Supplementary Materials

"Neural Correlates and Reinstatement of Immediate-, Short- and Long-Delay Memory Consolidation: A Comparison Between Children and Young Adults"

Iryna Schommartz[a,i], Philip F. Lembcke[b,c], Javier Ortiz-Tudela[a], Bauer[b], M., Angela M. Kaindl[c,d,e,f], Claudia Buss[b,h]*, and Yee Lee Shing[a,i]*

* Yee Lee Shing and Claudia Buss should be considered joint senior author.

[a]Department of Psychology, Goethe University Frankfurt, Frankfurt, Germany
[b] Charité – Universitätsmedizin Berlin, Department of Medical Psychology, Berlin, Germany
[c]Charité – Universitätsmedizin Berlin, Department of Pediatric Neurology, Berlin, Germany
[d]Charité – Universitätsmedizin Berlin, Center for Chronically Sick Children, Berlin, Germany
[e]Charité – Universitätsmedizin Berlin, Institute of Cell- and Neurobiology, Berlin, Germany
[f]Charité – Universitätsmedizin Berlin, Department of Pediatric Surgery, Berlin, Germany
[h]Development, Health and Disease Research Program, Department of Pediatrics, University of California Irvine, USA
[i]Center for Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt, Germany

***Corresponding authors at:** Department of Psychology, Goethe University Frankfurt, Frankfurt, Germany, Theodor-W.-Adorno-Platz, 6, 60323

E-mail adresses: schommartz@psych.uni-frankfurt.de (I. Schommartz),

shing@psych.uni-frankfurt.de (Y.L. Shing)

S1. Supplementary Methods

S1.1. Assessment of demographic and cognitive covariates

Other cognitive covariate tasks, such as cognitive switching and object-location memory, were run on each session but they are not included in the current paper.
***Day 0*:** After the experimental task, several subtests of the K-ABC II Test *(e.g., Atlantis, Rover, Rebus, Riddle and Atlantis delayed) were administered to children, while young adults were tested with the WAIS-IV Test.
Day 1: In addition, children performed several subtests of the K-ABC II Test *(e.g., Expressive Vocabulary, Triangles, Pattern Reasoning), and a cognitive switching task.
Day 14: Children performed several subtests of the K-ABC II Test *(e.g., Patterns, Verbal Knowledge, Word Order), and a object-location memory task.

In addition to the experimental paradigm, a sleep diary to assess the quality and duration of sleep was completed daily for the 14-day period between learning and long-delay.

S1.2. FMRI data pre-processing

The following description of the fMRI data pre-processing was generated by fMRIPrep 22.0.0:
Results included in this manuscript come from preprocessing performed using *fMRIPrep* 22.0.0 (Esteban et al., 2018, 2019; RRID:SCR_016216), which is based on *Nipype* 1.8.3 (Gorgolewski et al., 2011; Gorgolewski et al., 2016); RRID:SCR_002502).

S1.2.1.Preprocessing of B_0 inhomogeneity mappings

A total of 2 fieldmaps were found available within the input BIDS structure for this particular subject. A B_0-nonuniformity map (or *fieldmap*) was estimated based on two (or more) echo-planar imaging (EPI) references with topup (Andersson et al. (2003) ; FSL 6.0.5.1:57b01774).

S1.2.2. Anatomical data preprocessing

A total of 2 T1-weighted (T1w) images were found within the input BIDS dataset. All of them were corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al., 2010), distributed with ANTs 2.3.3 (Avants et al. (2008); RRID:SCR_004757). The T1w-reference was then skull-stripped with a *Nipype* implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 6.0.5.1:57b01774, RRID:SCR_002823; Zhang et al., (2001)). A T1w-reference map was computed after registration of 2 T1w images (after INU-correction) using mri_robust_template (FreeSurfer 7.2.0; Reuter et al., (2010)). Volume-based spatial normalization to two standard spaces (MNI152NLin6Asym, MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both T1w

reference and the T1w template. The following templates were selected for spatial normalization: *FSL's MNI ICBM 152 non-linear 6th Generation Asymmetric Average Brain Stereotaxic Registration Model* [Evans et al. (2012); RRID:SCR_002823; TemplateFlow ID: MNI152NLin6Asym], *ICBM 152 Nonlinear Asymmetrical template version 2009c* [Fonov et al. (2009); RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym].

S1.2.3. Functional data preprocessing

For each of the 5 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated by aligning and averaging 1 single-band references (SBRefs). Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 6.0.5.1:57b01774; Jenkinson et al. (2002)). The estimated *fieldmap* was then aligned with rigid-registration to the target EPI (echo-planar imaging) reference run. The field coefficients were mapped on to the reference EPI using the transform. BOLD runs were slice-time corrected to 0.346s (0.5 of slice acquisition range 0s-0.693s) using 3dTshift from AFNI (Cox & Hyde, (1997); RRID:SCR_005927). The BOLD reference was then co-registered to the T1w reference using mri_coreg (FreeSurfer) followed by flirt (FSL 6.0.5.1:57b01774; Jenkinson & Smith (2001) with the boundary-based registration (Greve & Fischl, 2009) cost-function. Co-registration was configured with six degrees of freedom. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al. (2014) and Jenkinson et al. (2002) (relative root mean square displacement between affines). FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* (following the definitions by Power et al. (2014)). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (*CompCor*; Behzadi et al. (2007)). Principal components are estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, a mask of pixels that likely contain a volume fraction of GM is subtracted from the aCompCor masks. This mask is obtained by thresholding the corresponding partial volume map at 0.05, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the

nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardized DVARS were annotated as motion outliers. Additional nuisance timeseries are calculated by means of principal components analysis of the signal found within a thin band (*crown*) of voxels around the edge of the brain, as proposed by Patriat et al. (2017). The BOLD time-series were resampled into several standard spaces, correspondingly generating the following *spatially-normalized, preprocessed BOLD runs*: MNI152NLin6Asym, MNI152NLin2009cAsym. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Automatic removal of motion artifacts using independent component analysis (ICA-AROMA; Pruim et al. (2015)) was performed on the *preprocessed BOLD on MNI space* time-series after removal of non-steady state volumes and spatial smoothing with an isotropic, Gaussian kernel of 6mm FWHM (full-width half-maximum). Corresponding "non-aggresively" denoised runs were produced after such smoothing. Additionally, the "aggressive" noise-regressors were collected and placed in the corresponding confounds file. All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer). Many internal operations of *fMRIPrep* use *Nilearn* 0.9.1 (Abraham et al. (2014); RRID:SCR_001362), mostly within the functional processing workflow.

S2. Supplementary behavioural results

Table S1

Statistical overview of the linear mixed effects model for recent memory retention rates for initially correctly learned items (corrected for chance performance).

	(A) Recent Memory Retention		**(B) Overall Memory Retention**	
Predictors	*F-value*$_{(DenDF)}$	*p-value*	*F-value*$_{(DenDF)}$	*p-value*
Session	$5.19_{(1,75)}$	**.026**		
Group	$47.44_{(1,83)}$	**<.001**	$55.00_{(1,85)}$	
Item Type			$229.17_{(3,250)}$	**<.001**
IQ	$2.39_{(1,88)}$	.125	$5.82_{(1,86)}$	**.018**
Sex	$1.73_{(1,89)}$	.191	$2.57_{(1,87)}$	.113
Session x Group	$1.77_{(1,75)}$	.187		
Item Type x Group			$17.35_{(3,250)}$	**<.001**

Random Effects

σ^2	59.91	57.36
$\tau_{00\ subNo}$	74.73	26.37
ICC	.56	.31
N_{subNo}	88	88
Observations	158	336
Marginal R^2 / Conditional R^2	0.335/ 0.704	.659/.767

Notes. Subject was included as random intercept. Group (children and young adults), Session recent memory retention (Day 1, Day 14 and Day1 and Day14 after 30 minutes), Item Type (baseline$_{learning}$, immediate, recent vs remote) were included as fixed effects. IQ, Sex, Handedness were included as covariates. [a]The following reference levels where used: for Session, Day 1/14; for Group, Children; for Item Type, baseline; for Sex, male; for Handedness, right-side handedness. IQ = Intelligence Quotient; σ2 – residuals, τ00 – variance of the random intercept. Type III Analysis of Variance Table with Satterthwaite's method.

S2.1. Memory Strength across Time

To analyse the time-related change in the memory strength, we employed the drift diffusion modelling approach (Forstmann et al., 2016; Fudenberg et al., 2020; Ratcliff & McKoon, 2008; Wagenmakers et al., 2007a). This approach utilizes performance accuracy and reaction time. We calculated following parameters: (i) the drift rate (v), which indicates memory strength or the average rate of evidence accumulation; (ii) the boundary (a) parameter, which indicates the amount of evidence required to decide or stringency of the decision; (iii) the non-decision time (Ter), which reflects sensorimotor processing time. The analysis was based on the EZ-diffusion model (Wagenmakers et al., 2007). In this model, the parameters are estimated based on memory performance accuracy, the mean and the variance of reaction time of the correct responses. With the derived parameters, we conducted linear mixed-effect models (LME model) for memory measures using the lmer function from the lme4 package in R (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017). All LME models were calculated with maximum-likelihood estimation and Subject as the random intercept to account for between-subject variability in the derived parameters of the drift diffusion model. For that, we included the within-subject factor of *Session* (Day 0, Day 1, and Day 14) and the between-subject factor of *Group* (children and young adults) in the LME models. All main and interaction effects were False Discovery Rate adjusted for multiple comparisons.

