OXFORD

ORIGINAL ARTICLE

The Importance of Knowing When You Don't Remember: Neural Signaling of Retrieval Failure Predicts Memory Improvement Over Time

Yana Fandakova[1,2], Silvia A. Bunge[2,3], Carter Wendelken[2], Peter Desautels[1], Lauren Hunter[1], Joshua K. Lee[1,4] and Simona Ghetti[1,4]

[1]Center for Mind and Brain, University of California, Davis, CA 95616, USA, [2]Helen Wills Neuroscience Institute, University of California, Berkeley, CA 94720, USA, [3]Department of Psychology, University of California, Berkeley, CA 94720, USA and [4]Department of Psychology, University of California, Davis, CA 95616, USA

Address correspondence to Yana Fandakova and Simona Ghetti, Center for Mind and Brain, University of California, 267 Cousteau Place, Davis, CA 95616, USA. Email: fandakova@ucdavis.edu (Y.F.) and sghetti@ucdavis.edu (S.G.)

Abstract

Just as the ability to remember prior events is critical for guiding our decision-making, so too is the ability to recognize the limitations of our memory. Indeed, we hypothesize that neural signaling of retrieval failure promotes more accurate memory judgments over time. To test this hypothesis, we collected longitudinal functional magnetic resonance imaging data from 8 to 9 years olds, 10 to 12 years olds, and adults, with two time points spaced approximately 1.4 years apart (198 scan sessions in total). Participants performed an episodic memory retrieval task in which they could either select a response or report uncertainty about the target memory detail. Children who engaged anterior insula more strongly during inaccurate or uncertain responses exhibited greater longitudinal increases in anterior prefrontal cortex activation for decisions to report uncertainty; both of these neural variables predicted improvements in episodic memory. Together, the results suggest that the brain processes supporting effective cognitive control and decision-making continue to develop in middle childhood and play an important role for memory development.

Key words: anterior prefrontal, children, fMRI, long-term memory, metacognition

Introduction

We routinely use our episodic memories to guide decisions and goal-directed behavior. As such, effective decision-making hinges on the ability to gauge the accuracy of these memories. To understand the neural substrates of episodic memory, research has focused predominantly on the process of recollection supporting successful retrieval of episodic details (e.g., Yonelinas 2002). However, given that memory errors are ubiquitous, it is also important to understand how individuals make decisions in the face of retrieval failure.

When retrieval is unsuccessful, subjective feelings of failure or uncertainty constitute potentially powerful cues for assessing ongoing performance (Nelson and Narens 1990), and for biasing our decision-making towards withholding a response or seeking additional information (Koriat and Goldsmith 1996). Recognizing the failure to recollect target information from memory may protect against memory distortion and may be particularly consequential in high-stakes situations, such as when eyewitnesses render testimony or students take exams.

Little is known about the brain mechanisms that underlie the decision to report retrieval failure, but brain regions supporting cognitive control are excellent candidates. Lateral prefrontal and parietal regions have been broadly implicated in the monitoring and control of thoughts and actions (e.g., Stuss and Knight 2012), including postretrieval monitoring and the evaluation of accumulated information during memory retrieval (Fleck et al. 2006; Gilboa et al. 2006; Cabeza et al. 2008).

More specifically, the anterior prefrontal cortex (APFC) has been related to metacognitive processing and the integration of information regarding ongoing performance (Stuss and Alexander 2007; Burgess and Wu 2012; Fleming and Dolan 2012). For example, individual differences in uncertainty monitoring of perceptual decisions have been associated with structural and functional variability in the APFC (Fleming et al. 2010, 2012). The APFC might, therefore, be essential to guide the decision to report uncertainty during memory retrieval (Koechlin and Hyafil 2007). However, little is known about what kinds of signals contribute to this decision.

Regions in the anterior insula (AI), and possibly the anterior cingulate cortex (ACC), may provide an initial signal that recollection of target information is not successful, thereby mobilizing additional evaluation and control processes (e.g., Medford and Critchley 2010; Menon and Uddin 2010). The AI and the ACC play a critical role in the monitoring of errors and ongoing performance (Braver et al. 2001; Medford and Critchley 2010; Ullsperger et al. 2010; Bastin et al. 2016). These areas have been implicated in the maintenance of task-relevant information (Dosenbach et al. 2008) as well as in detecting salient information and initiating cognitive control operations in lateral prefrontal and parietal regions (Menon and Uddin 2010). Critically, the AI may be particularly important for providing initial signals about relevant events, such as errors, that are further processed in cingulate and prefrontal areas (Singer et al. 2009; Ham et al. 2013; Bastin et al. 2016).

Here, we explore the hypothesis that the AI, together with cingulate regions, provides signals that episodic retrieval may fail to yield the target details, and that these signals mobilize APFC to guide the decision to claim uncertainty, rather than risking making a mistake. One powerful way to test this hypothesis is to take a developmental approach and ask whether the development of the neural correlates of retrieval failure signaling precede and, in fact, predict the development of control signals to guide decisions to withhold inaccurate responses. These questions can be addressed with a longitudinal study testing the same individuals over time during middle childhood, when substantial changes in memory and metacognition occur (Ghetti and Bunge 2012; Luna et al. 2015).

Evidence from behavioral studies is consistent with our hypothesis. Even preschoolers at times exhibit behaviors consistent with the monitoring of retrieval failure, such as hesitating when facing difficult decisions or responding more slowly for incorrect memory judgments (Lyons and Ghetti 2013; Hembacher and Ghetti 2014). The capacity to effectively regulate memory to avoid error, in contrast, appears to be robust later in childhood (Ghetti et al. 2010a; Koriat et al. 2014). An open question is whether the neural processes underlying retrieval failure emerge earlier in development than those underlying memory regulation, and whether these neural processes support memory development.

In this study, we used a longitudinal design combining behavioral and neuroimaging assessments to test the hypothesis that experiencing failure to retrieve target information prompts the development of control signals supporting the decision to report uncertainty, which in turn guides memory improvement over time. At the initial time point (T_1), we used functional magnetic resonance imaging (fMRI) to examine neural activity during the retrieval phase of an episodic memory task in 44 younger children (ages 8–9), 45 older children (ages 10–12), and 30 adults (ages 18–25). The majority of the participants performed the same task at a second time point (T_2) approximately 1.4 years later (37 younger children, 40 older children, 19 adults; 198 scan sessions total).

Participants encoded a series of items, each shown against a background of one of three possible scenes (Fig. 1*a*). The scene with which an item was presented constituted the target memory detail. Specifically, at retrieval, participants were presented with studied and novel items, and were instructed to indicate the scene with which the item had been paired at encoding, or —if they did not remember having seen the item before—to indicate that it was a novel stimulus (Fig. 1*b*). Critically, participants could select a "not sure" response option if they recognized the item as studied, but did not remember which scene it had been paired with. We refer to responses on which participants accurately identified the scene paired with an item as "correct source" responses, because participants accurately identified the contextual origin of their item memory. Correspondingly, we refer to responses on which participants selected the wrong scene as "incorrect source" responses, and to responses on which participants decided to report not sure instead of committing to a particular scene as "not sure source" responses.

Brain activity associated with not sure source responses was expected to include brain regions involved in both recollection failure and decisions to report uncertainty. To distinguish between these, we assessed the extent to which activity for incorrect source responses approximated those for not sure source responses. We reasoned that similar activity levels across not sure and incorrect source responses would be indicative of situations in which episodic retrieval fails and additional processing is needed. In contrast, we reasoned that enhanced activity for not sure relative to incorrect responses would be indicative of decision processes implemented in response to retrieval failure, because in our paradigm the option to report uncertainty competes with the option to commit to a particular scene on every trial.

