the mean regressor, we included age (in months, z-transformed) as a parametric regressor of interest. All other settings remained identical to the age-difference analyses. To control for individual differences in memory performance, z-transformed hit rates were also entered in the models as covariates. All analyses above were performed separately for congruent and incongruent SR > UR contrasts. If not otherwise stated, results were derived from the anatomical ROI's.

Finally, in an exploratory analysis complementing the separate analyses of schema-congruent and schema-incongruent events, we also performed a full factorial analysis directly testing for the interaction of age group (children, young adults), memory (SR, UR), and congruency (congruent, incongruent). Thus, we sought to test whether there were age-related differences in the memory x congruency interaction. For these analyses, we focused on those brain regions that showed age differences in the SR > UR contrasts for either the schema-congruent or the schema-incongruent condition, that is, the mPFC and the HC. We again used our a priori defined anatomical ROIs and performed the full factorial analysis in the same way as the previous analyses. That is, we used mixed-effects models in FSL (Woolrich et al., 2004) to delineate whether young adults showed greater mPFC activation for schema-congruent events as compared to schema-incongruent events than children (i.e., Young Adults $[(SR > UR)_{Congruent} > (SR > UR)_{Incongruent}] > Children [(SR > UR)_{Congruent} > (SR > UR)_{Incongruent}]$), and whether children showed greater HC activation for schema-incongruent events as compared to schema-congruent events than young adults (i.e., Children $[(SR > UR)_{Incongruent} > (SR > UR)_{Congruent}] > Young Adults [(SR > UR)_{Incongruent} > (SR > UR)_{Congruent}]$). To account for the relatively small group sizes and the corresponding low statistical power of the full-factorial design, statistical significance was evaluated at a slightly more lenient voxel

threshold of $Z=1.96$ (corresponding to an α level of 0.025). To facilitate the interpretation of the underlying patterns of the three-way interactions, percent signal change (PSC) was extracted from each cluster and separate follow-up tests for each age group were performed on the PSC values using within-subjects ANOVAs. Psychophysiological interaction (PPI) analyses were performed using FSL to assess age-related differences in connectivity. Based on our finding of stronger connectivity between striatum and DLPFC for incongruent events than for congruent events in young adults (Brod et al., 2015), we wished to test whether this connectivity pattern was similarly present in children. In addition, based on findings of differential HC–mPFC connectivity for schema-congruent vs. schema-incongruent events (Bein, Reggev, & Maril, 2014; van Kesteren et al., 2013), we tested for age-differences in HC–mPFC connectivity. Seed regions were defined in the left striatum and left hippocampus by building a sphere with a radius of 5 mm around peak activation in the SR > UR contrast across the two conditions and age groups, thus avoiding age group or condition biases. To determine age differences in connectivity that are modulated by congruency, we tested whether differences between incongruent SR and congruent SR trials were larger in the young adults than in the children, and vice versa (Young Adults [$SR_{\text{Congruent}} > SR_{\text{Incongruent}}$] > Children [$SR_{\text{Congruent}} > SR_{\text{Incongruent}}$]), and vice versa). The same mPFC and IPFC anatomical masks as used for the analyses above were used as ROI for the connectivity analyses.

Results

Learning and Memory Performance

In the first learning phase, children needed 3 blocks on average (range 2–5) to acquire the hierarchy of a group of 12 fribbles (performance criterion $\geq .90$), which is marginally more than the young adults who needed 2.6 blocks on average (range: 2–5; $t(47) = 1.99$, $p = .05$). In the second learning phase, during which pairs of all 36 fribbles were presented together within one block, children needed 2 blocks on average (range 1–3) to perform at 90%, significantly more than the young adults who needed 1.3 blocks on average (range: 1–

5; $t(47) = 3.19, p < .01$). During the repetition of the second learning phase on Day 2, however, children needed only on average 1.1 blocks (range: 1-2), which is comparable to the young adults ($t(47) = 0.75, p = .46$), who needed on average 1.3 blocks (range: 1 - 3). In addition, mean performance during the last learning phase was comparable between the two age groups (children: 93%, young adults: 94%, $t(47) = -.36, p > .7$). This supports the assumption that knowledge structures were similarly well established in children and young adult groups at the end of the learning phase.

During encoding, children classified 88% (SD = 7) of the outcomes correctly as congruent/incongruent, whereas the young adults participants classified 93% (SD = 4.5) correctly ($t(47) = 2.76, p = .01$). This difference is mostly due to some children who more often failed to answer within the response time limit of 2 sec during encoding, and not due to children responding less correctly. We confirmed this in an analysis looking only at the trials for which participants provided an answer. When doing this, accuracy differences between children and young adults were not reliable, and both age groups showed ceiling performance (Children: 95% correct; Adults: 96% correct, $t(47) = 1.08, p = .28$). Memory performance (retrieval phase) was above chance for both the congruent and incongruent condition in both age groups (for an overview of the results, see Table 1). Testing for differences between groups and conditions, the F-test revealed a main effect of congruency ($F(1,47) = 23.63, p < .001$), indicating better memory performance for congruent than for incongruent events, a non-significant trend towards better overall memory performance in the young adults ($F(1,47) = 3.85, p = .06$), and no interaction ($F(1,47) = .03, p = .85$). For the reaction times during memory retrieval, the F-test revealed no main effect of congruency ($F(1,47) = 1.58, p = .22$), a non-significant trend towards slower reaction times in the young adults ($F(1,47) = 3.62, p = .06$), and no interaction between congruency and age ($F(1,47) = 1.24, p = .27$). Thus, typically observed age differences in behavioural performance could be strongly attenuated.

fMRI Results

Age-related differences in the neural signature of remembering schema-congruent events

Given relatively comparable memory performance, the first question that we wished to address was whether children and young adults use similar brain regions to successfully remember congruent events. To this end, SR > UR contrasts for congruent trials were first computed for both age groups separately, thresholded voxel-wise at $p < .01$ and corrected for multiple comparisons using *a priori* defined anatomical masks of the HC, the IPFC, and the mPFC. As can be seen in Figure 2 (see also Table 2 for whole-brain results), while young adults (green) showed activation in both mPFC and left HC, children (yellow) showed significant activation in both the left and the right HC but not in the mPFC. To test directly for age-related differences in activation patterns, we examined differences between SR and UR trials that were larger in the young adults than in the children, and vice versa. To do this, above-threshold activation across the two groups was used as a pre-thresholded mask for the age comparison. As can be seen in Figure 2, young adults (blue) indeed showed stronger mPFC activation than children, but children and young adults did not significantly differ in hippocampal activation. Given the rather broad age range among the children, we conducted an additional between-subject analysis to examine age-related differences in the neural correlates of successfully retrieving congruent events (SR > UR contrast), statistically controlling for memory performance. This analysis revealed an age-related increase in brain activation within the children group for successfully remembered congruent events in an mPFC cluster (383 voxels) that overlapped with the mPFC cluster observed in the young adults ($r = .50$, Figure 2).