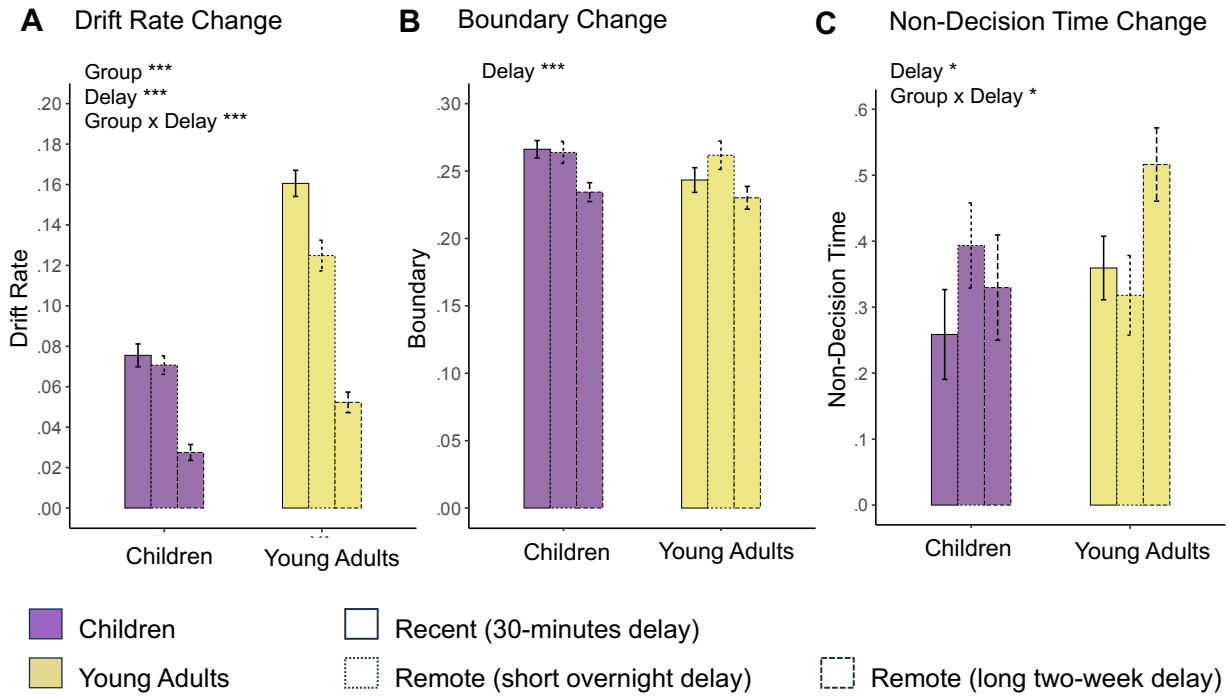
To characterize the change in memory strength across time within and between child and adult groups, we employed the drift diffusion modelling approach (Forstmann et al., 2016; Fudenberg et al., 2020; Ratcliff & McKoon, 2008; Wagenmakers et al., 2007a) that utilizes not only performance accuracy but also reaction time in complex tasks (Criss, 2010; Lerche & Voss, 2019; Palada et al., 2016; Zhou et al., 2021) and can be applied for different developmental groups (Ratcliff et al., 2011, 2012). We calculated (i) the drift rate (v), which indicates the average rate of evidence accumulation in favour of a correct decision. Thus, the drift rate reflects accessibility of memory representations: a higher value indicates a greater probability of making a correct decision, indicating stronger memory. Conversely, lower values suggest slower accumulation of evidence, possibly indicating difficulty in processing information or a lower signal-to-noise ratio strength (Turker & Swallow, 2022). Further, we calculated also (ii) the boundary (a) parameter, which indicates the amount of evidence required to decide. Larger boundary values mean that more information is needed before

deciding, leading to more accurate but slower decisions. Conversely, a smaller boundary value suggests that less information is needed, resulting in faster but potentially less accurate decisions. Lastly, (iii) the non-decision time (Ter) was calculated, reflecting the portion of response time that is not related to decision process. A low non-decision time suggests that most of response time is consumed by actual mnemonic decision process rather than peripheral processes. Conversely, a high non-decision time indicates that a large portion of response time is taken up by processes other than mnemonic decision-making.

All these parameters, namely the drift rate, boundary, and non-decision time, were calculated for children and young adults for recent (immediately retrieved), remote Day 1 and remote Day 2 memory items. For recent memory items, we aggregated the drift rates , the boundary, and non-decision time across two sessions, as there were no significant differences between sessions, as indicated by nonsignificant *Session* and *Session x Group* interactions (all $p > .13$). Additionally, we conducted LME model analyses for each parameter, with *Subject* as a random factor, and *Group* and *Delay* as fixed effects.

Firstly, the Linear Mixed Effects (LME) model for drift rate (v) explained a significant amount of variance $R^2 = .83$, 95% CI [.83 - .88]. We observed a significant main effect of *Group*, $F_{(1,84)} = 86.56$, $p < .001_{\text{FDR-adjusted}}$, $w^2 = .44$, indicating a lower overall drift rate in children compared to young adults, $b = -.06$, $t_{(89)} = -8.24$, $p < .001$. There was also a significant *Delay* effect, $F_{(2,156)} = 215.43$, $p < .001_{\text{FDR-adjusted}}$, $w^2 = .73$, showing an overall higher drift rate for recent items compared to remote Day 1 items, $b = .02$, $t_{(161)} = 5.54$, $p < .001$, and the drift rate was significantly higher for remote Day 1 compared to remote Day 14, $b = .06$, $t_{(165)} = 14.64$, $p < .001$. Additionally, there was a significant *Group x Delay* interaction, $F_{(2,156)} = 28.08$, $p < .001_{\text{FDR-adjusted}}$, $w^2 = .25$. Sidak-corrected post hoc tests revealed that the slope of decrease of the drift rate from recent to remote Day 1 was more pronounced in young adults compared to children, $b = -.03$, $t_{(161)} = -4.24$, $p = <.001$, and the slope of decrease of the drift rate from remote Day 1 to remote Day 14 was steeper in young adults, $b = -.03$, $t_{(165)} = -3.28$, $p = .008$. The results show overall lower memory strength in children compared to adults, indicating less effective long-term memory consolidation in children compared to young adults already immediately after learning and extending into longer delays. Albeit adults showed higher memory strength during all delays, the decline rate was faster compared to children, indicating with this profound changes in the memory strength of initially strong memories that stronger memories tend to lose more.

Figure S1



Delay-related Change in Memory Strength as Indicated by Drift Rate, Boundary and Non-decision Time Change Within and Between Children and Young Adults. (A) Drift Rate Change reflects the change in the memory strength or efficiency of evidence accumulation (retrieval processes) to choose a correct item location. (B) Boundary Change reflects the delay-related change in the stringency of retrieval-based decision process. (C) Non-decision time change reflects the delay-based change in sensorimotor processing during memory retrieval decision. *$p < .05$; **$p < .01$; ***$p < .001$(significant difference); non-significant differences were not specifically highlighted. Error bars indicate standard error based on the underlying LME-model.

Secondly, the LME model for the boundary (a) explained a significant amount of variance R2 = .43, 95% CI [.39 - .53]. It revealed a significant main effect of *Delay*, $F_{(2,159)} = 11.32$, p = $<.001_{FDR\text{-}adjusted}$, $w^2 = .11$. The overall boundary remained constant for recent to remote Day 1 items, b = -.008, $t_{(161)} = -1.24$, p = .520, but was significantly higher for remote Day 1 compared to remote Day 14 memories, b = .03, $t_{(167)} = 4.57$, p = <.001. Neither the *Group* effect nor the *Group x Delay* interaction was significant (all p > .227), indicating that the boundary and its change over time were similar in children and young adults. Overall, these findings indicate a slight decrease in boundary separation from a short to a long remote delay across both age groups. This decrease might suggest that participants are slightly more inclined to make mnemonic decisions with less evidence after two weeks.

Thirdly, the LME model for the non-decision time (Ter) explained a significant amount of variance R2 = .50, 95% CI [.44 - .60]. The LME revealed a significant main effect of *Delay*, $F_{(1,157)} = 3.57$, $p = .030_{FDR\text{-}adjusted}$, $w^2 = .03$. Sidak-adjusted post hoc tests revealed overall lower non-decision time for recent items compared to remote Day 14 items, b = -.13, $t_{(165)} = -2.63$, p = .028. There was no significant main effect of *Group* (p = .293), indicating similar non-decision time between children and adults. In addition, a significant *Group x Delay* interaction was observed, $F_{(2,157)} = 4.32$, $p = .022_{FDR\text{-}adjusted}$, $w^2 = .04$. The Sidak-adjusted post hoc tests

showed significantly higher non-decision time for remote Day 14 memories compared to remote Day 1 memories in young adult, b = .22, $t_{(164)}$ = 3.10, p = .013. This delay-related increase in adults was significantly higher compared to children, b = .28, $t_{(166)}$ = 2.83, p = .031. There were no other significant between or within group difference in the non-decision time (all p > .18). Overall, these findings suggest that overall increase in non-decision time over time was driven by the young adult group.

Table S2

Statistical overview of the main and interaction effects of the linear mixed effects model for drift diffusion parameters.