Materials and Methods

Participants

This study included 119 participants at T_1: 44 younger children (29 females, M_{age} = 8.99 years, SD_{age} = 0.64, 8–9.9 years), 45 older children (21 females, M_{age} = 10.73 years, SD_{age} = 0.55, 10–11.99 years), and 30 adults (17 females, 18–25 years). At T_2, 37 younger children performed the task again 0.7–2.7 years after their T_1 assessment (24 females, average T_1–T_2 delay = 1.31 years; 8.7–11.9 years at T_2). Of the older children, 40 performed the task again at T_2 0.8–2.9 after their T_1 assessment (17 females, average T_1–T_2 delay = 1.40 years; 10.9–13.9 years at T_2). Of the adults, 19 returned 0.77–2.80 years later for T_2 (11 females, average T_1–T_2 delay = 1.44 years; 19–23 years at T_2). Of the children who returned for a T_2 assessment, 30 younger children and 32 older children had fMRI data of high enough quality to be included in the longitudinal fMRI analyses. There were no differences between participants who returned for T_2 compared with those who did not in source memory performance or in a standardized measure of memory functioning at T_1 (i.e., WRAML2;
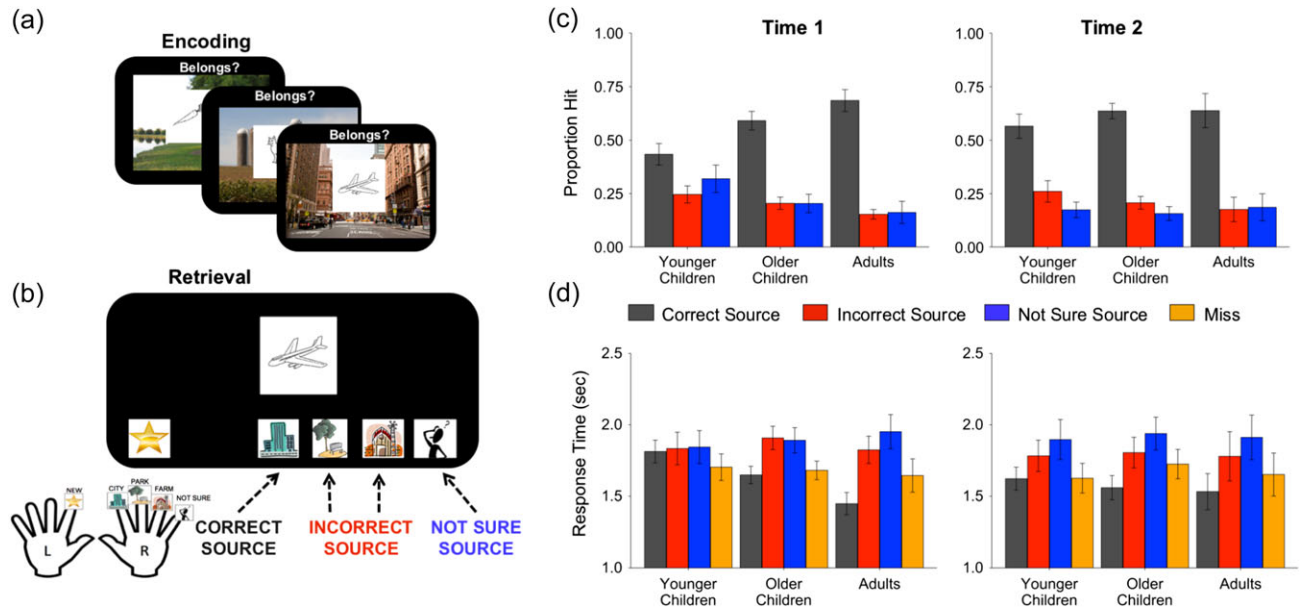
Figure 1. Experimental procedure and behavioral results. (*a*) At encoding, participants were instructed to remember the item-scene association between a picture drawing and the scene of a farm, a park or a city with which it was presented. (*b*) At retrieval, participants were presented with studied and novel items and indicated the scene that the item was studied with (by pressing the "park", "city", or "farm" button) or if they were unsure about the scene with which the item was studied (by pressing the not sure button, indicated by the uncertain character). Participants had to press the "new" button (indicated by a star) if they thought that the item was novel. (*c*) Age differences in proportions correct, incorrect, and not sure source responses at T_1 and T_2 for younger children (N T_1 = 44, N T_2 = 37), older children (N T_1 = 45, N T_2 = 40), and adults (N T_1 = 30, N T_2 = 19). T_1 and T_2 were spaced approximately 1.4 years apart. (*d*) Age differences in RTs for correct, incorrect, not sure source, and miss responses at T_1 and T_2. Data represented as mean ± SEM.

Sheslow and Adams 2003, Ps > 0.20). All children and adults included in this study were within the normal range of memory performance for their corresponding age group. The UC Davis Institutional Review Board approved the study. Informed consent was obtained from all participants and their parents.

Task Procedure

Participants performed three interleaved encoding and retrieval runs of a source memory task (Fig. 1*a*). Each encoding run included 48 picture drawings of an object or an animal presented on a scene background of a farm, a park, or a city. Participants were instructed to remember the drawing-scene pairs and to make a semantic judgment about whether the drawing belonged to the corresponding scene. Each retrieval run included the 48 drawings from the preceding encoding run along with 16 novel drawings that had not been seen before. At retrieval, each drawing was presented on a black background for 4000 ms, followed by a jittered fixation period (500–8000 ms). Participants were instructed to decide if a drawing was previously studied (i.e., seen in the preceding encoding run) or if it was new (i.e., never seen before). For drawings judged new, participants were to press the "new" button (indicated by a star) with their left hand. For drawings judged studied, participants were to indicate the scene with which the drawing was studied (by pressing the "park", "city", or "farm" button with their right hand) or to indicate that they did not remember the particular scene with which the drawing had appeared (by pressing the not sure button indicated by an uncertain cartoon character with their right hand; Fig. 1*b*). For instance, in the example depicted in Figure 1*a,b*, selecting the city button when presented with the drawing of a plane constituted a correct source response in which the item-scene pair was successfully retrieved. If instead participants chose the park or farm button,

this constituted an incorrect source response in which the item, but not the scene, was retrieved. If participants chose the not sure button, this constituted a not sure source response in which the item was successfully retrieved, but participants chose to report uncertainty about the scene with which it was studied. Neuroimaging data from the study phase of the source memory task were not considered for the purposes of this paper.

FMRI Data Acquisition

Data were collected with a Siemens 3 T MRI scanner using a 32-channel head coil at both time points. Functional images were acquired with a gradient EPI sequence (repetition time [TR] = 2000 ms, echo time [TE] = 23 ms, no interslice gap, field of view (FOV) = 204 mm, 37 slices per volume, voxel size = 3 mm isotropic). A T_1-weighted magnetization prepared rapid gradient echo (MP-RAGE) was acquired for co-registration of the functional images (TR = 2500 ms, TE = 3.24 ms, FOV = 224 mm, voxel size = 0.7 mm isotropic).

FMRI Data Analyses

Data were analyzed using SPM8 (www.fil.ion.ucl.ac.uk/spm). Preprocessing included slice-time correction, realignment to the first volume using rigid body motion correction with sinc interpolation, co-registration to MP-RAGE, and smoothing with a 6-mm FWHM isotropic Gaussian kernel. Volumes with head motion >1 mm or signal change >2% were replaced with interpolated values using ArtRepair (http://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html). If more than 25% of the volumes in one scan were replaced, the corresponding scan was excluded from further analyses. Participants with fewer than two valid functional scans were not included in

analyses. Among the participants included in analyses, there were reliable age differences in head movement, $F(2,117) = 15.74$, $P < 0.05$, $\eta_p^2 = 0.22$. Children displayed significantly more head movement than adults, and had a larger number of repaired scan volumes (adults vs. younger children $P < 0.05$, $d = 1.48$, adults vs. older children $P < 0.05$, $d = 1.15$). By contrast, there were no differences in head movement between younger and older children, $P = 0.22$, $d = 0.34$. Given the differences between children and adults, but not between younger and older children, we included the number of repaired scans as a covariate in the analyses including adults.

At T_1, a total of 38 additional participants were tested, but excluded from the analyses reported here due to chance behavioral performance (7 younger children, 2 older children, 1 adult), not using all response options (1 younger child, 2 older children), or motion (17 younger children, 8 older children). Out of the 77 children who returned for T_2, 15 children were excluded from neuroimaging analyses because they only completed behavioral assessments (4 younger children, 6 older children) or due to motion (3 younger children, 2 older children).

Univariate general linear model analysis
General linear models were estimated for correct source, incorrect source, and not sure source responses. These events were convolved with a canonical hemodynamic response function that modeled the duration of the event from event onset until the participant's response. Misses, correct rejections, false alarms, and trials with no response were modeled as events of no interest. Motion parameters were included in the model to regress out effects of head motion. Second-level group analyses with participants as a random factor were used to compare parameter estimates for not sure source > correct source in order to identify the regions supporting reports of uncertainty in memory outcomes. These analyses were performed across all participants (i.e., collapsing across age groups) and across time points (i.e., collapsing across T_1 and T_2). To account for the fact that some children ($n = 27$) and adults ($n = 13$) did not participate or provide usable data at T_2, we introduced subject-specific regressors modeling the time points available for each participant. Results were corrected for multiple comparisons using a cluster-based error correction at $P < 0.05$, with a height threshold of $P < 0.001$. Brain schematics were visualized with BrainNet Viewer (Xia et al. 2013).