Age-related differences in the neural signature of remembering schema-incongruent events

To address whether there were age-related differences in the use of brain regions to remember incongruent events successfully, SR > UR contrasts of incongruent trials were computed, again initially for the two age groups separately at a voxel-wise threshold of $p < .01$ (corrected for multiple comparisons). As can be seen in Figure 3 (see also Table 2 for

whole-brain results), this contrast revealed bilateral IPFC activation in both children (yellow) and young adults (green), but bilateral HC activation in the children group only. To test for the statistical significance of this pattern, we again examined differences between SR and UR trials that were larger in the young adults than in the children, and vice versa. As before, we used the combined activation mask of the two groups as a pre-thresholded mask to focus the search for age differences in task-positive regions. As can be seen in Figure 3 (left), children (red) showed stronger right HC activation than young adults. Contrary to our expectation, however, young adults did not show significantly greater IPFC activation than children. The complementary whole-brain analysis revealed that, in addition to PFC and HC, both age groups showed activation in bilateral striatum and posterior parietal areas (see Table 2), but there was no age difference in these regions. Finally, we tested whether there were age-related changes in the neural correlates of successfully retrieving incongruent events across middle childhood. The correlational analysis in the children group revealed no association between age and brain activation for incongruent events (SR > UR).

Age-related differences in the differential neural signatures of retrieving schema-congruent versus schema-incongruent events

Building upon the above findings of age-related differences in the retrieval of schema-congruent and schema-incongruent events in mPFC and HC, respectively, we also were interested in whether there were age-related differences in the memory (SR, UR) x congruency (congruent, incongruent) interaction in these regions. In an exploratory analysis (lowered voxel-wise threshold: $p < .025$, corrected for multiple comparisons), we, thus, tested whether young adults showed greater mPFC activation for schema-congruent events as compared to schema-incongruent events than children, and whether children showed greater HC activation for schema-incongruent events as compared to schema-congruent events than adults, using our pre-defined anatomical ROIs. Two above-threshold clusters were detected: one in the mPFC, and one in the left HC. To help interpretation of the underlying patterns of

the three-way interactions, percent signal change (PSC) was extracted from each cluster (see Figure 4) and follow-up within-subject ANOVAs were performed on the PSC values.

As can be seen in Panel A of Figure 4, we detected a cluster in the mPFC (peak [MNI 8, 36, -8], $Z = 2.84$, 57 voxels) that displayed age-related differences in the memory x congruency interaction. This cluster overlaps with the one that displayed stronger activation in young adults than children for SR vs. UR schema-congruent events (Figure 2), and is adjacent to the memory x congruency interaction cluster for young adults reported in Brod et al. (2015). Follow-up statistical tests indicated a trend towards a memory x congruency interaction in the young adults ($F(1,23) = 2.7$, $p = .11$), who showed greater mPFC activity for SR vs. UR congruent events and lower mPFC activity for SR vs. UR incongruent events. In children, the interaction did not differ reliably from zero ($F(1,24) = 0.07$, $p = .79$).

As can be seen in Panel B of Figure 4, another cluster was detected in the left hippocampus (peak [MNI -24, -16, -16], $Z = 2.58$, 26 voxels), contralateral to the left HC cluster that displayed stronger activation in the children for SR vs. UR schema-incongruent events. Follow-up tests indicated a memory x congruency interaction in young adults ($F(1,23) = 7.2$, $p = .01$), who showed greater left HC activation for SR vs. UR congruent events than for incongruent events. In the children group, the corresponding interaction was not statistically significant ($F(1,24) = 0.8$, $p = .38$). This finding, in conjunction with children's strong and bilateral HC activation for both remembered congruent and incongruent events (see Table 2), indicates that HC contributions to successful memory retrieval were prominent in children, independent of congruency.

Age-related differences in connectivity for remembering schema-congruent vs. schema-incongruent events

Based on our earlier findings of stronger connectivity between striatum and DLPFC for incongruent SR events as compared to congruent SR events in young adults (Brod et al.,

2015), we examined whether this connectivity pattern was similarly present in the group of children (thresholded voxel-wise at $p < .01$, corrected for multiple comparisons). In addition, based on findings of differential HC–mPFC connectivity for schema-congruent vs. schema-incongruent events (Bein, Reggev, & Maril, 2014; van Kesteren et al., 2013), we tested for age-differences in HC–mPFC connectivity. Seed regions (5 mm spheric mask) were defined based on peak activation in the SR > UR contrast across the two conditions and age groups. For incongruent as compared to congruent SR events, connectivity between striatum and IPFC was significantly stronger in the young adults than in the children (see Figure 5), who did not show any reliable connectivity differences between congruent and incongruent events. No cluster was found for the opposite contrast. Also, HC–mPFC connectivity did not differ between the two conditions, nor were there any age-related differences.

Discussion

The results of this study provide new insights into the differences in neural correlates of schema-related memory processing between children aged 8–12 years and young adults. By experimentally inducing task-relevant knowledge, we avoided the common confound between knowledge and age. When successfully retrieving schema-*congruent* events, children showed less mPFC activity than adults, and their mPFC activity during successful retrieval correlated positively with their age. When successfully retrieving schema-*incongruent* events, children showed stronger HC activation as well as weaker connectivity between the striatum and the dorsolateral PFC than adults. These findings were corroborated by an exploratory full-factorial analysis investigating age differences in the retrieval of schema-congruent versus schema-incongruent events, comparing the two conditions directly. Consistent with the findings of the separate analyses, two clusters, one in the medial PFC, one in the HC, were identified that exhibited a memory x congruency x age group interaction. For the HC, a left HC cluster was detected that displayed a memory x congruency interaction in adults, but not in children, who showed strong memory (SR > UR) activation under both conditions. The

present results indicate an age-related increase in the relative importance of prefrontal brain regions for the retrieval of schema-related events, and a strong and general contribution of the hippocampus to memory retrieval in children. Overall, these findings are in good agreement with the two-component model of episodic memory development proposed by Shing et al. (2010).