	Main Effect of Group		**Main Effect of Delay**		**Group x Delay Interaction**		
Regions of Interest	$F_{(DF)}$	p	$F_{(DF)}$	*p*	$F_{(DF)}$	*p*	*R2*
V	69.56(1,84)	<.001	215.43(2,156)	<.001	28.08(2,156)	<.001	.829
A	1.42(1,85)	.293	11.32(2,159)	<.001	1.50(2,159)	.227	.429
Ter	$1.12_{(1,85)}$	.293	$3.57_{(2,157)}$	.030	$4.32_{(2,157)}$	.022	.500

Notes. Subject was included as random effect. Group (children, young adults), Delay (recent, remote (Day 1), remote (Day 14)), and their interaction were included as fixed effect. The following reference levels where used: for Delay, recent; for Group, Children; V – drift rate; A – boundary; Ter – Non-decision Time; F – F-value; DF – degrees of freedom; p – p-value; R2 – amount of variance explained by the model (Stoffel et al., 2021). All main and interaction effects are False Descovery Rate corrected for multiple comparisons. Type III Analysis of Variance Table with Satterthwaite's method. *p < .05; ** <.01, ***<.001 (significant difference).

S3.1. Supplementary fMRI univariate analysis

Table S3

Regions exhibiting stronger activation for remote vs. recent items in (i) young adults, (ii) children, (iii) children vs young adults, and (iv) young adults vs children on Day 1 (short delay). To capture the involved brain region better, local maxima are presented in addition to cluster maxima for the largest clusters.

Day 1 (Short Delay)					
		Young adults			
Region	**x**	**y**	**x**	**Z-max**	**# voxels**
Left Middle Frontal Gyrus	- 44	2	40	6.67	2990
Left Insula Cortex	- 34	22	2	6.58	
Left Inferior Frontal Gyrus, Pars Opercularis	- 44	6	34	6.03	
Left Lateral Occipital Cortex	- 28	- 76	36	6.82	2272
Left Superior Parietal Lobule	- 34	- 50	44	5.11	
Left Fusiform Gyrus	- 44	- 60	- 12	6.7	1661
Left Parahippocampal Gyrus	- 34	- 34	- 16	4.58	
Right Cerebellum	30	- 60	- 28	6.03	1049
Right Lateral Occipital Cortex	34	- 72	40	5.96	943
Right Inferior Parietal Lobule	38	- 78	26	4.3	
Right Parahippocampal Gyrus	32	- 34	- 16	5.29	718

Right Inferior Temporal Gyrus	52	- 54	- 10	5.17	
Left Superior Frontal Gyrus	- 4	16	48	5.04	405
Right insular cortex	30	24	2	5.25	279
Right Middle Frontal Gyrus, Pars Triangularis	40	30	20	3.61	
Right precentral Gyrus	42	2	30	4.97	146
Right Middle Frontal Gyrus, Pars Opercularis	50	16	32	3.41	
Left Frontal Orbital Cortex	- 26	32	- 10	4.51	123
Left Cingulate Gyrus	- 4	2	28	4.86	103
Children					
Right Temporal Occipital Fusiform Cortex	26	- 44	- 8	5.1	658
Right Parahippocampal Gyrus	30	- 36	- 16	4.93	
Right Precuneus	8	- 52	6	4.79	
Left Temporal Fusiform Gyrus	- 34	- 42	- 12	5.59	500
Left Parahippocampal Gyrus	- 18	- 42	- 10	4.91	
Left Precuneus Cortex	- 14	- 60	10	4.47	160
Left Lateral Occipital Cortex	- 36	- 84	26	4.95	112
Children > Young Adults					
Right precuneus	4	- 48	30	5.25	1051
Left precuneus	- 4	- 48	40	4.68	
Right Superior Parietal Lobule	12	- 32	50	4.99	203
Right Parietal Operculum Cortex	54	- 30	24	3.32	149
Young Adults > Children					
Left Precentral Gyrus, Middle Frontal Gyrus	- 44	2	40	4.8	501
Left Inferior Frontal Gyrus	- 54	14	10	3.39	
Left Frontal Operculum Cortex	- 34	22	2	5.48	260
Right Cerebellum	12	- 76	- 20	4.7	141
Left Medial Frontal Gyrus	- 2	16	48	4.2	118
Left/Right Insular Cortex	32	22	2	4.66	113
Left/Right Lateral Occipital Cortex	- 26	- 74	36	4.5	107

Table S4

Regions exhibiting stronger activation for remote vs. recent items in (i) young adults, (ii) children, (iii) children vs young adults, and (iv) young adults vs children on Day 14 (long delay). To capture the involved brain region better, local maxima are presented in addition to cluster maxima for the largest clusters.

Day 14 (Long Delay)					
Young Adults					
Region	**x**	**y**	**x**	**Z-max**	**# voxels**
Left/Right Occipital Fusiform Gyrus	- 46	- 58	- 16	7.62	19227
Left Lateral Occipital Cortex	- 30	- 60	- 14	7.25	
Left Middle Frontal Gyrus, Pars Opercularis,				7.17	2890
Left Superior Frontal Gyrus	- 6	12	56	6.78	
Right Inferior Frontal Gyrus, Pars Opercularis, Pars Trinagularis	46	12	28	6	691
Left Insular Cortex	- 32	22	2	6.7	501
Left Caudate	- 10	4	10	5.58	456
Right Frontal Orbital Cortex	34	28	0	6.11	298
Right Cerebellum	16	- 44	- 46	4.97	250
Right Caudate	8	12	2	5.27	215
Left Cerebellum	- 34	- 68	- 54	6.1	211

Children					
Left Temporal Fusiform Gyrus	- 34	- 26	- 24	4.91	580
Left anterior Parahippocampal Gyrus, Hippocampus	- 36	- 18	- 24	4.4	
Left Lateral Occipital Cortex	- 48	- 58	- 16	4.25	
Right Temporal Occipital Fusiform Cortex	40	- 54	- 18	4.34	448
Right Lateral Occipital Cortex	50	- 70	- 12	4.2	
Children > Young Adults					
Right/Left angular gyrus	62	- 40	44	4.8	847
Right/Left Lateral Occipital Cortex	46	- 66	48	4.44	
Right Superior Frontal Gyrus	20	30	58	4.58	640
Right/Left Superior Temporal Gyrus				4.73	493
Right Precuneous	8	- 52	30	4.51	332
Right Medial Frontal Cortex	8	50	- 2	4.35	287
Right Middle Temporal Gyrus	66	- 18	- 20	4.17	203
Left Middle Frontal Gyrus	- 20	36	38	4.31	154
Left Cingulate Gyrus	- 14	- 50	30	4.36	138
Young Adults > Children					
Right/Left Cerebellum	14	- 72	- 22	5.77	3162
Left Occipital Fusiform Gyrus	- 20	- 90	- 14	5.22	1229
Left Lateral Occipital Cortex	- 30	- 80	36	5.62	620
Left Middle Frontal Gyrus, Inferior Frontal Gyrus	- 44	12	30	4.8	387
Right Precuneous	18	- 58	20	4.39	205
Left Superior Frontal Gyrus	- 6	12	56	5.12	165
Left Posterior Parahippocampal Gyrus, Hippocampus	- 28	- 32	- 18	3.9	96

Table S5

Regions exhibiting stronger activation for remote vs. recent items that decreases over time (i) in young adults stronger than in children (ii) children stronger than in adults; that increases over time (iii) in young adults stronger than in children, and (iv) in children stronger than in young adults. To capture the involved brain region better, local maxima are presented in addition to cluster maxima for the largest clusters.

Decrease Across Time					
Young Adults > Children					
Region	**x**	**y**	**x**	**Z-max**	**# voxels**
Right Superior Parietal Lobule, Agular Gyrus	42	- 50	58	3.69	946
Right Middle Frontal Gyrus	42	56	2	4.16	546
Left Middle Frontal Gyrus	- 38	24	48	3.9	379
Right Superior Frontal Gyrus	8	48	30	3.44	329
Children > Adults					
Left Lateral Occipital Cortex	- 32	- 88	6	4.81	4474
Left Hippocampus, Posterior Parahippocampal Gyrus	- 30	- 30	- 6	4.09	
Right Lateral Occipital Cortex, Occipital Fusiform Gyrus, Lingual Gyrus	30	- 86	4	4.73	1717
Increase Over Time					

Young Adults > Children					
Left Lateral Occipital Cortex	- 32	- 88	6	4.81	4474
Left Hippocampus	- 30	- 30	- 6	4.09	
Left Lingual gyrus	- 10	- 56	- 6	4.04	
Right Lateral Occipital Cortex, Occipital Fusiform Gyrus, Precuneus	- 30	86	4	4.73	1717
Children > Young Adults					
Right Superior Parietal Lobule, Angular Gyrus	42	- 50	58	3.69	946
Right Middle Frontal Gyrus	42	56	2	4.16	546
Left Middle Frontal Gyrus, Superior Frontal Gyrus	- 38	24	48	3.9	379
Right Superior Frontal Gyrus, Paracingulate Gyrus	8	48	30	3.44	329

Table S6

Full statistical overview of LME model for univariate analysis.