ROI Analysis
Marsbar (http://marsbar.sourceforge.net/) was used to carry our ROI analyses. The ROIs included the bilateral AI, the bilateral APFC, the ACC, and the right posterior parietal cortex (PPC). Clusters identified from the not sure source > correct source whole-brain contrast across all participants and across time points were masked with anatomical masks from the AAL atlas corresponding to our preselected ROIs, thereby restricting our functional ROIs to the anatomical regions of a priori interest at the group level. Thus, the AI ROI was created using anatomical masks of left and right insula; the APFC ROI was created by masking whole-brain activity with left and right middle frontal gyrus; the ACC ROI was created by using a mask of the ACC; and the PPC ROI was created by masking whole-brain activity with inferior parietal lobe and supramarginal gyrus. There were no unique effects involving hemisphere (see Supplementary Figure S2c–d); therefore, we analyzed activation averaged across left and right ROIs for AI and APFC. We assessed age differences in activity of these ROIs in a mixed analysis of variance

(ANOVA) with ROI (AI vs. ACC vs. APFC vs. PPC) and trial type (incorrect source vs. not sure source) as within-subject factors and age group (younger children vs. older children vs. adults) as a between-subject factor. Post-hoc tests assessed the magnitude of the difference between not sure and incorrect responses within each ROI in each age group. Although this comparison is generally more likely to detect higher values for not sure source responses, in this analysis we were particularly interested in the degree to which these brain regions varied in the magnitude of the difference between the two trial types across age groups. All reported statistical tests are two-sided. Reported post-hoc comparisons survive corrections for multiple comparisons using the Bonferroni correction.

Multiple Regression Analyses

Multiple regression analyses were performed to examine the contributions of task-related brain activity to memory performance at T_1 and to change in memory performance over time. The behavioral measure used in these analyses was the source accuracy score, computed as the number of accurate item-scene associations, relative to all source guesses (i.e., incorrect and correct source trials) and therefore did not include not sure source trials. Regression models performed across the entire child sample at T_1 examined whether source accuracy at T_1 was predicted by children's exact age and task-related activity in the corresponding ROI. Regression models examining memory improvement over time across all children examined whether change in source accuracy between T_1 and T_2 was predicted by children's age at T_1, interval between T_1 and T_2, source accuracy scores at T_1, APFC contrast for [not sure > correct and incorrect] at T_1, AI contrast for [not sure and incorrect > correct] at T_1, PPC contrast for [not sure > correct and incorrect] at T_1, ACC contrast for [not sure and incorrect > correct] at T_1, and change in APFC [not sure > correct and incorrect] between T_1 and T_2. Given the intermediate activity profile in ACC, we performed control analysis in which we instead included ACC contrast for [not sure > correct and incorrect] at T_1, and the reported results did not differ.

Latent Change Score Models

Latent change score models were used to examine the dynamics of longitudinal change in task-related activity among children. Estimating latent baseline and change scores in a structural equation modeling framework has several advantages as it allows testing hypotheses about individual differences in change, while simultaneously accounting for the effects of relevant covariates and measurement error (McArdle and Nesselroade 1994; Raz et al. 2005). Structural equation modeling relies on full information maximum likelihood estimation procedures, which allows for incomplete data, while not necessitating restrictive assumptions about the covariance structure of the data.

We constructed a latent change score model in which we estimated latent factors of AI and APFC activity for each time point. The contrast estimates for left and right AI from [not sure and incorrect source > correct source] were allowed to load on a common latent factor at T_1 and T_2 (see Supplementary Figure S3). Contrast estimates for left and right APFC from [not sure source > incorrect and correct sources] were allowed to load on common latent factors at both T_1 and T_2. Factor loadings were fixed to be the same across T_1 and T_2, residual variances were set equal across time, and intercepts were included in the

model. We estimated baseline levels at T_1 and latent change for each of the AI and APFC factors. We also included cross-time effects across regions such that baseline levels at T_1 in each region predicted latent change in both AI and APFC. Thus, APFC change was predicted by both the AI factor at T_1 and the APFC factor at T_1, while also simultaneously being correlated with change in AI activity. Participants' age was included as an additional predictor of baseline levels and change in both AI and APFC factors. In addition, individual delay between T_1 and T_2 was included as a predictor of change in AI and APFC. The resulting model was estimated using Mplus5.1 (Muthén and Muthén 1998–2007) via maximum likelihood. Model fit was assessed with several indices including χ^2/(degrees of freedom) < 2, comparative fit index >0.90 (CFI), and a root-mean-square error of approximation (RMSEA) ≤0.08 that included 0.05 within its 90% confidence interval (RMSEA). Cross-time effects were tested by restricting the corresponding paths to zero and comparing the model fit of the resulting nested models to the freely estimated models via the χ^2 statistic with degrees of freedom equal to the difference in the number of free parameters. Additional control models are reported in Supplementary Results.

Functional Connectivity

Beta correlations (Rissman et al. 2004) were used to examine connectivity among the brain regions involved in not sure source responses. The canonical HRF was fit to each occurrence of not sure source responses and parameter estimates were sorted to derive condition-specific beta-series for each ROI, represented as 5 mm spheres centered at activation maxima of the ROIs at the group level. The estimated connectivity across AI-APFC ROIs was calculated by averaging the corresponding connectivity coefficients for each subject. The same procedure was applied to estimate average functional connectivity during not sure source responses at T_1 and T_2.

Results

Age Differences and Longitudinal Change in Source Memory Accuracy

A comparison of younger children (8–9 years), older children (10–12 years), and adults at T_1 revealed that source memory accuracy (i.e., the number of items for which the corresponding scene was correctly identified, divided by the total number of correctly recognized items) increased across age groups, $F_{(2,116)} = 26.39$, $P = 0.0001$, $\eta_p^2 = 0.31$ (Fig. 1c). Younger children exhibited significantly lower source memory accuracy than older children, who in turn exhibited lower source memory accuracy than adults (Ps < 0.05). Across all children, source memory accuracy increased with age, $r = 0.43$, $P = 0.0001$.

At T_1, the age groups also differed in the likelihood of choosing the not sure source response, $F_{(2,116)} = 8.88$, $P = 0.0001$, $\eta_p^2 = 0.13$, as well as in their source incorrect responses, $F_{(2,116)} = 7.16$, $P = 0.001$, $\eta_p^2 = 0.11$ (Fig. 1c). Younger children were more likely to report source uncertainty and to make incorrect source judgments than both older children and adults (Ps < 0.01), who did not differ from each other (Ps > 0.20). To confirm that uncertainty responses reflected metacognitive evaluations and not age differences in item memory strength, we performed additional control analyses showing that there was no age difference in participants' use of the not sure source option when making false alarms to new items (see Supplementary Materials).

The use of the not sure source response option can be further characterized by calculating a not sure selection score as the probability that participants selected the not sure source response when they failed to recollect the scene with which an item was presented at study (i.e., across incorrect and not sure source responses). Thus, this score represents participants' tendency to report uncertainty when recollection of target information failed, while accounting for individual differences in memory accuracy. The not sure selection score did not differ among age groups at T_1 ($P = 0.41$), suggesting that, in aggregate, children and adults used the not sure source response option similarly when source retrieval failed, even though adults exhibited higher memory accuracy overall (see Supplementary Results for additional behavioral results).

Improvements in source memory accuracy between T_1 and T_2 differed as a function of age group, as revealed by a significant age group by time point (T_1 vs. T_2) interaction, $F_{(2,93)} = 11.89$, $P = 0.0001$, $\eta_p^2 = 0.20$ (Fig. 1c). Stronger improvements in source memory accuracy were observed in younger children, $F_{(1,36)} = 25.12$, $P = 0.0001$, $\eta_p^2 = 0.41$, who at T_2 performed similarly to older children at T_1 ($P > 0.45$). Smaller longitudinal improvements were observed in older children, $F_{(1,39)} = 3.62$, $P = 0.065$, $\eta_p^2 = 0.09$. In contrast, no improvements were observed in adults.

These results suggest that source memory accuracy continues to develop during middle and late childhood. The fact that the adults did not improve from T_1 to T_2 suggests that the memory improvements in younger and older children do not reflect practice effects with the task, but rather a developmental change in episodic memory. Furthermore, an analysis examining the effects of time interval on change in performance revealed that children improved more when the delay between T_1 and T_2 was longer, which is the opposite of the pattern expected for practice-related improvements (see Supplementary Materials).