The important role of the mPFC in the successful retrieval of schema-congruent events observed in the present study is in good agreement with available evidence on the importance of the mPFC for memory. Previous studies in humans and animals have shown that mPFC is involved in memory integration and generalization (DeVito, Lykken, Kanter, & Eichenbaum, 2010; Kosciak & Tranel, 2012; Kumaran & Maguire, 2009; for reviews, see Schlichting & Preston, 2015; van Kesteren et al., 2012). While generalization is highly important for behavioural flexibility in everyday life (Schlichting & Preston, 2015), overgeneralization can lead to an increased number of false memories. A recent study (Warren, Jones, Duff, & Tranel, 2014) revealed that patients with focal mPFC lesions show less false memories in the Deese-Roediger-McDermott (DRM) paradigm than healthy controls. In the DRM paradigm, which reliably induces false memories in healthy adults by presenting words in the retrieval phase that are highly semantically related to words of the study phase, false memories are a result of semantic memory generalization based on schemas. Interestingly, most developmental studies using the DRM paradigm report that children show less memory intrusions than adults, thus suggesting an increase in memory generalization across childhood (e.g. Brainerd & Reyna, 2002; Paz-Alonso, Ghetti, Donohue, Goodman, & Bunge, 2008; but see Ghetti, Qin, & Goodman, 2002). At the same time, other studies report that the structure and function of mPFC structure continues to mature into adolescence (Barbalat, Bazargani, & Blakemore, 2013; Shaw et al., 2008; Supekar et al., 2010). In line with these findings, the results from the present study suggest that the increasing engagement of mPFC during

retrieval of schema-*congruent* events in middle childhood is associated with neural mechanisms that enhance schematic memory generalization.

Remembering incongruent events resembles a source memory situation during which the specific encoding context has to be retrieved and a specific task-relevant detail, in our case that the generally “slower” fribble won the contest, has to be recovered. Furthermore, interference from the schema has to be overcome (Brod et al., 2013). Accordingly, in young adults, successful memory for incongruent as compared to congruent events was associated with enhanced activity in the DLPFC, striatum, and posterior parietal cortex (Brod et al., 2015), which are areas typically found in source memory tasks (Han et al., 2012; Hayes, Buchler, Stokes, Kragel, & Cabeza, 2011; Henson, Rugg, Shallice, Josephs, & Dolan, 1999; Ranganath, Johnson, & D’Esposito, 2000). In addition, young adults showed enhanced connectivity between striatum and DLPFC for remembered schema-incongruent events, which fits the notion that frontostriatal circuits matter in situations that pose high control demands on memory retrieval, for example when interference suppression is required (Scimeca & Badre, 2012). Research into the development of source memory has revealed steep performance increases throughout childhood (Cycowicz, Friedman, & Snodgrass, 2001; de Chastelaine, Friedman, & Cycowicz, 2007; Schneider & Pressley, 1999), which have been linked to the maturation of IPFC-mediated strategic and cognitive control processes (Cycowicz et al., 2001; Czernochowski, Mecklinger, & Johansson, 2009; Shing et al., 2010). However, fMRI evidence for this line of reasoning has been sparse. We expected children to show less DLPFC activation when successfully retrieving incongruent events than young adults. However, young adults did not show reliably greater DLPFC activation than children in this condition, nor were there any apparent age group differences in the 3-way interaction. Rather, we observed greater connectivity between striatum and DLPFC for remembered incongruent vs. congruent events in young adults than in children. Taken together, findings regarding age-differences in DLPFC contributions to schema-incongruent memory retrieval

remain inconclusive. The expected decreased DLPFC connectivity in children stands in contrast to the unexpected null finding concerning age-differences in DLPFC activity. However, age-related differences for schema-incongruent events became apparent in the HC. Children showed strong, bilateral HC activity, whereas the young adults showed left-lateralized HC activation, which resulted in significant age-related differences in the right HC. Furthermore, a follow-up ANOVA revealed a three-way interaction (see Panel B in Figure 4): In young adults, activation differences favoring SR over UR events were greater for congruent than for incongruent events. In contrast, children showed strong activation differences favoring SR over UR that did not interact reliably with condition (see also Table 2). These findings provide evidence for a strong, general, and bilateral contribution of the HC to successful memory retrieval during middle childhood. This finding is consistent with the proposition that the HC plays a particularly prominent role for episodic memory performance during childhood (Shing et al., 2010; Ofen, 2012). In fact, our findings provide tentative evidence that this reliance is enhanced compared to the young adults. These data are consistent with an earlier finding of an age-related HC reduction in the subsequent memory effect for encoding verbal material (Maril et al., 2010). One explanation for stronger HC-reliance in children might be the immaturity of other brain regions and of the connections to these brain regions, such as fronto-parietal regions (see DeMaster et al., 2013; Ghetti & Bunge, 2012), which later in development may facilitate processing within the HC. An alternative, but not mutually exclusive explanation might be an age-related change of processing in the HC, either in localization or in type (e.g., perceptual vs. semantic, see DeMaster & Ghetti, 2013).

Overall, our findings are in good agreement with the two-component model of episodic memory development (Shing et al., 2010, 2008). The two-component model distinguishes an early developing associative component, assumed to perform associative binding, from a later developing strategic component, which refers to memory control

operations. The development of the associative component is seen as being closely linked to the HC, which is assumed to be already highly functional by middle childhood (Ofen et al., 2007; Ofen et al., 2012). The development of the strategic component is linked to the lateral PFC, which displays a protracted functional development (Shaw et al., 2008; Sowell et al., 2003). These maturational differences are revealed by the finding that children's worse episodic memory performance can be linked to lower levels of strategy use, but performance differences are reduced when children's and young adults' strategy use is equated via training (Shing et al., 2008). From this point of view, the enhanced activation of the HC for the retrieval of incongruent events seen in the children group may reflect their strong reliance on the already well-functioning associative component.

The present experimental paradigm was designed with the goal to minimize age differences in task-relevant knowledge, and, indeed, children and young adults did not differ in performance on the final learning block. However, for the congruency decision during the encoding phase, children responded less often (in time) than young adults, possibly because the question phase was too short for some of them, which resulted in lower accuracy for the children. This is a potential concern because encoding classification was not taken into account in the fMRI analysis because overall trial numbers were low. When trials with no responses were excluded, performance was at ceiling in both age groups. This suggests that both groups had comparable knowledge and perceived congruency similarly. This strength of the present study makes it hard to be compared directly to age-comparative studies that did not address or control for this important confound. In such studies, memory for congruent events was typically found to increase more steeply with age than memory for incongruent events (Stangor & McMillan, 1992). The extent to which this pattern reflects age-related increases in knowledge is unknown (Brod et al., 2013). We acknowledge, however, that the gain in internal validity achieved by experimental knowledge induction comes at the price of

a potential loss in external validity. Future studies may profit from inducing bodies of knowledge that more closely mimic the complexity of real-world knowledge.

Another limitation of the present study is that analyses are confined to the retrieval phase of episodic memory. The present version of the paradigm does not allow the identification of subsequent memory effects, which would require more trials and substantially longer study–test delays (cf. Stark & Squire, 2000; Uncapher & Rugg, 2005). It would be of great interest to investigate age differences in subsequent memory effects on the basis of experimentally induced knowledge.