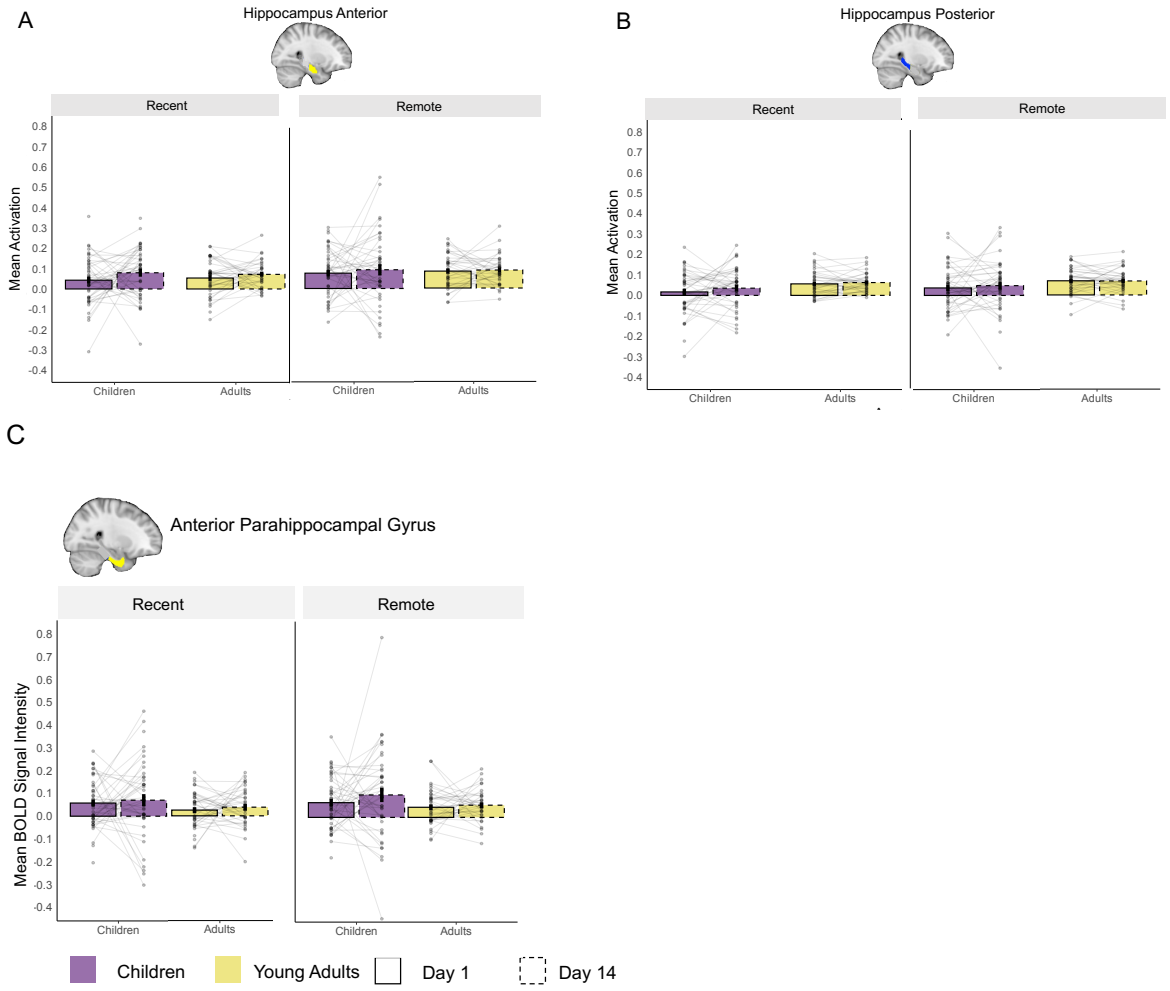
	Main Effect of Group		**Main Effect of Session**		**Group x Session Interaction**		
Regions of Interest	$F_{(DF)}$	*p*	$F_{(DF)}$	*p*	$F_{(DF)}$	p	*R2*
Hippocampus Anterior	$.01_{(1,161)}$	$.911_{(.955)}$	$.34_{(1,161)}$	$.560_{(.622)}$	$.03_{(1,161)}$	$.856_{(.880)}$	.022∩
Hippocampus Posterior	$.60_{(1,161)}$	$.430_{(.614)}$	$.40_{(1,161)}$	$.527_{(.622)}$	$.02_{(1,161)}$	$.880_{(.880)}$	.035∩
Parahippocampal Gyrus Anterior	$.32_{(1,161)}$	$.573_{(.714)}$	$.02_{(1,161)}$	$.892_{(.892)}$	$.53_{(1,161)}$	$.466_{(.583)}$	.041∩
Parahippocampal Gyrus Posterior	$2.97_{(1,83)}$	$.088_{(.176)}$	$2.48_{(1,100)}$	$.118_{(.197)}$	$9.54_{(1,83)}$	**$.002_{(.020)}$**	.200⊥
Medial Prefrontal Cortex	$7.61_{(1,86)}$	**$.007_{(.023)}$**	$.42_{(1,99)}$	$.517_{(.622)}$	$1.16_{(1,83)}$	$.284_{(.406)}$	.369⊥
Ventrolateral Prefrontal Cortex	$31.35_{(1,82)}$	**$<.001_{(<.001)}$**	$10.68_{(1,99)}$	**$.001_{(.005)}$**	$1.61_{(1,83)}$	$.207_{(.345)}$	.309⊥
Cerebellum	$1.54_{(1,161)}$	$.215_{(.358)}$	$4.67_{(1,161)}$	$.032_{(.080)}$	$7.68_{(1,161)}$	**$.006_{(.020)}$**	.100∩
Retrosplenial Cortex	$.003_{(1,161)}$	$.955_{(.955)}$	$3.14_{(1,161)}$	$.078_{(.156)}$	$8.56_{(1,161)}$	**$.004_{(.020)}$**	.087∩
Precuneus	$5.09_{(1,161)}$	**$.011_{(.027)}$**	$6.50_{(1,161)}$	**$.011_{(.036)}$**	$1.61_{(1,161)}$	$.205_{(.345)}$	.099∩
Lateral Occipital Cortex	$9.12_{(1,82)}$	**$.003_{(.015)}$**	$16.76_{(1,97)}$	**$<.001_{(<.001)}$**	$6.42_{(1,81)}$	**$.013_{(.032)}$**	.324⊥

	Main Effect of Sex		**Main Effect of Handedness**		**Main Effect of IQ**		**Main Effect Of Reaction Time**	
Regions of Interest	$F_{(DF)}$	*p*	$F_{(DF)}$	*p*	$F_{(DF)}$	p	$F_{(DF)}$	*p*
Hippocampus Anterior	$1.07_{(1,161)}$	.302	$.36_{(1,161)}$	.695	$.978_{(1,161)}$	.324	$.03_{(1,161)}$	.856
Hippocampus Posterior	$.66_{(1,161)}$	.419	$1.1_{(1,161)}$	.337	$2.01_{(1,161)}$	.158	$.00_{(1,161)}$	.990
Parahippocampal Gyrus Anterior	$3.88_{(1,161)}$	.051	$.09_{(1,161)}$	.901	$2.63_{(1,161)}$	.107	$.59_{(161)}$	.466
Parahippocampal Gyrus Posterior	$1.26_{(1,84)}$	.263	$.03_{(1,93)}$	.962	$1.28_{1,84)}$	.259	$.09_{(1,155)}$	.764
Medial Prefrontal Cortex	$.62_{(1,87)}$	.430	$.50_{(1,95)}$	.607	$5.16_{(1,87)}$	**.024**	$.22_{(1,160)}$	.635
Ventrolateral Prefrontal Cortex	$1.11_{(1,83)}$	.294	$.71_{(1,92)}$	.494	$.09_{(1,83)}$	.764	$.20_{(1,154)}$	.654
Cerebellum	$3.15_{(1,161)}$	.077	$.21_{(1,161)}$	.806	$.781_{(1,161)}$	.378	$.11_{(1,161)}$	.741
Retrosplenial Cortex	$.84_{(1,161)}$	.361	$.46_{(1,161)}$	.631	$.00_{(1,161)}$	.996	$1.84_{(1,161)}$	.177
Precuneus	$.35_{(1,161)}$	.553	$.20_{(1,161)}$	.817	$.08_{(1,161)}$	.776	$.137_{(1,161)}$	.712
Lateral Occipital Cortex	$.10_{(1,83)}$	.752	$.76_{(1,92)}$	.468	$3.04_{(1,83)}$	.084	$.005_{(1,159)}$	.944

Notes. Notes. Subject was included as random effect. Group (children, young adults), Session (Day 1 remote > recent, Day 14 remote > recent), and their interaction were included as fixed effect. The following reference levels where used: for Session – Day 1; for Group – Children; F – F-value; DF – degrees of freedom; p – p-value; FDR_adj – False Discovery Rate adjusted; R2 – amount of variance explained by the model (∩- marginal; ⊥ - conditional). Type III Analysis of Variance Table with Satterthwaite's method. *p < .05; ** <.01, ***<.001

(significant difference). All p-values of main and interactions effects were FDR-adjusted for multiple comparisons.

Figure S2.



Mean Blood Oxygen Level-Dependent (BOLD) Signal Intensity

The figure presents the mean blood oxygen level-dependent (BOLD) signal intensity for recent and remote memories on Day 1 and Day 14 in children and adults in **(A)** anterior hippocampus; (**B**) posterior hippocampus; **(C)** anterior parahippocampal gyrus. *Note:* Bars represent the average BOLD signal intensity. The colour indicated the age groups: purple for children and khaki yellow for young adults. Solid-lined bars represent data from Day 1, while dashed-lined bars depict data from Day 14. Across all panels, mean of individual subject data are shown with transparent points. The connecting faint lines reflect within-subject differences across sessions. Error bars indicate standard error of the mean.

Table S7

Test of neural activation during object presentation separately for recent and remote memories for significance (higher than zero).