Age Differences and Longitudinal Change in Response Times

Signals of retrieval failure are expected for both incorrect and not sure source responses; in both cases further processing should occur (Kelley and Lindsay 1993), and thus these responses were expected to be slower in response times (RTs) than correct source decisions. Incorrect and not sure source responses were also expected to be slower than miss responses because, although inaccurate, miss responses involve judgments of lack of item memory, without the ensuing need to select among scene options. At T_1, we observed a main effect of trial type (correct source vs. incorrect source vs. not sure source vs. miss) on RTs, $F_{(3, 114)} = 45.82$, $P < 0.001$, $\eta_p^2 = 0.55$, which was qualified by an interaction with age group, $F_{(6, 230)} = 8.90$, $P < 0.001$, $\eta_p^2 = 0.19$ (Fig. 1d). In older children and adults, not sure source responses were slower than correct source and miss responses (Ps < 0.05), and were comparable to incorrect source responses (Ps > 0.05). In younger children, not sure source responses were comparable to incorrect and correct source responses (Ps > 0.50), but were significantly slower than miss responses (Ps < 0.05), suggesting greater overall hesitation and slowing when younger children engaged in a source decision compared with when they judged lack of item memory.

At T_2, we again observed a main effect of trial type on RTs, $F_{(3, 89)} = 45.67$, $P < 0.001$, $\eta_p^2 = 0.61$. However, the trial type × age group interaction was not reliable ($P = 0.23$) with all age groups responding more slowly for incorrect and not sure source responses relative to both correct source and miss responses

(Ps < 0.05; Fig. 1*d*). Taken together, these behavioral results support longitudinal improvements in source memory, and indicate that participants of all age groups exhibited greater hesitation and slowing when making an incorrect or not sure source response at T_2.

Age Differences in Brain Activity Underlying Not Sure Source Responses

Next, we examined whole-brain activity for not sure source responses relative to correct source responses, collapsing across time points and ages. This whole-brain contrast revealed clusters in regions of a priori interest in bilateral AI, APFC, ACC and right PPC, and additional activations in temporal and orbitofrontal regions (Fig. 2). The ACC cluster corresponded to a ventral ACC area that overlaps partially with the dorsal ACC region that has been associated to performance monitoring (e.g., Ridderinkhof et al. 2004), and is similar to a ventral region that has been implicated in confidence judgments of retrieval (e.g., Chua et al. 2006; Hebscher et al. 2015).

Next, we examined age and trial type differences across ROIs at T_1. To distinguish recollection failure from the decision to report uncertainty, we excluded source correct trials and investigated the relative differences between not sure and incorrect source trials across ROIs across age groups. A mixed ANOVA with age group (younger children vs. older children vs. adults), trial type (not sure source vs. incorrect source) and ROI (AI vs. ACC vs. APFC vs. PPC) revealed a trial type × ROI interaction, $F_{(3, 108)} = 8.997$, $P = 0.0001$, $\eta_p^2 = 0.20$, and an age group × trial type × ROI interaction, $F_{(6, 218)} = 2.11$, $P = 0.05$, $\eta_p^2 = 0.06$ (including head movement covariate: $P = 0.06$, $\eta_p^2 = 0.06$), indicating that activity profiles across ROIs differed as a function of both trial type and age group. To investigate these interaction effects further, we examined each ROI separately.

In the APFC, the main effect of trial type was significant, $F_{(1,113)} = 9.04$, $P = 0.02$, $\eta_p^2 = 0.07$, indicating enhanced activity on trials in which participants reported uncertainty about source relative to incorrect source trials (Fig. 3*a*). Critically, we

observed a significant age group × trial type interaction, $F_{(2,113)} = 3.697$, $P = 0.03$, $\eta_p^2 = 0.06$ (including head movement covariate: $P = 0.045$, $\eta_p^2 = 0.05$), such that enhanced activity on not sure source trials relative to incorrect source trials was evident only in older children, $P = 0.001$, and adults, $P = 0.05$. In younger children, there were no differences between incorrect and not sure source responses, $P = 0.99$. Together, these findings suggest that the APFC is involved in the decision to report uncertainty, and that this function is not fully developed until middle childhood.

AI activity was comparable for not sure source and incorrect source trials, in which participants did not successfully retrieve the source (Fig. 3*b*). There were no main effects of trial type or interactions with age group in this ROI (Ps > 0.50). Supplementary whole-brain analyses revealed enhanced activity in bilateral AI for incorrect source responses relative to correct source responses (see Supplementary Figure S1*b*). Together, these results are consistent with the interpretation that the AI supports recollection failure signaling across the three age groups.

PPC activity was enhanced for not sure source relative to incorrect source responses ($P = 0.0004$; Fig. 3*c*) as in the APFC, but unlike the APFC, there were no age differences in activity in this region ($P > 0.50$). Finally, the ACC showed an attenuated difference between not sure and incorrect source responses, $P = 0.06$ (Fig. 3*d*), with no age group × trial type interaction ($P > 0.50$).

Overall, the key findings from these analyses were 2-fold. First, the APFC was engaged more strongly when participants indicated that they were uncertain about the source with which an item was studied than when they made an incorrect source judgment or selected the correct source, consistent with engagement in decision-making so as to prevent inaccurate reports. This APFC activity profile was not evident until age 10–12 years, consistent with protracted development of this region. Second, AI activity was enhanced when participants failed to retrieve the correct source—that is, for both incorrect and not sure source responses—compared with when they selected the correct source response. The AI activity profile, consistent with a role in signaling recollection failure, was observed in all age groups.

These results are consistent with our proposal that these brain regions play different roles in uncertainty monitoring and follow different developmental trajectories, but more direct evidence will come from analyses examining longitudinal change in these areas.
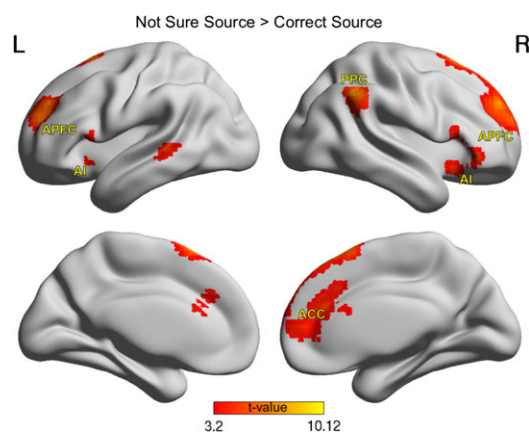

Figure 2. Brain activity observed during reporting of uncertainty in source memory retrieval. Voxel-wise maps represent the difference between not sure source and correct source trial activity, thresholded to display significant clusters (voxel-wise $P < 0.001$, cluster-corrected $P < 0.05$) across all participants at T_1 ($N = 119$) and T_2 ($N = 79$), including subject-specific terms to account for individual repeated observations. For results of the opposite contrast, see Supplementary Figure S1*a*. MNI space, L = left, R = right, AI = anterior insula, ACC = anterior cingulate cortex, APFC = anterior prefrontal cortex, PPC = posterior parietal cortex.