Summary and Outlook

This study documents, for the first time, differences between children and young adults in neural activation patterns associated with the retrieval of memories that are congruent or incongruent with a schema. Study results suggest an age-related increase in the relative importance of prefrontal brain regions for the retrieval of schema-related events, and a strong and general contribution of the hippocampus to memory retrieval in children. Clearly, these results need to be replicated and extended, preferably with longitudinal designs that also address individual differences in memory development (Schneider, 2011). Methodologically, the present study demonstrates that the experimental induction of task-relevant knowledge is a useful tool to arrive at a more valid assessment of age differences in the effects of schemas on memory retrieval. Future renditions of this approach may seek to implement richer knowledge structures that bear greater resemblance to real-world knowledge.

References

- Barbalat, G., Bazargani, N., & Blakemore, S.-J. (2013). The Influence of Prior Expectations on Emotional Face Perception in Adolescence. *Cerebral Cortex*, *23*(7), 1542–1551.
- Bein, O., Reggev, N., & Maril, A. (2014). Prior knowledge influences on hippocampus and medial prefrontal cortex interactions in subsequent memory. *Neuropsychologia*, *64*, 320–330.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–96.
- Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-Trace Theory and False Memory. *Current Directions in Psychological Science*, *11*(5), 164–169.
- Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 717–726.
- Brod, G., Lindenberger, U., Werkle-Bergner, M., & Shing, Y. L. (2015). Differences in the neural signature of remembering schema-congruent and schema-incongruent events. *NeuroImage*. *117*, 358–366.
- Brod, G., Werkle-Bergner, M., & Shing, Y. (2013). The influence of prior knowledge on memory: a developmental cognitive neuroscience perspective. *Frontiers in Behavioral Neuroscience*, *7*, 139.
- Chai, X. J., Ofen, N., Jacobs, L. F., & Gabrieli, J. D. E. (2010). Scene complexity: influence on perception, memory, and development in the medial temporal lobe. *Frontiers in Human Neuroscience*, *4*, 21.
- Chi, M. T. H. (1978). Knowledge Structures and Memory Development. *Children's Thinking: What Develops?*
- Cycowicz, Y. M., Friedman, D., & Snodgrass, J. G. (2001). Remembering the color of objects: An ERP investigation of source memory. *Cerebral Cortex*, *11*(4), 322–334.
- Czernochowski, D., Mecklinger, A., & Johansson, M. (2009). Age-related changes in the control of episodic retrieval: an ERP study of recognition memory in children and adults. *Developmental Science*, *12*(6), 1026–1040.
- Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, *8*(2-3), 109–114.
- de Chastelaine, M., Friedman, D., & Cycowicz, Y. M. (2007). The development of control processes supporting source memory discrimination as revealed by event-related potentials. *Journal of Cognitive Neuroscience*, *19*(8), 1286–1301.
- Deichmann, R., Gottfried, J. A., Hutton, C., & Turner, R. (2003). Optimized EPI for fMRI

- studies of the orbitofrontal cortex. *Neuroimage*, *19*(2), 430–441.
- DeMaster, D. M., & Ghetti, S. (2013). Developmental differences in hippocampal and cortical contributions to episodic retrieval. *Cortex*, *49*(6), 1482–1493.
- Demaster, D., Pathman, T., Lee, J. K., & Ghetti, S. (2013). Structural Development of the Hippocampus and Episodic Memory: Developmental Differences Along the Anterior/Posterior Axis. *Cerebral Cortex*, 3036–3045.
- DeVito, L. M., Lykken, C., Kanter, B. R., & Eichenbaum, H. (2010). Prefrontal cortex: role in acquisition of overlapping associations and transitive inference. *Learning & Memory*, *17*(3), 161–7.
- Ghetti, S., & Bunge, S. A. (2012). Neural changes underlying the development of episodic memory during middle childhood. *Developmental Cognitive Neuroscience*, *2*(4), 381–95.
- Ghetti, S., DeMaster, D. M., Yonelinas, A. P., & Bunge, S. A. (2010). Developmental Differences in Medial Temporal Lobe Function during Memory Encoding. *Journal of Neuroscience*, *30*(28), 9548–9556.
- Ghetti, S., Qin, J., & Goodman, G. S. (2002). False memories in children and adults: Age, distinctiveness, and subjective experience. *Developmental Psychology*, *38*(5), 705–718.
- Ghosh, V. E., Moscovitch, M., Melo Colella, B., & Gilboa, A. (2014). Schema Representation in Patients with Ventromedial PFC Lesions. *Journal of Neuroscience*, *34*(36), 12057–12070.
- Gogtay, N., Nugent, T. F., Herman, D. H., Ordonez, A., Greenstein, D., Hayashi, K. M., ... Thompson, P. M. (2006). Dynamic mapping of normal human hippocampal development. *Hippocampus*, *16*(8), 664–672.
- Golarai, G., Ghahremani, D. G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J. L., Gabrieli, J. D. E., & Grill-Spector, K. (2007). Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nature Neuroscience*, *10*(4), 512–22.
- Han, S., O'Connor, A. R., Eslick, A. N., & Dobbins, I. G. (2012). The role of left ventrolateral prefrontal cortex during episodic decisions: semantic elaboration or resolution of episodic interference? *Journal of Cognitive Neuroscience*, *24*(1), 223–234.
- Hayes, S. M., Buchler, N., Stokes, J., Kragel, J., & Cabeza, R. (2011). Neural correlates of confidence during item recognition and source memory retrieval: evidence for both dual-process and strength memory theories. *Journal of Cognitive Neuroscience*, *23*(12), 3959–3971.
- Henson, R. N. A., Rugg, M. D., Shallice, T., Josephs, O., & Dolan, R. J. (1999). Recollection and familiarity in recognition memory: an event-related functional magnetic resonance imaging study. *The Journal of Neuroscience*, *19*(10), 3962–3972.
- Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine

- registration of brain images. *Medical Image Analysis*, 5(2), 143–156.
- Koscik, T. R., & Tranel, D. (2012). The Human Ventromedial Prefrontal Cortex Is Critical for Transitive Inference. *Journal of Cognitive Neuroscience*, 24(5), 1191–1204.
- Kumaran, D., & Maguire, E. A. (2009). Novelty signals: A window into hippocampal information processing. *Trends in Cognitive Sciences*, 13(2), 47–54.
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics:(de) constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920–933.
- Maril, A., Avital, R., Reggev, N., Zuckerman, M., Sadeh, T., Ben Sira, L., & Livneh, N. (2011). Event congruency and episodic encoding: a developmental fMRI study. *Neuropsychologia*, 49(11), 3036–45.
- Maril, A., Davis, P. E., Koo, J. J., Reggev, N., Zuckerman, M., Ehrenfeld, L., ... Rivkin, M. J. (2010). Developmental fMRI study of episodic verbal memory encoding in children. *Neurology*, 75(23), 2110–6.
- Navawongse, R., & Eichenbaum, H. (2013). Distinct pathways for rule-based retrieval and spatial mapping of memory representations in hippocampal neurons. *The Journal of Neuroscience*, 33(3), 1002–1013.
- Ofen, N. (2012). The development of neural correlates for memory formation. *Neuroscience and Biobehavioral Reviews*.
- Ofen, N., Chai, X. J., Schuil, K. D. I., Whitfield-Gabrieli, S., & Gabrieli, J. D. E. (2012). The development of brain systems associated with successful memory retrieval of scenes. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 32(29), 10012–20.
- Ofen, N., Kao, Y.-C., Sokol-Hessner, P., Kim, H., Whitfield-Gabrieli, S., & Gabrieli, J. D. E. (2007). Development of the declarative memory system in the human brain. *Nature Neuroscience*, 10(9), 1198–205.
- Paz-Alonso, P. M., Ghetti, S., Donohue, S. E., Goodman, G. S., & Bunge, S. a. (2008). Neurodevelopmental correlates of true and false recognition. *Cerebral Cortex (New York, N.Y. : 1991)*, 18(9), 2208–16.
- Piaget, J. (1929). *The child's conception of the world*. Lanham, ML: Rowman & Littlefield.
- Preston, A., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, 23(17), R764–73.
- R Core Team. (2014). R: A Language and Environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
- Ranganath, C., Johnson, M. K., & D'Esposito, M. (2000). Left anterior prefrontal activation increases with demands to recall specific perceptual information. *The Journal of Neuroscience*, 20(22), RC108.

- Schlichting, M. L., & Preston, A. R. (2015). Memory integration: neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, *1*, 1–8.
- Schneider, W. (2011). Memory development in childhood. In U. Goswami (Ed.), *The Wiley-Blackwell Handbook of Childhood Cognitive Development* (2nd ed., pp. 347–376). New York, NY: Wiley-Blackwell.
- Schneider, W., & Pressley, M. (1999). *Memory development between two and twenty*, 2nd edition. New York, NY: Springer.
- Scimeca, J. M., & Badre, D. (2012). Striatal Contributions to Declarative Memory Retrieval. *Neuron*, *75*(3), 380–392.
- Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., ... Wise, S. P. (2008). Neurodevelopmental Trajectories of the Human Cerebral Cortex. *Journal of Neuroscience*, *28*(14), 3586–3594.
- Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Müller, V., Li, S.-C., & Lindenberger, U. (2010). Episodic memory across the lifespan: The contributions of associative and strategic components. *Neuroscience & Biobehavioral Reviews*, *34*(7), 1080–1091.
- Shing, Y. L., Werkle-Bergner, M., Li, S.-C., & Lindenberger, U. (2008). Associative and strategic components of episodic memory: a life-span dissociation. *Journal of Experimental Psychology. General*, *137*(3), 495–513.
- Sloutsky, V. M., & Fisher, A. V. (2004). When development and learning decrease memory: Evidence against category-based induction in children. *Psychological Science*, *15*(8), 553–558.
- Smith, S., Jenkinson, M., & Woolrich, M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23*, 208–219.
- Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience*, *6*(3), 309–315.
- Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*, *111*(1), 42–61.
- Stark, C. E. L., & Squire, L. R. (2000). fMRI activity in the medial temporal lobe during recognition memory as a function of study-test interval. *Hippocampus*, *10*(3), 329–337.
- Supekar, K., Uddin, L. Q., Prater, K., Amin, H., Greicius, M. D., & Menon, V. (2010). Development of functional and structural connectivity within the default mode network in young children. *NeuroImage*, *52*(1), 290–301.
- Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. a, Wood, E. R., ... Morris, R. G. M. (2007). Schemas and memory consolidation. *Science*, *316*(5821), 76–82.
- Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., ... Morris, R. G. M.

- (2011). Schema-dependent gene activation and memory encoding in neocortex. *Science*, 333(6044), 891–5.
- Uncapher, M. R., & Rugg, M. D. (2005). Encoding and the durability of episodic memory: a functional magnetic resonance imaging study. *The Journal of Neuroscience*, 25(31), 7260–7267.
- van Kesteren, M. T. R., Beul, S. F., Takashima, A., Henson, R. N., Ruiter, D. J., & Fernández, G. (2013). Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. *Neuropsychologia*, 51(12), 2352–9.
- van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., & Fernández, G. (2010). Retrieval of associative information congruent with prior knowledge is related to increased medial prefrontal activity and connectivity. *The Journal of Neuroscience*, 30(47), 15888–94.
- van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, 35(4), 211–219.
- Warren, D. E., Jones, S. H., Duff, M. C., & Tranel, D. (2014). False recall is reduced by damage to the ventromedial prefrontal cortex: implications for understanding the neural correlates of schematic memory. *The Journal of Neuroscience*, 34(22), 7677–82.
- Weiskopf, N., Hutton, C., Josephs, O., & Deichmann, R. (2006). Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: A whole-brain analysis at 3 T and 1.5 T. *Neuroimage*, 33(2), 493–504.
- Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for fMRI group analysis using Bayesian inference. *Neuroimage*, 21(4), 1732–1747.

Tables

Table 1. Mean and standard deviations of behavioural results from each condition at retrieval, separately for young adults and children. Accuracy refers to the mean percentage of remembered trials in congruent and incongruent conditions. RT = reaction time.

Trial Type	Congruent		Incongruent	
	Accuracy	RT	Accuracy	RT
Young Adults	.72 (.09)	2412 (345)	.62 (.08)	2441 (334)
Children	.69 (.12)	2223 (398)	.58 (.11)	2225 (411)

Table 2. Regions observed in the successful retrieval (SR) > unsuccessful retrieval (UR) contrasts for the congruent and incongruent condition, respectively, separately for the children and young adults groups. MNI coordinates are reported.

Region	x	y	z	# voxels	Z-Max
Congruent SR-UR					
Children					
Left Hippocampus	-20	-8	-18	56	3.3
Right Hippocampus	16	-8	-18	35	3.2
Young Adults					
medial PFC	-12	42	0	1979	3.8
Left Hippocampus	-24	-16	-16	71	3.4
Incongruent SR-UR					
Children					
Bilateral Putamen / Caudate	20	12	-4	11844	4.2
Left Precentral Gyrus / Middle Frontal Gyrus	-44	-16	46	1969	3.7
Left Frontal Pole	-24	54	-8	1687	3.8
Right Central Opercular Cortex	56	-12	12	394	3.1
Left Inferior Frontal Gyrus	-44	8	26	270	3.4
Left Superior Parietal Lobe	-34	-54	46	245	2.7
Bilateral Cingulate Gyrus	-4	38	8	166	3
Bilateral Superior Frontal Gyrus	-4	14	58	163	3
Right Middle Frontal Gyrus	42	28	32	86	2.7
Left Hippocampus	-28	-18	-16	178	3.8
Right Hippocampus	30	-22	-14	113	3.3
Young Adults					
Precuneus / Lateral Occipital Cortex	2	-68	52	22962	4
Right Middle Frontal Gyrus / Frontal Pole	46	42	18	1914	3.6
Right Putamen / Caudate	12	16	4	1074	4.2
Right Central Opercular Cortex	54	-6	8	735	3.8
Right Lingual Gyrus	10	-64	-2	446	3.5