	Recent	Remote
	Young Adults	

ROI	*Day*	*mean*	*T test*	$p_{(FDRadj)}$	*mean*	*T test*	$p_{(FDRadj)}$
Hippocampus Anterior	Day 1	.054	3.76	.0005	.083	6.42	<.0001
	Day 14	.072	6.25	<.0001	.089	6.95	<.0001
Hippocampus Posterior	Day 1	.056	5.79	<.0001	.069	6.91	<.0001
	Day 14	.063	7.71	<.0001	.068	6.66	<.0001
Parahippocampal Gyrus Anterior	Day 1	.025	1.98	.028	.044	3.45	.0083
	Day 14	.038	2.59	.009	.053	4.31	<.0001
				Children			
		mean	*T test*	$p_{(FDRadj)}$	*mean*	*T Test*	$p_{(FDRadj)}$
Hippocampus Anterior	Day 1	.043	2.51	.0099	.080	4.47	<.0001
	Day 14	.080	4.09	<.0001	.092	3.37	.0012
Hippocampus Posterior	Day 1	.017	1.09	.141	.037	2.41	.011
	Day 14	.035	2.28	.015	.048	2.45	.011
Parahippocampal Gyrus Anterior	Day 1	.058	3.69	.0005	.065	3.81	.0004
	Day 14	.070	2.64	.0083	.099	3.14	.0024

Notes.To test for significance we used one-sample permutation t-test for more robust calculations with Monte-Carlo permutation percentile confidence interval. All p-values for False Discovery Rate (FDR) corrected for multiple comparisons. ROI – region of interest; p – p-value; FDRadj – False Discovery Rate adjustment; *p < .05; ** <.01, ***<.001 (significant difference).

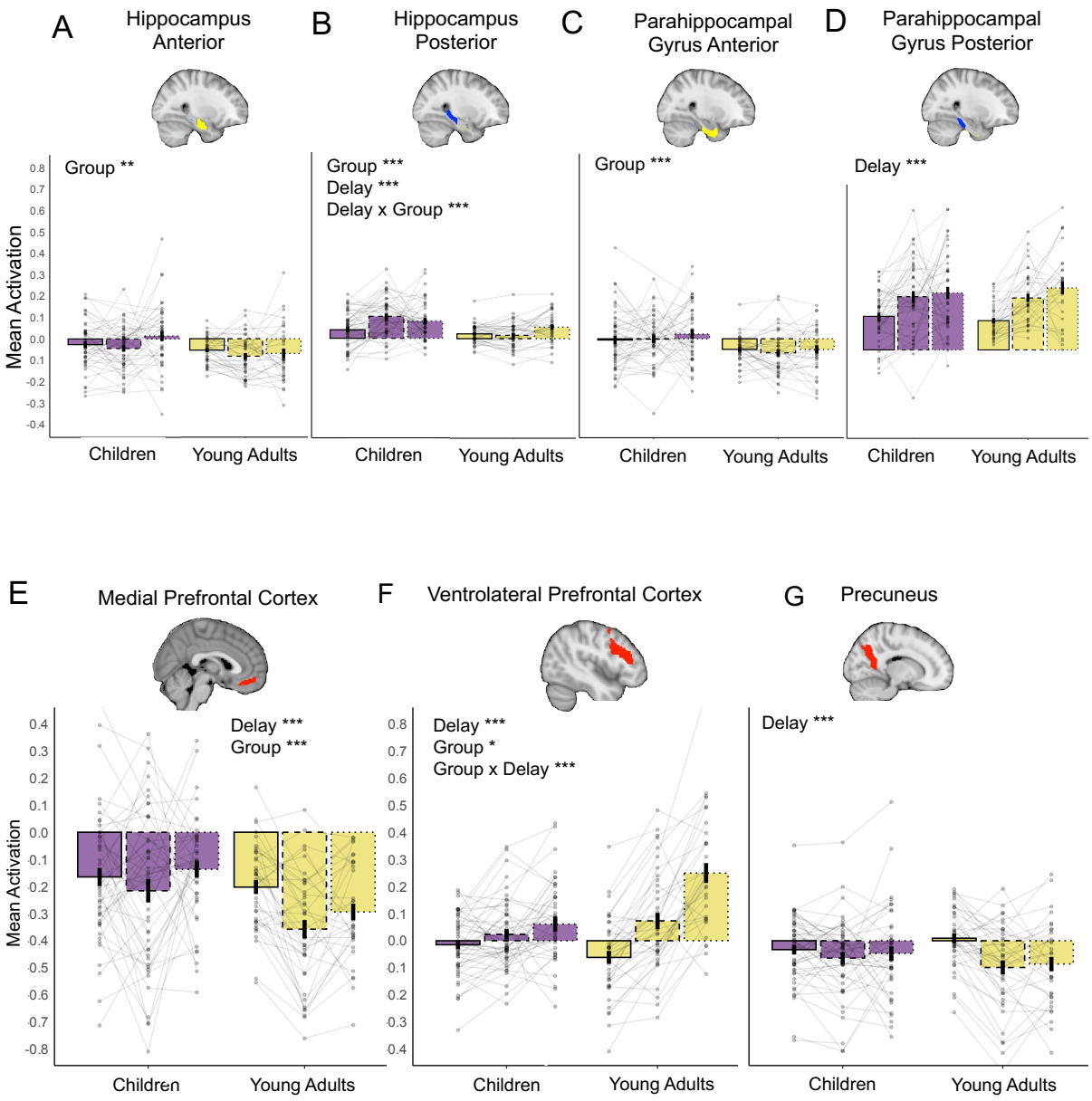
Table S8

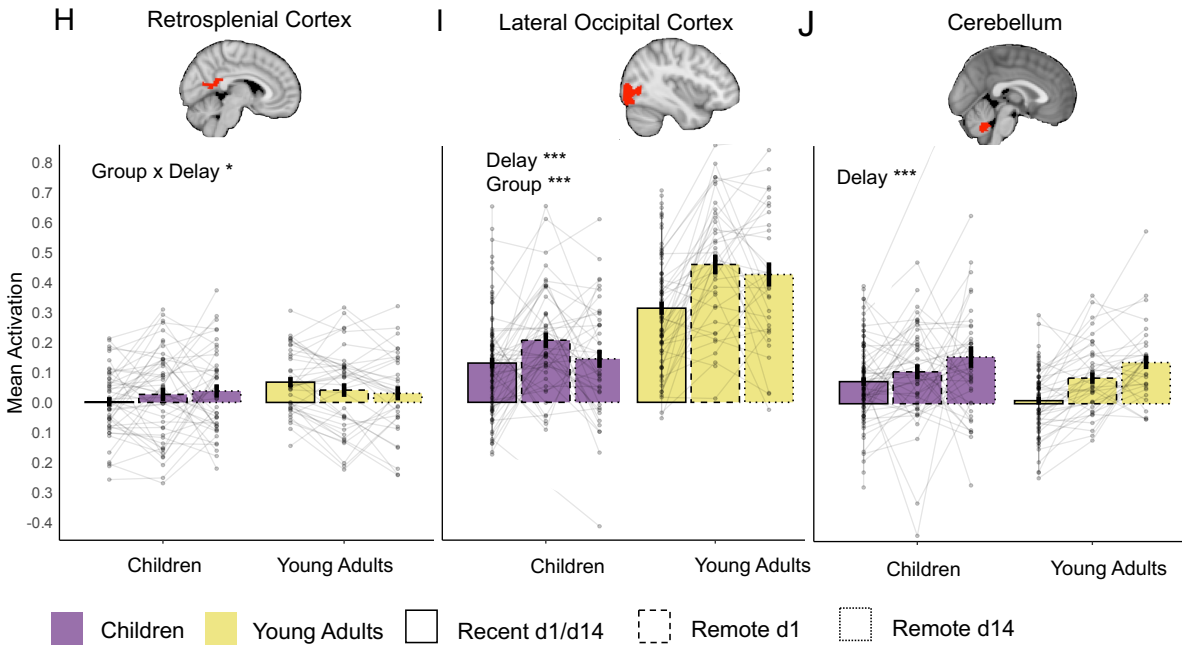
Statistical overview of the main and interaction effects of the linear mixed effects model for scene-specific reinstatement.