Activity During Retrieval Failure Contributes to Concurrent Memory Performance

Next, we performed multiple regression analysis including age and ROI activity at T_1 as predictors of source accuracy at T_1. To understand the mechanisms of source memory development, these analyses focused on children, whose behavioral performance showed improvement over time along with marked individual differences. The overall model was significant, $F_{(5,83)} = 2.81$, $P = 0.02$, $R_{adj}^2 = 0.10$. Recollection failure activity in the AI (indicated by the contrast [not sure and incorrect source > correct source]) predicted memory accuracy at T_1 across all children, $\beta = 0.35$, $P = 0.01$. Enhanced activity in APFC and PPC associated with the decision to report source uncertainty (indicated by the contrast [not sure source > correct and incorrect source]) was not related to memory performance at T_1; nor was ACC activity for either of the two contrasts (Ps > 0.50).
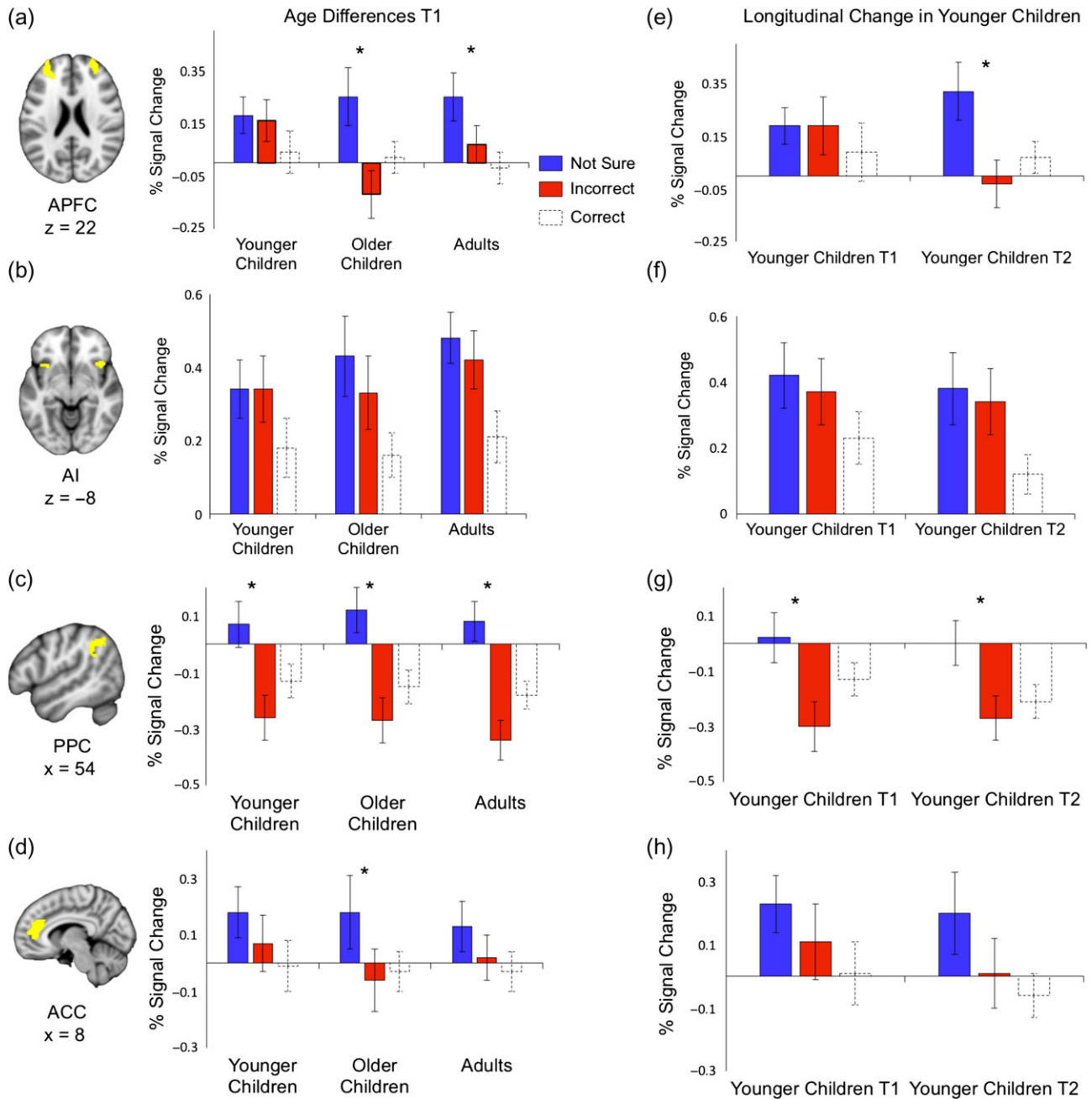
Figure 3. Age differences in task-related activity across trial types and age groups. (a–d) Activity differences across correct source, incorrect source, and not sure source trials in younger children, older children, and adults at T_1 in (a) bilateral APFC, (b) bilateral AI, (c) right PPC, and (d) ACC. (e–h) Activity differences across correct source, incorrect source, and not sure source trials at both T_1 and T_2 for those younger children who provided neuroimaging data at both time points (N = 30) in (e) APFC, (f) AI, (g) PPC, and (h) ACC. There were no differences by hemisphere in APFC and AI at both time points (see Supplementary Figure 2c–d), and data were collapsed across left and right. For corresponding time courses from these ROIs, see Supplementary Figure 2a,b. Brain images depict regions of interest from which signal change was obtained. Data represented as mean ± SEM. MNI space, *$P < 0.05$.

Longitudinal Change in APFC Activity Underlying Decisions to Report Uncertainty

Given that younger children showed a different activation profile in the APFC than either older children or adults, we predicted that APFC activation would change the most in younger children over time, such that at T_2, when their age approximated that of older children at T_1, they would show a difference between not sure and incorrect source trials. Indeed, the increase in the difference in APFC activity between not sure

and incorrect source trials was more pronounced in younger children than in older children, as revealed by a time point (T_1 vs. T_2) × trial type (incorrect vs. not sure source) × age group (younger children vs. older children) interaction, $F_{(1, 59)} = 5.63$, $P = 0.02$, $\eta_p^2 = 0.09$ (Fig. 3e). In younger children, APFC activity difference between not sure and incorrect source trials increased over time and at T_2 became indistinguishable from the activation pattern of older children at T_1 ($P = 0.80$), indicating a developmental change in this brain region over time. No reliable age differences or longitudinal changes were observed in any of the

other ROIs (Ps > 0.40, Fig. 3f–h). Taken together, the longitudinal results support the conclusions from the cross-sectional age group comparisons, showing a developmental change in APFC activity over the time period examined here.

Neural Activity Underlying Retrieval Failure and Uncertainty Reporting Predicts Memory Improvement Over Time

Next, we tested the hypothesis that neural activity associated with retrieval failure signaling and uncertainty reporting predicted developmental improvements in source memory accuracy. Given that activation in the APFC showed age-related change over this developmental window, we sought to determine whether this change contributed to behavioral improvements in source memory occurring during the same period. We conducted a multiple regression analysis in which we predicted improvement in source memory accuracy from T_1 to T_2 from AI activity for recollection failure at T_1 (indicated by the contrast [not sure and incorrect source > correct source]), APFC activity for source uncertainty at T_1 (indicated by the contrast [not sure source > correct and incorrect source]), and change in APFC activity for source uncertainty from T_1 to T_2. Age at T_1, time interval between T_1 and T_2, source accuracy at T_1, PPC and ACC activity at T_1 were also included in the model. The overall model was significant, $F_{(8,50)} = 5.89$, $P = 0.0001$, $R_{adj}^2 = 0.40$. Two brain variables emerged as significant predictors of memory improvement: AI activity at T_1, $\beta = 0.38$, $P = 0.006$ (Fig. 4a), and change in APFC activity over time, $\beta = 0.32$, $P = 0.037$ (Fig. 4b), predicted source accuracy improvement. PPC and ACC activities were not associated with change in source accuracy over time (Ps > 0.20).

AI Activity at T_1 Predicts Longitudinal Changes in APFC Activity

Having established that both enhanced AI activity underlying the failure to recollect target details from memory as well as longitudinal change in APFC activity associated with reporting of uncertainty predicted children's source memory improvement, we next probed the nature of the developmental relation between task-related activity in these brain regions. We hypothesized that the earlier-developing recollection failure signals in the AI support the development of control signals to guide decisions to report uncertainty (and avoid errors) in the APFC over time. We used structural equation models to test this hypothesis. This approach is well suited for testing the predicted cross-time relations between AI activation at T_1 and change in APFC activation from T_1 to T_2, while simultaneously testing the converse relation (i.e., between APFC activation at T_1 and change in AI activation from T_1 to T_2) and estimating the effects of covariates and measurement error (Fig. 4c; Supplementary Figure S3). Model fit was acceptable, $\chi^2 = 49.51$, df = 35, RMSEA = 0.068, 90% CI [0.00–0.11], CFI = 0.96. Greater activity in the AI at T_1 predicted greater increase in APFC activity over time, $\beta = 0.34$, $\chi^2(1) = 7.26$, $P = 0.007$ (Fig. 4c). In contrast, APFC activity at T_1 was not related to change in AI activity over time, $\beta = -0.15$, $\chi^2(1) = 2.3$, $P = 0.13$, and the two paths were significantly different, $\chi^2(1) = 11.03$, $P = 0.001$ (see Supplementary Results for alternative models). Finally, when change in source accuracy was added to the structural equation model, the results were replicated such that increased AI activity at T_1 remained a reliable predictor of change in APFC activity ($P = 0.002$), which in turn predicted source accuracy improvement ($P = 0.04$) over time. Taken together, children who displayed stronger AI activity at T_1
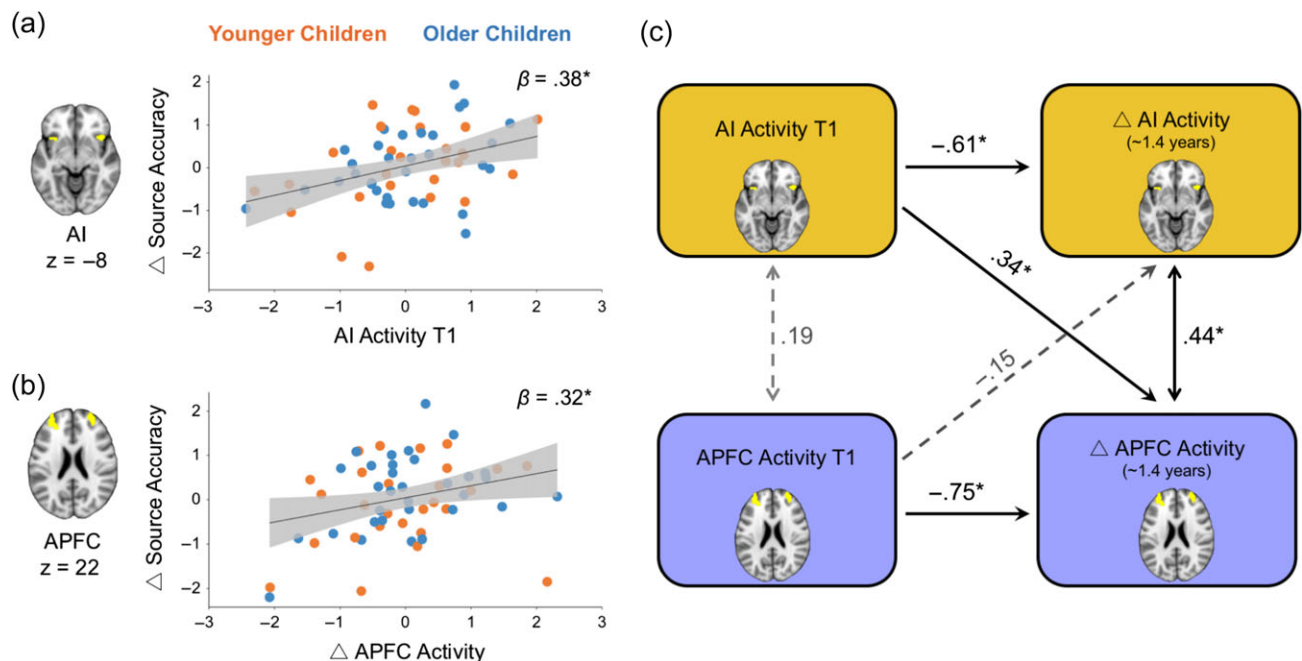