Figures

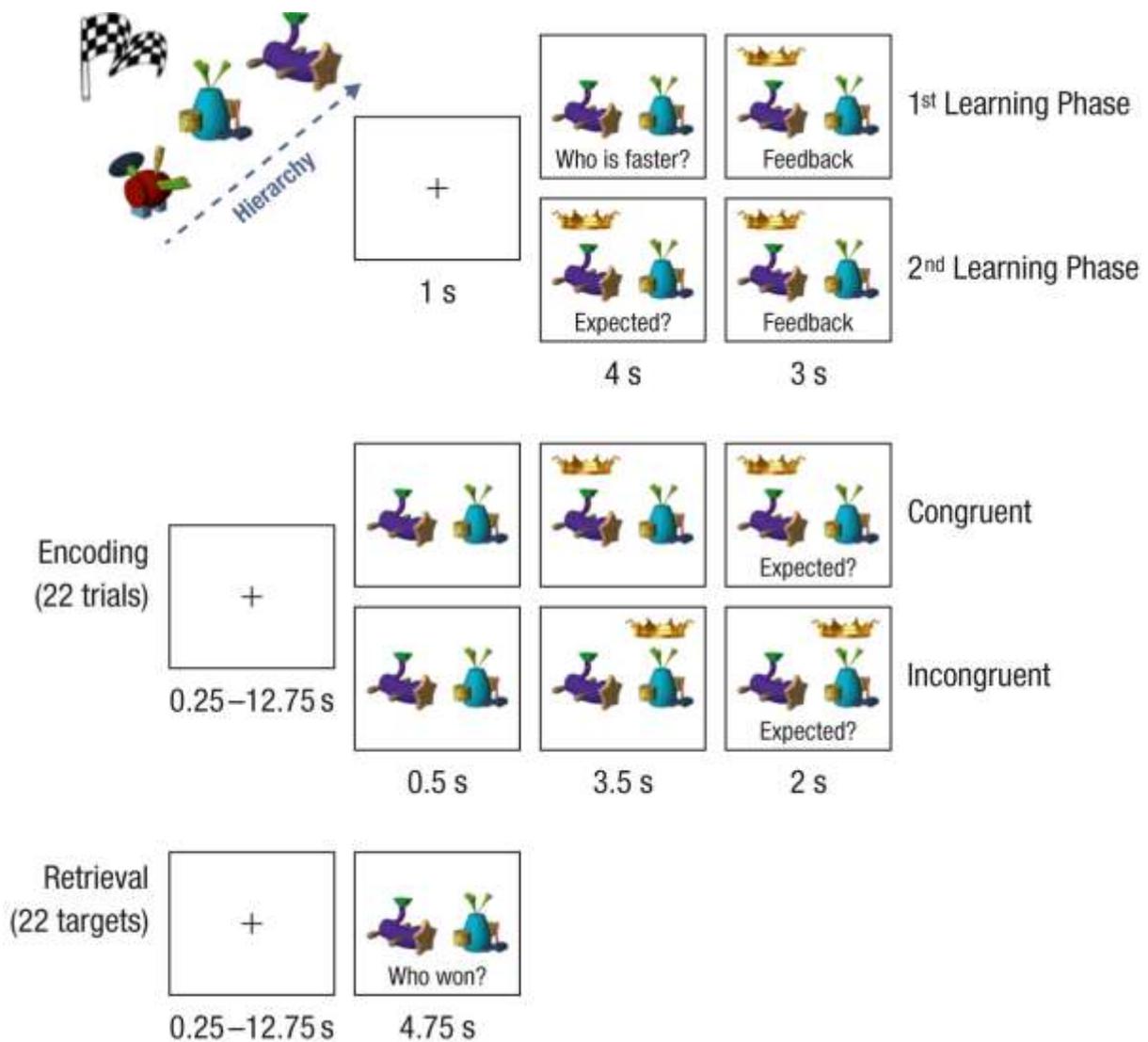


Figure 1. Experimental Design. On Day 1, participants acquired the hierarchy in the course of a trial-and-error learning task. In the first part, to facilitate learning, the 36 frubbles were subdivided into 3 x 12 exemplars and studied in different blocks. In the second part, all 36 exemplars were presented within one block, and participants were shown the outcome of the race (by a crown over the winner) and had to indicate whether they had expected the outcome or not based on the first learning part. In the memory task on Day 2, participants had to remember outcomes that were either congruent (50%) or incongruent (50%) with the hierarchy learned on Day 1. The MRI session consisted of alternately presented encoding and

retrieval blocks. During encoding, participants had to indicate whether they expected the outcome based on the hierarchy. During retrieval, they were presented with the same pairs again (in pseudorandom order) and had to indicate which of the two exemplars had won the race.