	Main Effect of Group		**Main Effect of Session**		**Group x Session Interaction**		**Main Effect of BOLD activation**		
Regions of Interest	$F_{(DF)}$	*p*	$F_{(DF)}$	*p*	$F_{(DF)}$	*p*	$F_{(DF)}$	*p*	*R2*
HCa	$27.21_{(1,86)}$	**<.001**	$100.70_{(2,159)}$	**<.001**	$.94_{(2,159)}$	.393	$.92_{(1,226)}$	.339	.411
HCp	$27.19_{(1,87)}$	**<.001**	$98.18_{(2,159)}$	**<.001**	$1.71_{(2,158)}$	.183	$.97_{(1,240)}$	.324	.417
PHGa	$23.14_{(1,87)}$	**<.001**	$97.74_{(2,159)}$	**<.001**	$1.62_{(2,159)}$	.201	$1.05_{(1,221)}$	.307	.397
PHGp	$15.70_{(1,82)}$	**<.001**	$94.40_{(2,163)}$	**<.001**	$1.85_{(2,155)}$	.161	$.25_{(1,240)}$	.619	.371
mPFC	$8.89_{(1,90)}$	**.0044**	$72.811_{(2,161)}$	**<.001**	$.935_{(2,152)}$	.395	$2.24_{(1,221)}$	.136	.634
vlPFC	$15.18_{(1,90)}$	**<.001**	$71.36_{(2,172)}$	**<.001**	$1.23_{(2,165)}$	.295	$.003_{(1,242)}$	.955	.591
CE	$9.54_{(1,87)}$	**.0038**	$59.99_{(2,166)}$	**<.001**	$1.17_{(2,162)}$	.313	$.679_{(1,228)}$	.411	.520
RSC	$9.27_{(1,89)}$	**.0038**	$79.40_{(2,162)}$	**<.001**	$1.86_{(2,162}$	.159	$.101_{(1,242)}$	.751	.564
PC	$11.35_{(1,85)}$	**.0016**	$74.33_{(2,161)}$	**<.001**	$1.57_{(1,160)}$	.190	$.008_{(1,223)}$	.925	.580
LOC	$1.22_{(1,100)}$	.271	$64.96_{(2,167)}$	**<.001**	$1.05_{(2,162)}$	.350	$1.33_{(1,220)}$	.249	.523

Notes. Subject was included as a random effect. Group (children, young adults), Delay (recent, remote (Day 1), remote (Day 14)), and their interaction were included as fixed effect. The following reference levels where used: for Delay, recent; for Group, Children; F – F-value; DF – degrees of freedom; p – p-value; FDR_adj – False Discovery Rate adjusted; R2 – amount of variance explained by the model (Stoffel et al., 2021); mPFC – medial prefrontal cortex; vlPFC – ventrolateral prefrontal cortex; HCa – anterior hippocampus; HCp – posterior hippocampus; PHGa – anterior parahippocampal cortex; PHGp – posterior parahippocampal cortex; CE – cerebellum; PC – precuneus; RSC – retrosplenial cortex; LOC – lateral occipital cortex.. Type III Analysis of Variance Table with Satterthwaite's method. *p < .05; ** <.01, ***<.001 (significant difference). All main and interactions p-values were FDR-adjusted for multiple comparisons. All main and interactions p-values were FDR-adjusted for multiple comparisons.

Figure S3


A
Hippocampus Anterior
Group **
B
Hippocampus Posterior
Group ***
Delay ***
Delay x Group ***
C
Parahippocampal Gyrus Anterior
Group ***
D
Parahippocampal Gyrus Posterior
Delay ***
Mean Activation
Children
Young Adults
E
Medial Prefrontal Cortex
Delay ***
Group ***
F
Ventrolateral Prefrontal Cortex
Delay ***
Group *
Group x Delay ***
G
Precuneus
Delay ***



Mean Neural Activation for Correctly Recalled Memories during Scene Presentation Time Window. The figure presents mean signal intensity for correctly recalled recent, short delay remote and long delay remote memories in children and adults in **(A)** anterior hippocampus; **(B)** posterior hippocampus; (**C**) anterior parahippocamla gyrus; **(D)** posterior parahippocampal gyrus; **(E)** medial prefrontal cortex; **(F)** ventrolateral prefrontal cortex; **(G)** precuneus; **(H)** retrosplenial cortex; **(I)** lateral occipital cortex; **(J)** cerebellum. *Note:* Bars represent the average signal difference. The colour indicated the age groups: purple for children and khaki yellow for young adults. Solid-lined bars represent data from Day 1, while dashed-lined bars depict data from Day 14. Across all panels, mean of individual subject data are shown with transparent points. The connecting faint lines reflect within-subject differences across sessions. Error bars indicate standard error of the mean. *$p < .05$; **$p < .01$; ***$p < .001$(significant difference); non-significant differences were not specifically highlighted. Significance main and interaction effects are highlighted by the corresponding asterisks. All main and interactions p-values were FDR-adjusted for multiple comparisons.

Table S9

Statistical overview of the main and interaction effects of the linear mixed effects model for scene-based univariate neural analysis

	Main Effect of Group		**Main Effect of Delay**		**Group x Delay Interaction**		
Regions of Interest	$F_{(DF)}$	p	$F_{(DF)}$	*p*	$F_{(DF)}$	*p*	*R2*
HCa	**7.16(1,94)**	**.009**	2.35(2,238)	.097	3.02(2,238)	.051	.320
HCp	**11.67(1,97)**	**.0009**	**8.25(2,241)**	**.0003**	**7.19(2,241)**	**.0009**	.374
PHGa	**11.02(1,90)**	**.001**	.42(2,234)	.660	.927(2,234)	.397	.326
PHGp	.012(1,95)	.914	**36.46(2,240)**	**<.001**	.749(2,240)	.474	.377
Medial Prefrontal Cortex	**10.28**$_{(1,85)}$	**.002**	**7.21**$_{(2,163)}$	**<.001**	2.94 $_{(2,163)}$	.056	.105
Ventrolateral Prefrontal Cortex	**5.96**$_{(1,88)}$	**.016**	**55.14**$_{(2,164)}$	**<.001**	**20.47**$_{(2,164)}$	**<.001**	.262
Cerebellum	1.98$_{(1,80)}$	.163	**13.63**$_{(2,158)}$	**<.001**	.065$_{(2,158)}$	.522	.084
Retrosplenial Cortex	1.05$_{(1,88)}$	.308	.00$_{(2,164)}$	.999	**3.28**$_{(2,164)}$	**.039**	.023
Precuneus	.19$_{(1,88)}$	.666	**12.01**$_{(2,163)}$	**<.001**	**4.54**$_{(2,163)}$	**.012**	.056
Lateral Occipital Cortex	**54.52**$_{(1,88)}$	**<.001**	**17.09**$_{(2,163)}$	**<.001**	**3.55**$_{(2,163)}$	**.031**	.338

Notes. Subject was included as random effect. Group (children, young adults), Delay (recent, remo te (Day 1), remote (Day 14)), and their interaction were included as fixed effect. The following reference levels where used: for Delay, recent; for Group, Children; mPFC – medial prefrontal cortex; vlPFC – ventrolateral prefrontal cortex; HCa – anterior hippocampus; HCp – posterior hippocampus; PHGa – anterior parahippocampal cortex; PHGp – posterior parahippocampal cortex;CE – cerebellum; PC – precuneus; RSC – retrosplenial cortex; LOC – lateral occipital cortex. F – F-value; DF – degrees of freedom; p – p-value; R2 – amount of variance explained by the model (Stoffel et al., 2021). Type III Analysis of Variance Table with Satterthwaite's method. *p < .05; ** <.01, ***<.001 (significant difference).

Table S10

Test of gist-like reinstatement index for significance (higher than zero).

	Recent Pre-activation			Short-Delay Pre-activation			Long-Delay Pre-activation		
	Children								
ROI	*mean*	*p*	$p_{(FDRadj)}$	*mean*	*p*	$p_{(FDRadj)}$	*mean*	*p*	$p_{(FDRadj)}$
mPFC	.010	.084	.168	**.028**	**.008**	**.034**	**.029**	**.011**	**.034**
vlPFC	.007	.068	.135	.010	.038	.135	.016	.047	.135
HCa	.006	.035	.104	**.009**	**.003**	**.018**	.003	.320	.385
HCp	.002	.203	.311	.007	.013	.079	.002	.381	.458
PHGa	.006	.035	.105	.008	.012	.069	.007	.196	.294
PHGp	.006	.062	.186	.008	.023	.140	-.001	.560	.672
CE	.005	.140	.280	.005	.191	.287	-.0004	.516	.620
PC	.007	.059	.179	**.012**	**.007**	**.042**	.006	.223	.334
RSC	.008	.008	.050	.009	.025	.076	.010	.083	.166
LOC	.002	.305	.365	.010	.045	.140	.009	.164	.246
	Young Adults								
	mean	*p*	$p_{(FDRadj)}$	*mean*	*p*	$p_{(FDRadj)}$	*mean*	*p*	$p_{(FDRadj}$
mPFC	.0003	.450	.502	-.00002	.502	.502	.006	.126	.189
vlPFC	-.003	.786	.787	-.0003	.535	.712	.001	.593	.712
HCa	.003	.115	.172	.004	.094	.172	-.004	.863	.863
HCp	.002	.207	.311	.005	.048	.144	-.003	.822	.822
PHGa	.004	.063	.127	.001	.369	.443	-.004	.774	.774
PHGp	.003	.152	.304	-.005	.908	.908	.006	.242	.365
CE	.003	.098	.280	.005	.063	.280	-.003	.735	.736
PC	.0004	.446	.445	.003	.124	.247	.001	.382	.445
RSC	.003	.302	.361	.002	.236	.354	-.0008	.582	.583
LOC	.004	.065	.140	.004	.070	.140	.001	.403	.402

Notes.To test for significance we used one-sample permutation t-test for more robust calculations with Monte-Carlo permutation percentile confidence interval. The p-values of child group were corrected for False Discovery Rate (FDR) for multiple comparisons. ROI – region of interest; p – p-value; FDRadj – False Discovery Rate adjustment; mPFC – medial prefrontal cortex; vlPFC – ventrolateral prefrontal cortex; HCa – anterior hippocampus; HCp – posterior hippocampus; PHGa – anterior parahippocampal cortex; PHGp – posterior parahippocampal cortex; CE – cerebellum; PC – precuneus; RSC – retrosplenial cortex; LOC – lateral occipital cortex. *p < .05; ** <.01, ***<.001 (significant difference).