Figure 4. Relationships among AI activity, APFC activity, and source memory improvement over time. (a) AI activity at T_1 predicted source accuracy improvement over the 1.4-year period across younger and older children. (b) Change in APFC activity between T_1 and T_2 predicted source accuracy improvement over the 1.4-year period in younger and older children. Gray area represents 95% confidence interval, solid lines represent the best linear fit across all children; different colors for younger and older children only for reference. X- and Y-axes represent standardized residuals from regression models including participants' age, T_1–T_2 delay, source accuracy at T_1, ACC, PPC and APFC activity at T_1. (c) Summary of main results from structural equation model relating task-related activity at T_1 and longitudinal change across AI and APFC in all children (see Supplementary Figure 3 for full model). *P < 0.05.

demonstrated a greater increase in the selective engagement of APFC when reporting uncertainty over time.

In additional analyses, we examined the extent to which variability in functional connectivity during not sure source responses was associated with the decision to report uncertainty. At T_1, stronger connectivity between the AI and the APFC was related to a higher not sure source response rate among younger and older children, $r = 0.28$, $P = 0.01$ (see Supplementary Figure S4a). Similar relations were observed at T_2, $r = 0.26$, $P = 0.04$ (see Supplementary Figure S4b). These results were not related to age but rather to individual differences in connectivity, as average connectivity was similar across ages at both time points (Ps > 0.05; see also Supplementary Materials). Together, these results indicate that greater connectivity among the key brain regions supporting not sure source decisions was associated with an increased likelihood for reporting uncertainty in source memory at any given time point.

Discussion

The ability to evaluate the accuracy of our memories and make decisions to regulate our behavior affects daily functioning in a variety of situations, from deciding what information to study next to providing accurate eyewitness testimony in the courtroom. This ability emerges during the preschool years and becomes more robust in middle childhood, as children become able to monitor even subtle differences in memory quality (Ghetti et al. 2010a; Fandakova et al. 2013). To better understand how the capacity to assess memory accuracy affects the development of source memory, we conducted a longitudinal study that examined the neural mechanisms supporting the development of decision-making when recollection of target information fails over middle and late childhood. Our results suggest that these regulatory processes have important functional implications for memory development and highlight the utility of a developmental approach for understanding how we monitor and control our memory.

Corroborating cross-sectional research on memory development (Ghetti and Bunge 2012), our longitudinal study demonstrated that children's source memory accuracy improved over the 1.4-year period. Notably, these memory improvements were predicted by task-related activity in AI and APFC regions. Across all age groups, AI activity was enhanced when source retrieval failed, regardless of whether participants reported being uncertain. In contrast, APFC activity was enhanced for uncertain relative to correct and incorrect source trials in older children and adults, but not in younger children, indicating protracted development of the prefrontal mechanisms supporting uncertainty evaluation and cognitive control (e.g., Ofen et al. 2012; Luna et al. 2015). Importantly, over the 1.4-year period, younger children exhibited the largest improvements in memory performance, and also showed the largest increase in APFC activity associated with decisions to report uncertainty. Consistent with our hypothesis, longitudinal changes in APFC activity were predicted by enhanced AI activity at T_1. These results provide compelling new evidence that the interplay between signals of recollection failure and decision-making processes contributes to source memory development in middle childhood.

A large body of literature has examined neural mechanisms associated with memory retrieval, emphasizing recollection as the process that affords more precise memory information (Diana et al. 2007). For this reason, the neurocognitive development of recollection has been the primary focus of research in developmental cognitive neuroscience (e.g., Ofen et al. 2012; DeMaster and Ghetti 2013). However, prior work has largely neglected the potential benefits of recognizing the failure to retrieve accurate target information as an additional route to initiating regulatory behavior and maximizing memory performance. In adults, familiarity-based responses engage anterior insular and fronto-parietal regions (e.g., Kim 2010), which are thought to reflect greater demands for postretrieval monitoring when facing recollection failure (Henson et al. 2000; Badre and Wagner 2004; Fleck et al. 2006). Our research corroborates and extends these findings by demonstrating that when accurate recollection of target information fails, control and decision-making processes may operate on familiarity signals to regulate accuracy, thereby contributing to memory improvement.

In a separate line of research, the AI has been implicated in evidence accumulation during processing of ambiguous perceptual stimuli (e.g., Grinband et al. 2006; Ploran et al. 2007), error or performance monitoring (e.g., Preuschoff et al. 2008; Ullsperger et al. 2010), and decision-making under uncertainty (e.g., Critchley et al. 2001). Our results offer new insights into the functional role of the AI in episodic retrieval, supporting the idea that it may be involved in the monitoring of ongoing mnemonic processes to detect salient events that need further processing and evaluation in the context of a current goals and task demands (Menon and Uddin 2010). More broadly, our longitudinal approach provides valuable insights into the interplay between memory and decision-making processes, uncovering brain–behavior relations that would likely have gone unnoticed in a study of adults at a single time point.

Furthermore, our longitudinal evidence shows for the first time that activity associated with ongoing memory processing and failure to recollect target memory details predicts developmental change in memory and decision-making during middle childhood, a period marked by notable improvements in memory and metacognition (Ghetti and Angelini 2008; Daugherty and Ofen 2015). While such developmental links have been hypothesized based on behavioral evidence from different age groups (Hembacher and Ghetti 2014), our results suggest a possible neural mechanism by which children become sensitive to factors that influence ongoing stimulus processing such as salience, prior errors or retrieval fluency, to guide subsequent memory decisions.

Even preschoolers can monitor and report on their own uncertainty (Lyons and Ghetti 2011; Hembacher and Ghetti 2014; Vo et al. 2014) using such cues as fluency of processing or decision effort (e.g., Koriat 2007). These cues, which are often implicit, may drive behavior in the absence of overt self-reflection even in toddlers facing uncertain situations (i.e., Koriat 2007; Goupil et al. 2016). In older children and adults, however, the experience of uncertainty guides decision-making above and beyond these implicit signals (Son and Metcalfe, 2000; Hembacher and Ghetti 2016). This prior literature, together with the current results, suggests the intriguing possibility that component processes of metacognition become functional at different time points in development. From this perspective, differences in APFC activation profiles between younger children and older participants may reflect different paths to decide to report source uncertainty: whereas older children and adults may rely more precisely on the integration and evaluation of multiple cues in APFC to reach a decision, younger children may make decisions without the same degree

of integration in the APFC, responding more directly to signals of retrieval failure. This is not to say that the APFC is not involved at all in younger children, given our connectivity results implicating AI-APFC connectivity in overall tendencies to select not sure responses, but signals of retrieval failure in the AI in younger children may constitute an important stepping stone to assessing uncertainty. Based on this account, we hypothesize that younger children would perform poorly on experimental conditions that place heavy demands on evaluation and integration of multiple cues, as a result of immature APFC function. Future studies are needed to test this prediction, and to address the interactions between recollection failure and decision-making processes within the time course of individual decisions in order to pinpoint the exact mechanisms through which these processes function.