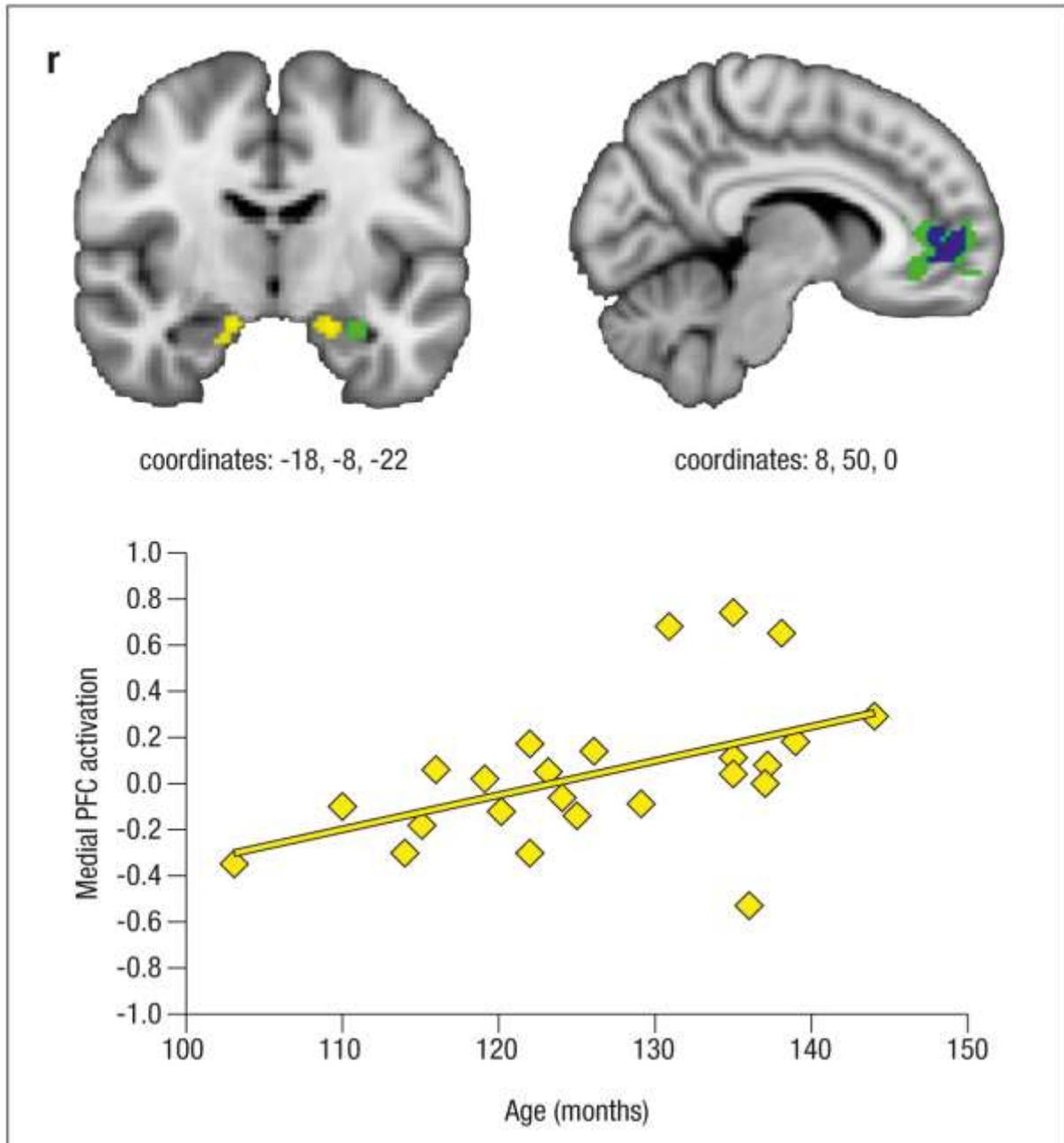


Figure 2. Successful memory retrieval for congruent events. Regions showing enhanced activation for congruent SR vs. UR trials encompassed mPFC and left HC in the young adults (green), and both bilateral HC in children (yellow). Medial PFC activation was stronger in young adults than in children (blue), whereas HC activation did not differ between the two

age groups. Contrasts were thresholded voxel wise at $Z > 2.3$ and corrected for multiple comparisons at $p < .05$ using 3dClustSim. In the children group, mPFC activation for congruent SR vs. UR trials was positively correlated with age ($r = .50$, plotted for illustrative purposes).

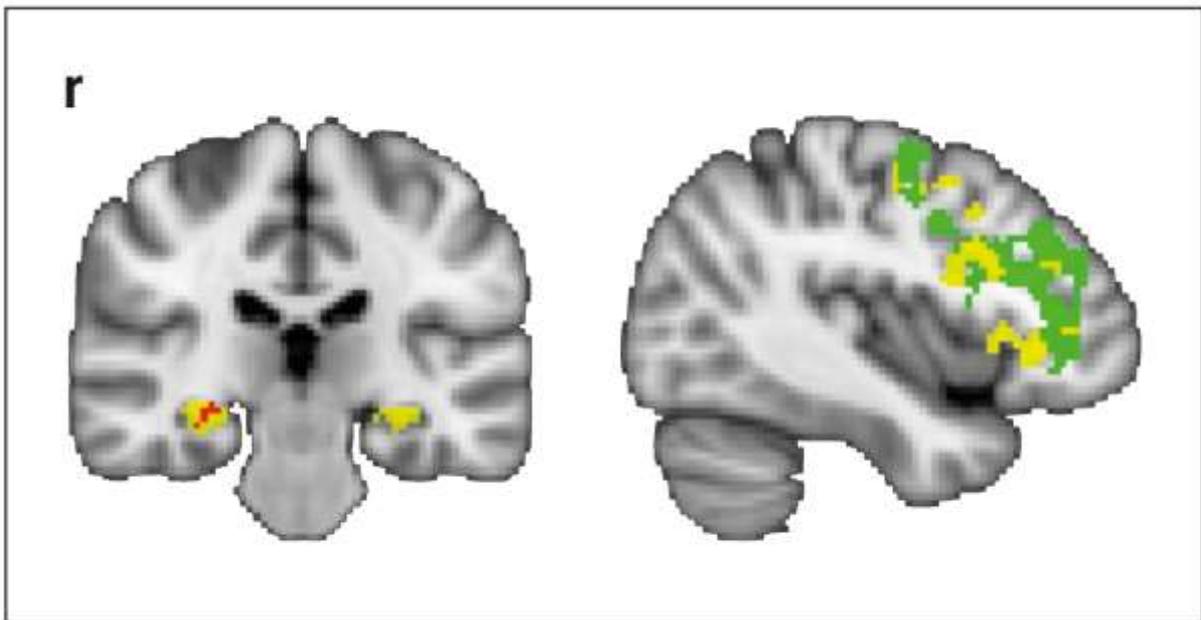


Figure 3. Successful memory retrieval for incongruent events. For young adults, enhanced activation for incongruent SR vs. UR trials was found in the DLPFC (green), whereas children showed activation in the DLPFC and in the HC (yellow). Right HC activation was stronger in children than in young adults (red, left), whereas DLPFC activation did not differ reliably between the two age groups. Contrasts were thresholded voxel wise at $Z > 2.3$ and corrected for multiple comparisons at $p < .05$ using 3dClustSim.