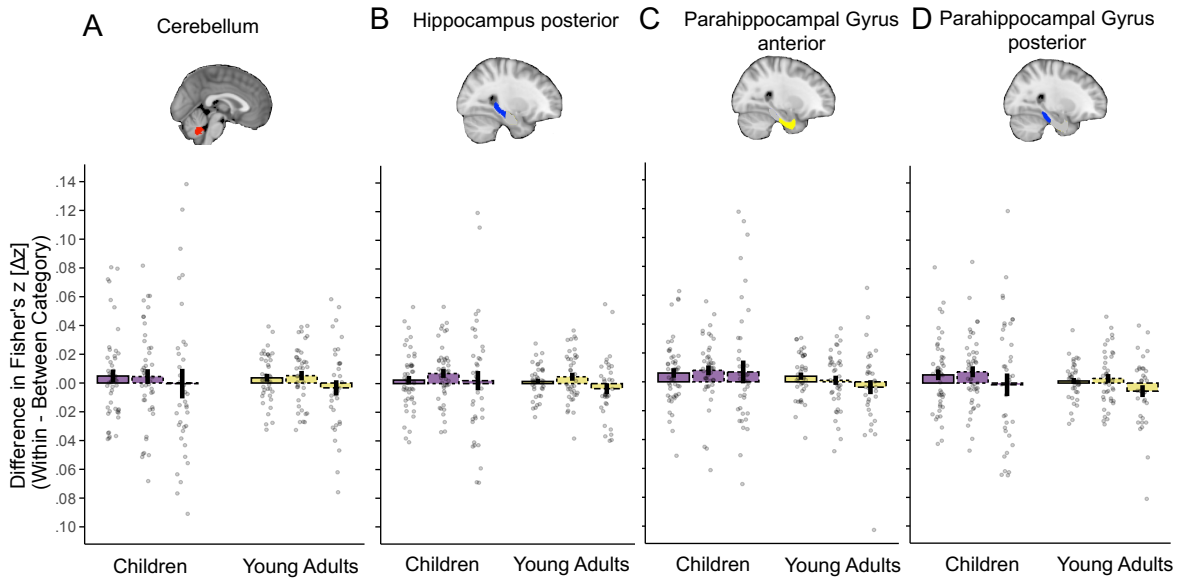
Table S11

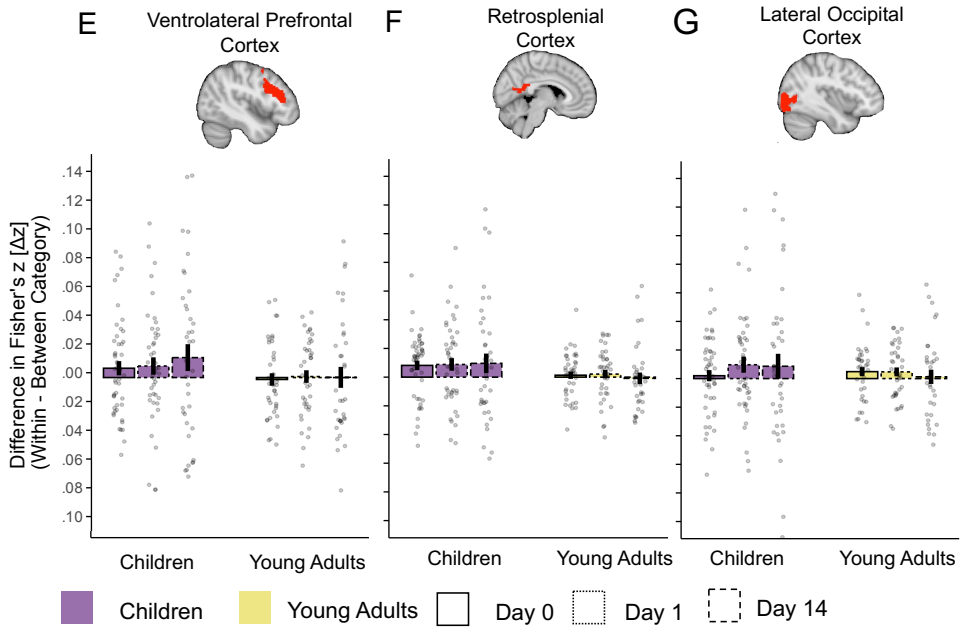
Statistical overview of the main and interaction effects of the linear mixed effects model for gist-like reinstatement.

	Main Effect of Group		**Main Effect of Session**		**Group x Session Interaction**		**Main Effect of BOLD activation**		
Regions of Interest	$F_{(DF)}$	*p*	$F_{(DF)}$	*p*	$F_{(DF)}$	*p*	$F_{(DF)}$	*p*	*R2*
HCa	$2.91_{(1,83)}$	**.111**	$1.88_{(2,162)}$	**.253**	$.15_{(2,162)}$	.859	$1.05_{(1,238)}$	.918	.071
mPFC	$6.77_{(1,75)}$	**.033**	$1.79_{(2,150)}$	**.253**	$.52_{(2,149)}$	.597	$.005_{(1,238)}$	.942	.146
PC	$2.59_{(1,79)}$	**.111**	$.56_{(2,161)}$	**.574**	$.08_{(1,160)}$	.921	$.137_{(1,240)}$	.942	.048

Notes. Subject was included as a random effect. Group (children, young adults), Delay (recent, remote (Day 1), remote (Day 14)), and their interaction were included as fixed effect. The following reference levels where used: for Delay, recent; for Group, Children; F – F-value; DF – degrees of freedom; p – p-value; FDR_adj – False Discovery Rate adjusted; R2 – amount of variance explained by the model (Stoffel et al., 2021); mPFC – medial prefrontal cortex; HCa – anterior hippocampus; PC – precuneus. Type III Analysis of Variance Table with Satterthwaite's method. *p < .05; ** <.01, ***<.001 (significant difference). All main and interactions p-values were FDR-adjusted for multiple comparisons.

Figure S4





Gist-like Reinstatement.
Gist-like reinstatement is reflected by the difference in Fisher's z (Δ z) between within-category and between-category representational similarity during fixation time window, where participants were instructed to reinstate the scene associated with the learned object before the actual scenes were shown. Higher values mean higher gist-like reinstatement. The index was tested for significance against zero and all results were FDR corrected for multiple comparisons. Significant reinstatement of gist-like information is highlighted by a green rectangle (**A**) Cerebellum; (**B**) Hippocampus Posterior; (**C**) Parahippocampal Gyrus Anterior; (**D**) Parahippocampal Gyrus Posterior; (**E**) Ventrolateral Prefrontal Cortex; (**F**) Retrosplenial Cortex; (**G**) Lateral Occipital Cortex. *$p < .05$; **$p < .01$; ***$p < .001$(significant difference); non-significant difference was not specifically highlighted. Error bars indicate standard error.

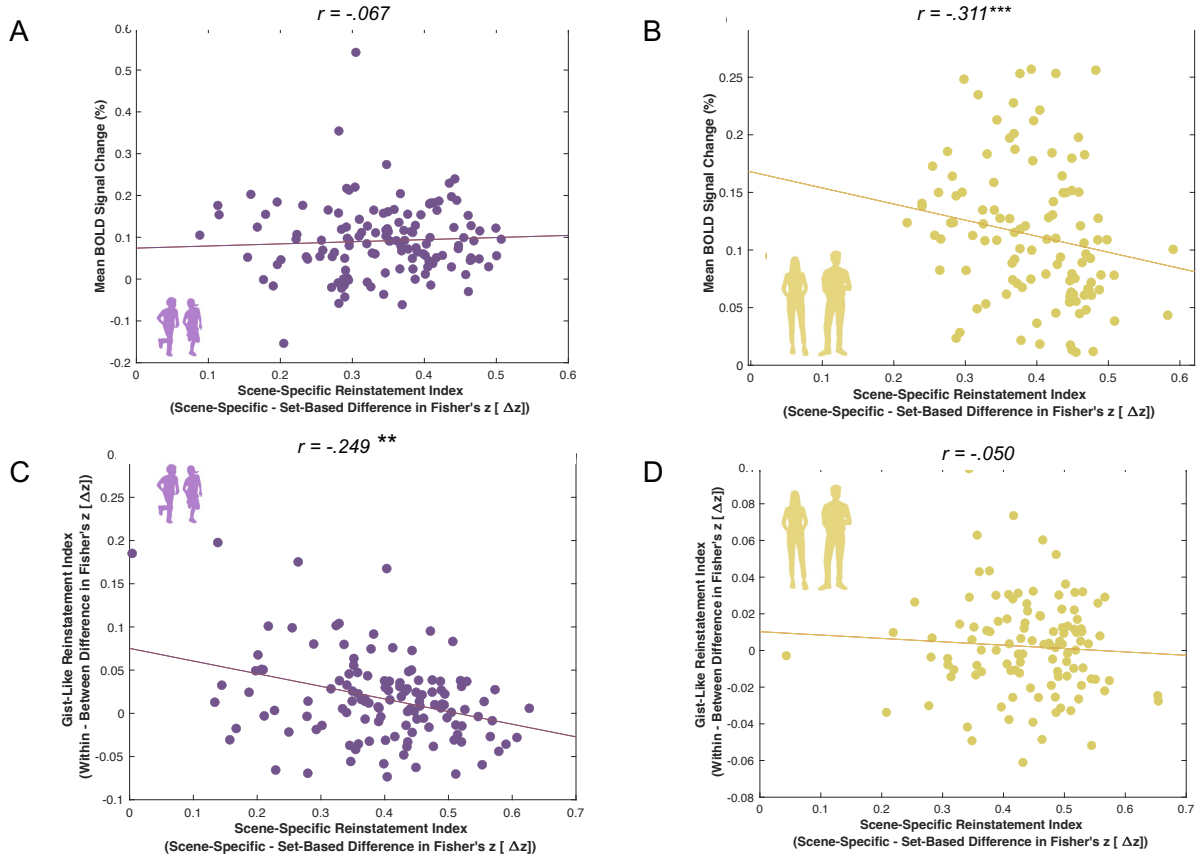
S3.1. Neural-Neural Correlations

Finally, building upon our findings on less pronounced neocortical neural upregulation in children in comparison to adults and gist-like neocortical reinstatement present uniquely in children, we explored whether changes in reinstatement patterns and neural activation are differentially related in children and adults. We hypothesized that young adults may potentially show higher neural upregulation during retrieval which may go along with delay-related attenuations in scene-specific reinstatement. On the other hand, children may show higher gist-like reinstatement that goes along the attenuation of scene-specific reinstatement. With this aim, the Spearman's rank order correlation analysis was employed. The changes in mean neural activation and scene-specific reinstatement were aggregated across all ROIs. The changes in gist-like reinstatement were based on the mPFC in both age groups as the only ROI showing significant group difference.