The observed APFC effects were in close proximity to areas previously associated with metacognitive judgments (Fleming et al. 2012). Recent accounts of APFC function have suggested that it may represent a primary hub of metacognitive processing (Burgess and Wu 2012) that maintains monitoring signals from insular and cingulate areas and compares them to representations of current goals or personal beliefs (Fleming and Dolan 2012). Notably, APFC activity patterns in adults were consistent with a graded response across trials types (Fig. 3a; Supplementary Figure S2b), suggesting that activity in this area may be sensitive to different levels of subjective confidence (Fleming et al. 2012). In addition, abundant literature on the cognitive mechanisms of metacognitive control has linked participants' choice to report uncertainty to shifts in decision criteria (e.g., Koriat and Goldsmith 1994, 1996; Hanczakowski et al. 2013). Indeed, criterion-setting is critical for the determination of source uncertainty (cf. Rotello and Macmillan 2007; O'Connor et al. 2010; Kantner and Lindsay 2012), and may well contribute to the APFC activity patterns observed here. Future research involving manipulations of subjective memory uncertainty could dissociate the neural mechanisms supporting the representations of uncertainty, task goals, and criterion shifts as well as their role in the selection of future actions, such as withholding of memory responses or re-studying information that one has not yet learned (Nelson and Narens 1990). One intriguing possibility is that changes in APFC activity may foster the exploration of novel or more effective strategies by integrating monitoring signals with person-specific and context-specific influences (Badre et al. 2012). These processes, while beneficial throughout the lifespan, may be particularly important during childhood, when a number of learning strategies are acquired for the first time (Siegler 1996).

Individual differences in AI activity predicted memory accuracy improvements independently and beyond individual differences in APFC activity change. This result does not require us to posit a specialized role of AI in memory; indeed, such a claim would be inconsistent with the broader literature. Individuals who show more pronounced retrieval failure activity in the AI may be less likely to base their responses on weak evidence; in our case, they may be less likely to base their responses on scene familiarity without recollecting additional, more diagnostic details. In addition, uncertain memory outcomes can potentially improve future learning and memory success through the generation of prediction error signals in the AI (Singer et al. 2009; Menon and Uddin 2010).

The PPC and ACC were also engaged when participants from all age groups reported uncertainty. Their activation profiles did not change over time and were not related to memory performance. The PPC could play an ancillary role in the decision to report uncertainty through its involvement in working memory and top-down control (D'Esposito 2007). While the exact role of PPC in evaluating source failure needs to be investigated in future research, especially given known functional heterogeneity within this region (Uncapher and Wagner 2008), one of the few studies that explored neural activity during metacognitive judgments of confidence found similar results to ours: activity in both APFC and PPC was enhanced during uncertain trials, but only activity in APFC predicted individual differences in metacognition (Fleming et al. 2012). This overlap in results suggests that the PPC may index lack-of-recollection of target details and increased attentional demands for further evaluation by APFC. This interpretation is consistent with a recent study by Hutchinson et al. (2014), which revealed enhanced activity associated with decision uncertainty in the superior parietal and supramarginal cortex, in close proximity to the PPC cluster reported in this study, along with superior frontal and anterior insular regions. These results were interpreted as reflecting the deployment of goal-directed attention in these parietal regions for responses associated with difficult recognition judgments and the lowest subjective confidence (Hutchinson et al. 2014). Combining these findings with our results, an intriguing possibility is that an increased demand on attention may be one of the cues used by the APFC, influencing the decision to report uncertainty. While this idea remains to be tested directly in future research, it is consistent with our developmental finding of age differences in the APFC but not PPC.

The ACC cluster identified in this study, which overlaps partially with areas previously associated with metacognitive monitoring (Chua et al. 2006), may be involved in the coordination of brain regions supporting performance monitoring and decision-making. Insular and cingulate areas are frequently coactivated when monitoring of performance is necessary (Medford and Critchley 2010). Thus, we expected that these areas would show similar involvement during source retrieval failure in our paradigm. At the same time, several studies have also demonstrated a dissociation in the activity profiles between these areas (Ullsperger et al. 2010), which would be consistent with theoretical frameworks suggesting that the insula is differentially involved in interoception (Craig 2009; Singer et al. 2009) and that the anterior cingulate is primarily involved in the evaluation of the expected value of cognitive control (cf. Shenhav et al. 2013). Our results are consistent with ACC engagement in expected value of control, such that it may integrate signals about the current state of retrieval success or failure from the AI and provide an estimation of expected value of control to prefrontal regions.

Our results speak to the role of recognizing retrieval failure as an important factor to understand memory development, which is a multifaceted process. Age differences in source accuracy have been related to hippocampal development during correct source retrieval (DeMaster and Ghetti 2013) along with differences in available attentional resources (Bunge and Crone 2009), and environmental influences (Luby et al. 2016). Future studies are needed to address the mutual influences among these mechanisms over longer periods extending into early childhood and adolescence (Ordaz et al. 2013). Furthermore, while our results provide initial evidence of the distinct developmental trajectories of AI and APFC activity, future research is necessary to examine the extent to which age-related changes in task-related activity are related to gray and white matter development in these regions. Finally, given the high expression levels of dopamine receptors across insular,

cingulate, and prefrontal areas, age-related changes in dopamine neurotransmission (Rothmond et al. 2012) may also contribute to developmental changes in memory regulation.

In a broader context, educational findings suggest that, compared with errorless learning approaches, the introduction of learning challenges leads to short-term increases in error rates and better long-term learning outcomes (Bjork et al. 2013). Our results suggest that, during development, experiences with retrieval failure may help enhance long-term memory outcomes by boosting decision-making processes and their neural mechanisms. Providing feedback to children might promote the ability to make a connection between experienced memory states and their actual accuracy, thereby potentially fostering the calibration of overt behavioral responses to retrieval failure and uncertainty signals. Future research should address whether interventions aimed at enhancing performance monitoring or decision-making result in improved metacognitive regulation, especially in populations with known deficits in memory (Ghetti et al. 2010b) or behavioral regulation (Castel et al. 2011).

Together, our findings underscore the necessity for children to fine-tune the ability to introspect on their internal memory states so that they can refrain from judgment when relevant memories do not come to mind. By combining longitudinal behavioral and neuroimaging assessments, our results bolster the claim that the ability to monitor and effectively control retrieval failure contributes to memory function, and should be integrated into theories of episodic memory development.

Supplementary Material

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/

Funding

Notes

Conflict of Interest: None declared.

References

Badre D, Wagner AD. 2004. Selection, integration, and conflict monitoring; assessing the nature and generality of prefrontal cognitive control mechanisms. Neuron. 41:473–487.

Badre D, Doll BB, Long NM, Frank MJ. 2012. Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. Neuron. 9:595–607.

Bastin J, Deman P, David O, Gueguen M, Benis D, Minotti L, et al. 2016. Direct recordings from human anterior insula reveal its leading role within the error-monitoring network. Cereb Cortex. 2016:1–13.

Bjork RA, Dunlosky J, Kornell N. 2013. Self-regulated learning: beliefs, techniques, and illusions. Annu Rev Psychol. 64: 417–444.

Braver TS, Barch DM, Gray JR, Molfese DL, Snyder A. 2001. Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. Cereb Cortex. 11 (9): 825–836.

Bunge SA, Crone EA. 2009. Neural correlates of the development of cognitive control. In: Rumsey J, Ernst M, editors. Neuroimaging in developmental clinical neuroscience. Cambridge (UK): University Press. p. 22–37.

Burgess PW, Wu H. 2012. Rostral prefrontal cortex (Brodmann area 10): metacognition in the brain. In: Stuss DT, Knight RT, editors. Principles of frontal lobe function. New York: Oxford University Press. p. 524–544.

Cabeza R, Ciaramelli E, Olson IR, Moscovich M. 2008. The parietal cortex and episodic memory: an attentional account. Nat Rev Neurosci. 9:613–625.

Castel AD, Lee SS, Humphreys KL, Moore AN. 2011. Memory capacity, selective control, and value-directed remembering in children with and without attention-deficit/hyperactivity disorder (ADHD). Neuropsychology. 25:15–24.

Chua EF, Schacter DL, Rand-Giovannetti E, Sperling RA. 2006. Understanding metamemory: neural correlates of the cognitive process and subjective level of confidence in recognition memory. Neuroimage. 29 (4):1150–1160.

Craig AD. 2009. How do you feel – now? The anterior insula and human awareness. Nat Rev Neurosci. 10:59–70.