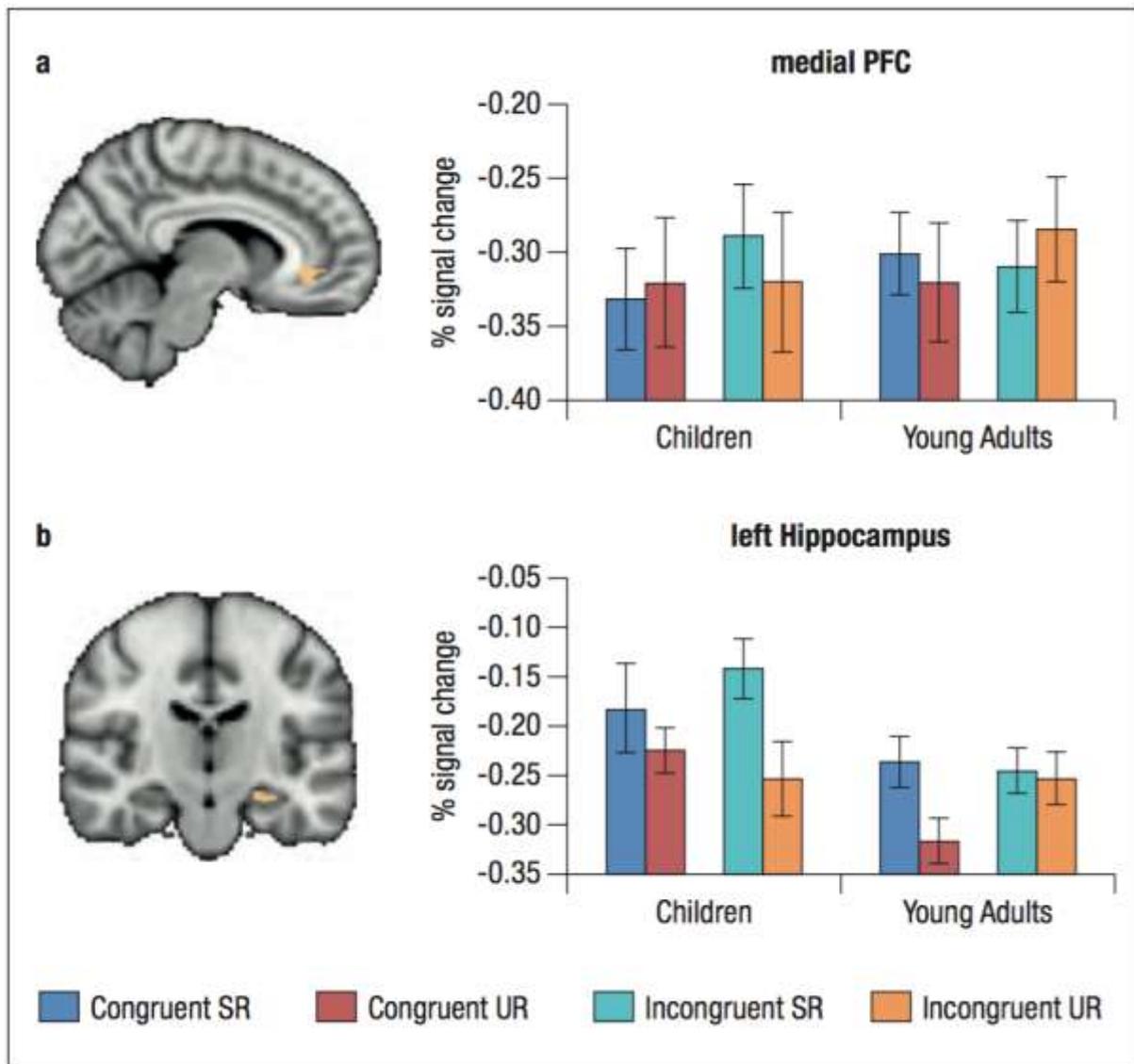


Figure 4. Exploratory memory x age group x congruency analysis. This analysis revealed age differences in the differential retrieval of schema-congruent versus schema-incongruent events. Two clusters exhibiting a memory x congruency x age group interaction were detected: one in the mPFC (Panel a), and one in the left HC (Panel b). a) the mPFC cluster overlapped with the one identified in the young adults vs. children contrast for schema-congruent events (Figure 2), and is adjacent to the memory x congruency interaction cluster for young adults reported in Brod et al. (2015). Follow-up statistical tests indicated a trend towards a memory x congruency interaction in the young adults ($F(1,23) = 2.7, p = .11$), who showed greater mPFC activity for SR vs. UR congruent events and lower mPFC activity for SR vs. UR incongruent events. In children, the interaction did not differ reliably from zero

($F(1,24) = 0.07, p = .79$). b) the left hippocampus cluster was located contralateral to the left HC cluster that displayed stronger activation in the children schema-incongruent events. Follow-up tests indicated a memory x congruency interaction in young adults ($F(1,23) = 7.2, p = .01$), who showed greater left HC activation for SR vs. UR congruent events than for incongruent events. In the children group, the corresponding interaction was not statistically significant ($F(1,24) = 0.8, p = .38$). Error bars are between-subject standard errors.

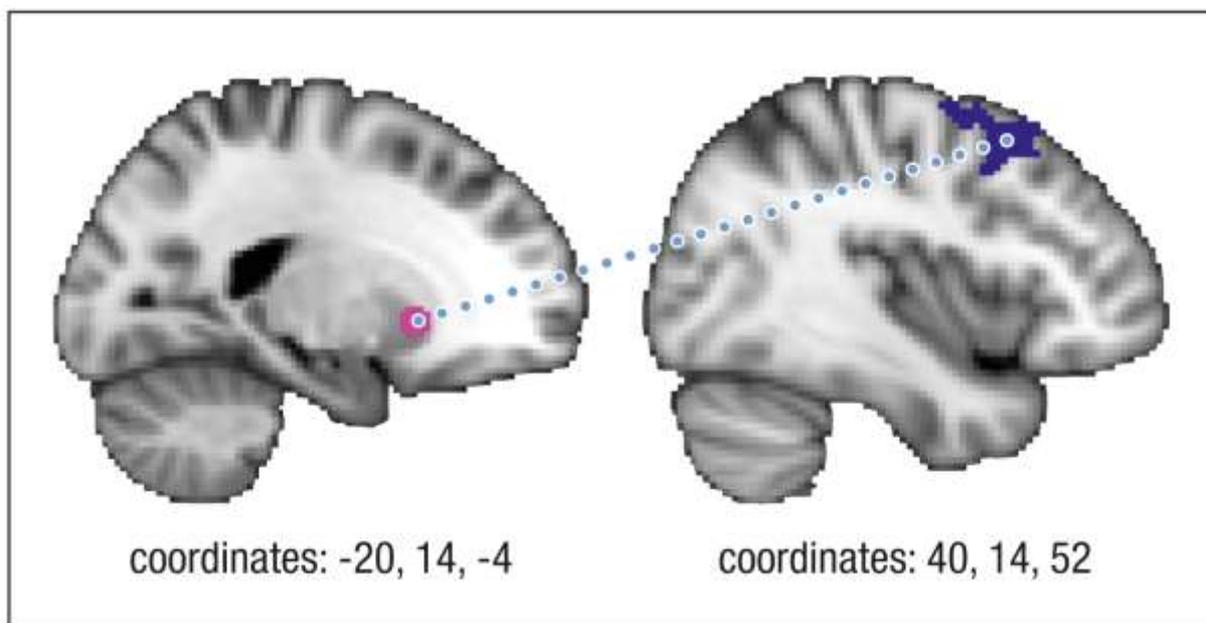


Figure 5. Psychophysiological interaction (PPI) analysis between striatum and DLPFC. PPI analyses with seeds in the left putamen revealed stronger striatum–DLPFC connectivity in young adults than in children (blue) for schema-incongruent > schema-congruent SR trials. Contrast were thresholded voxel wise at $Z > 2.3$ and corrected for multiple comparisons at $p < .05$ using 3dClustSim.