First, the results revealed that a higher mean neural activation during retrieval was negatively related to scene-specific reinstatement in young adults, $r = -.311$, $p = .00018_{FDR\ adjusted}$; Fig.10B), indicating that when as scene-specific reinstatement decreases, neural activation increases in adults. No such association was observed in children (Fig.10A), $r = .067$, $p = .459_{FDR\ adjusted}$). A Fisher's Z-transformation was used to compare the two

correlation coefficients. The results indicated a significant difference between the correlations in adult and child groups, Z = -2.97, p = .002. The results suggest that the relationship between neural activation and scene-specific reinstatement during retrieval varies significantly with age. Adults show a significant negative association, indicating that higher neural activation is linked to less scene-specific reinstatement. Children do not show this association, which could imply developmental differences in how neural activation during retrieval relates to the reinstatement of specific scenes.

Figure S5



Relation between scene-specific neural reinstatement and memory performance.
Reinstatement brain profiles were derived with the partial least square correlation analysis as a participant's expression of the latent brain pattern across implicated ROIs for reinstatement indices that share the most variance with either short-delay or long-delay memory accuracy variations. Short delay scene-specific reinstatement indices were significantly positively related to short-delay memory accuracy in children (**A;** in purple) but not in young adults (**B**; in yellow). Long delay scene-specific reinstatement indices were significantly positively related to long-delay memory accuracy in children (**C;** in purple) and in young adults. (**D**; in yellow). R = correlation coefficient, p = p-value. All p-values were FDR-adjusted for multiple comparisons.

Second, the results revealed that a higher gist-like neural reinstatement during retrieval was negatively related to scene-specific reinstatement in children, r = -.249, p = $.008_{\text{FDR adjusted}}$; Fig.10C), indicating that when as scene-specific reinstatement decreases, gist-like reinstatement increases in children. No such association was observed in young adults (Fig.10D), r = -.050, p = $.599_{\text{FDR adjusted}}$). A Fisher's Z-transformation was used to compare the

two correlation coefficients. The results indicated no significant difference between the correlations in adult and child groups, $Z = -1.56$, $p = .118$. This may indicate that although children show a moderate negative correlation, and young adults show a weak and non-significant negative correlation, the statistical analysis indicates that these differences are not substantial enough to be considered distinct from each other. This could imply similar underlying neural mechanisms regarding these processes between children and young adults.

Taken together, these results indicate at distinct reinstatement-activation neural relationship during retrieval in children and young adults. While less differentiated detail-rich neural reinstatement is related to higher neural engagement in adults without observable gist-like reinstatement, in children lower differentiated neural reinstatement goes in hand with higher gist-like more generic reinstatement. This may indicate a differential consolidation-related functional neural reorganization of memories in child and adult groups.

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*. https://doi.org/10.3389/fninf.2014.00014

Andersson, J. L. R., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage*, *20*(2), 870–888. https://doi.org/10.1016/S1053-8119(03)00336-7

AVANTS, B., EPSTEIN, C., GROSSMAN, M., & GEE, J. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. https://doi.org/10.1016/j.media.2007.06.004

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/JSS.V067.I01

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, *37*(1), 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042

Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine*, *10*(4–5), 171–178. https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L

Criss, A. H. (2010). Differentiation and response bias in episodic memory: Evidence from reaction time distributions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 484–499. https://doi.org/10.1037/a0018435

Esteban, O., Blair, R., Markiewicz, C. J., Berleant, S. L., Moodie, C., Ma, F., & Isik, A. I. (2018). *fMRIPrep 22.0.0.*

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Evans, A. C., Janke, A. L., Collins, D. L., & Baillet, S. (2012). Brain templates and atlases. *NeuroImage*, *62*(2), 911–922. https://doi.org/10.1016/j.neuroimage.2012.01.024

Fonov, V., Evans, A., McKinstry, R., Almli, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, *47*, S102. https://doi.org/10.1016/S1053-8119(09)70884-5

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions. *Annual Review of Psychology*, *67*(1), 641–666. https://doi.org/10.1146/annurev-psych-122414-033645

Fudenberg, D., Newey, W., Strack, P., & Strzalecki, T. (2020). Testing the drift-diffusion model. *Proceedings of the National Academy of Sciences*, *117*(52), 33141–33148. https://doi.org/10.1073/pnas.2011446117

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Frontiers in Neuroinformatics*, *5*. https://doi.org/10.3389/fninf.2011.00013

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., …

Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*(1), 160044. https://doi.org/10.1038/sdata.2016.44

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, *48*(1), 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, *17*(2), 825–841. https://doi.org/10.1006/nimg.2002.1132

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–156. https://doi.org/10.1016/S1361-8415(01)00036-6

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/JSS.V082.I13

Lanczos, C. (1964). Evaluation of Noisy Data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, *1*(1), 76–85. https://doi.org/10.1137/0701007

Lerche, V., & Voss, A. (2019). Experimental validation of the diffusion model based on a slow response time paradigm. *Psychological Research*, *83*(6), 1194–1209. https://doi.org/10.1007/s00426-017-0945-8

Palada, H., Neal, A., Vuckovic, A., Martin, R., Samuels, K., & Heathcote, A. (2016). Evidence accumulation in a complex task: Making choices about concurrent multiattribute stimuli under time pressure. *Journal of Experimental Psychology: Applied*, *22*(1), 1–23. https://doi.org/10.1037/xap0000074

Patriat, R., Reynolds, R. C., & Birn, R. M. (2017). An improved model of motion-related signal changes in fMRI. *NeuroImage*, *144*, 74–82. https://doi.org/10.1016/j.neuroimage.2016.08.051

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, *84*, 320–341. https://doi.org/10.1016/j.neuroimage.2013.08.048

Pruim, R. H. R., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage*, *112*, 267–277. https://doi.org/10.1016/j.neuroimage.2015.02.064

Ratcliff, R., Love, J., Thompson, C. A., & Opfer, J. E. (2012). Children Are Not Like Older Adults: A Diffusion Model Analysis of Developmental Changes in Speeded Responses. *Child Development*, *83*(1), 367–381. https://doi.org/10.1111/j.1467-8624.2011.01683.x

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of aging and IQ on item and associative memory. *Journal of Experimental Psychology: General*, *140*(3), 464–487. https://doi.org/10.1037/a0023810

Reuter, M., Rosas, H. D., & Fischl, B. (2010). Highly accurate inverse consistent registration: A robust approach. *NeuroImage*, *53*(4), 1181–1196. https://doi.org/10.1016/j.neuroimage.2010.07.020

Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., Eickhoff, S. B., Hakonarson, H., Gur, R. C., Gur, R. E., & Wolf, D. H. (2013). An improved framework for confound regression and filtering for control of motion artifact

in the preprocessing of resting-state functional connectivity data. *NeuroImage*, *64*, 240–256. https://doi.org/10.1016/j.neuroimage.2012.08.052

Stoffel, M. A., Nakagawa, S., & Schielzeth, H. (2021). partR2 : partitioning R ^{2} in generalized linear mixed models. *PeerJ*, *9*, e11414. https://doi.org/10.7717/peerj.11414

Turker, H. B., & Swallow, K. M. (2022). Diffusion Decision Modeling of Retrieval Following the Temporal Selection of Behaviorally Relevant Moments. *Computational Brain & Behavior*, *5*(3), 302–325. https://doi.org/10.1007/s42113-022-00148-z

Tustison, N. J., Avants, B. B., Cook, P. A., Yuanjie Zheng, Egan, A., Yushkevich, P. A., Gee, J. C., Zheng, Y., Egan, A., Yushkevich, P. A., Gee, J. C., Yuanjie Zheng, Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320. https://doi.org/10.1109/TMI.2010.2046908

Wagenmakers, E.-J., Van Der Maas, H. L. J., & Grasman, R. P. P. P. (2007a). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*(1), 3–22. https://doi.org/10.3758/BF03194023

Wagenmakers, E.-J., Van Der Maas, H. L. J., & Grasman, R. P. P. P. (2007b). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*(1), 3–22. https://doi.org/10.3758/BF03194023

Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, *20*(1), 45–57. https://doi.org/10.1109/42.906424

Zhou, J., Osth, A. F., Lilburn, S. D., & Smith, P. L. (2021). A circular diffusion model of continuous-outcome source memory retrieval: Contrasting continuous and threshold accounts. *Psychonomic Bulletin & Review*, *28*(4), 1112–1130. https://doi.org/10.3758/s13423-020-01862-0