Critchley HD, Mathias CJ, Dolan RJ. 2001. Neural activity in the human brain relating to uncertainty and arousal during anticipation. Neuron. 29:537–545.

D'Esposito M. 2007. From cognitive to neural models of working memory. Philos Trans R Soc Lond B Biol Sci. 362:761–772.

Daugherty AM, Ofen N. 2015. That's a good one! Belief in the efficacy of mnemonic strategies contributes to age-related increase in associative memory. J Exp Child Psych. 136: 17–29.

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

Diana RA, Yonelinas AP, Ranganath C. 2007. Imaging recollection and familiarity in the medial temporal lobe: a three-component model. Trends Cogn Sci. 11:379–386.

Dosenbach NUF, Fair DA, Cohen AL, Schlaggar BL, Petersen SE. 2008. A dual-networks architecture of top-down control. Trends Cogn Sci. 12:99–105.

Fandakova Y, Shing YL, Lindenberger U. 2013. Differences in binding and monitoring mechanisms contribute to lifespan age differences in false memory. Dev Psychol. 49:1822–1832.

Fleck MS, Daselaar SM, Dobbins IG, Cabeza R. 2006. Role of prefrontal and anterior cingulate regions in decision-making processes shared by memory and nonmemory tasks. Cereb Cortex. 16:1623–1630.

Fleming SM, Dolan RJ. 2012. The neural basis of metacognitive ability. Philos Trans R Soc Lond B Biol Sci. 367:1338–1349.

Fleming SM, Huijgen J, Dolan RJ. 2012. Prefrontal contributions to metacognition in perceptual decision making. J Neurosci. 32:6117–6125.

Fleming SM, Weil RS, Nagy Z, Dolan RJ, Rees G. 2010. Relating introspective accuracy to individual differences in brain structure. Science. 329:1541–1543.

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

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

Ghetti S, Castelli P, Lyons KE. 2010a. Knowing about not remembering: developmental dissociations in lack-of-memory monitoring. Dev Sci. 13:611–621.

Ghetti S, Lee JK, Sims CE, Demaster DM, Glaser NS. 2010b. Diabetic ketoacidosis and memory dysfunction in children with type 1 diabetes. J Pediatr. 156:109–114.

Gilboa A, Alain C, Stuss DT, Melo B, Miller S, Moscovitch M. 2006. Mechanisms of spontaneous confabulations: a strategic retrieval account. Brain. 129:1399–1414.

Grinband J, Hirsch J, Ferrera VP. 2006. A neural representation of categorization uncertainty in the human brain. Neuron. 49:757–763.

Goupil L, Romand-Monnier M, Kouider S. 2016. Infants ask for help when they know they don't know. Proc Natl Acad Sci. http://doi.org/10.1073/pnas.1515129113.

Ham T, Leff A, de Boissezon X, Joffe A, Sharp DJ. 2013. Cognitive control and the salience network: an investigation of error processing and effective connectivity. J Neurosci. 33: 7091–7098.

Hanczakowski M, Pasek T, Zawadzka K, Mazzoni G. 2013. Cue familiarity and "don't know" responding in episodic memory tasks. J Mem Lang. 69:368–383.

Hebscher M, Barkan-Abramski M, Goldsmith M., Aharon-Peretz J, Gilboa A. 2015. Memory, decision-making, and the ventromedial prefrontal cortex (vmPFC): the roles of subcallosal and posterior orbitofrontal cortices in monitoring and control processes. Cereb Cortex. http://doi.org/10.1093/cercor/bhv220.

Hembacher E, Ghetti S. 2014. Don't look at my answer: subjective uncertainty underlies preschoolers' exclusion of their least accurate memories. Psychol Sci. 25:1768–1776.

Hembacher E, Ghetti S. 2016. Subjective experience guides betting decisions beyond accuracy: evidence from a metamemory illusion. Memory. 1–11.

Henson R, Rugg MD, Shallice T, Dolan R. 2000. Confidence in recognition memory for words: dissociating right prefrontal roles in episodic retrieval. J Cogn Neurosci. 12:913–923.

Hutchinson JB, Uncapher MR, Weiner KS, Bressler DW, Silver MA, Preston AR, Wagner AD. 2014. Functional heterogeneity in posterior parietal cortex across attention and episodic memory retrieval. Cereb Cortex. 24:49–66.

Kantner J, Lindsay DS. 2012. Cross-situational consistency in recognition memory response bias. Psychon Bull Rev. 21: 1272–1280.

Kelley CM, Lindsay DS. 1993. Remembering mistaken for knowing: ease of retrieval as a basis for confidence in answers to general knowledge questions. J Mem Lang. 32:1–24.

Kim H. 2010. Dissociating the roles of the default-mode, dorsal, and ventral networks in episodic memory retrieval. Neuroimage. 50:1648–1657.

Koriat A. 2007. Metacognition and consciousness. In: Zelazo PD, Moscovitch M, Thompson E, editors. The Cambridge handbook of consciousness. New York: Cambridge University Press. p. 289–325.

Koechlin E, Hyafil A. 2007. Anterior prefrontal function and the limits of human decision-making. Science. 318:594–598.

Koriat A, Ackerman R, Adiv S, Lockl K, Schneider W. 2014. The effects of goal-driven and data-driven regulation on metacognitive monitoring during learning: a developmental perspective. J Exp Psychol Gen. 143:386–403.

Koriat A, Goldsmith M. 1994. Memory in naturalistic and laboratory contexts: distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. J Exp Psychol Gen. 123:297–315.

Koriat A, Goldsmith M. 1996. Monitoring and control processes in the strategic regulation of memory accuracy. Psychol Rev. 103:490–517.

Lyons K, Ghetti S. 2013. I don't want to pick: Introspection on uncertainty underlies early strategic behavior. Child Dev. 84:726–736.

Luby JL, Belden A, Harms MP, Tillman R, Barch DM. 2016. Preschool is a sensitive period for the influence of maternal support on the trajectory of hippocampal development. Proc Natl Acad Sci USA. 13 (20):5742–5747.

Luna B, Marek S, Larsen B, Tervo-Clemmens B, Chahal R. 2015. An integrative model of the maturation of cognitive control. Ann Rev Neurosci. 38:151–170.

McArdle JJ, Nesselroade JR. 1994. Using multivariate data to structure developmental change. In: Cohen SH, Reese HW, editors. Life-span developmental psychology: methodological contributions. Hillsdale (NJ): Lawrence Erlbaum Associates. p. 223–267.

Medford N, Critchley HD. 2010. Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. Brain Struct Funct. 214:535–549.

Menon V, Uddin LQ. 2010. Saliency, switching, attention and control: a network model of insula function. Brain Struct Funct. 214:655–667.

Muthén LK, Muthén BO. 1998–2007. Mplus user's guide. 5th ed. Los Angeles (CA): Muthén and Muthén.

Nelson TO, Narens L. 1990. Metamemory: a theoretical framework and some new findings. In: Bower GH, editor. The psychology of learning and motivation. New York: Academic Press. p. 125–173.

O' Connor AR, Han S, Dobbins IG. 2010. The inferior parietal lobule and recognition memory: expectancy violation or successful retrieval? J Neurosci. 30:2924–2934.

Ofen N, Chai XJ, Schuil KD, Whitfield-Gabrieli S, Gabrieli JD. 2012. The development of brain systems associated with successful memory retrieval of scenes. J Neurosci. 32: 10012–10020.

Ordaz SJ, Foran W, Velanova K, Luna B. 2013. Longitudinal growth curves of brain function underlying inhibitory control through adolescence. J Neurosci. 33:18109–24.

Ploran EJ, Nelson SM, Velanova K, Donaldson DI, Petersen SE, Wheeler ME. 2007. Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. J Neurosci. 27:11912–11924.

Preuschoff K, Quartz SR, Bossaerts P. 2008. Human insula activation reflects risk prediction errors as well as risk. J Neurosci. 28:2745–2752.

Raz N, Lindenberger U, Rodrigue KM, Kennedy KM, Head D, Williamson A, Dahle C, Gerstorf D, Acker JD. 2005. Regional brain changes in aging healthy adults: general trends, individual differences, and modifiers. Cereb Cortex. 15: 1676–1689.

Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S. 2004. The role of the medial frontal cortex in cognitive control. Science. 15:443–447.

Rissman J, Gazzaley A, D'Esposito M. 2004. Measuring functional connectivity during distinct stages of a cognitive task. NeuroImage. 23:752–763.

Rotello CM, Macmillan NA. 2007. Response bias in recognition memory. In: Benjamin AS, Ross BH, editors. The psychology of learning and motivation: skill and strategy in memory use. Vol. 48. London: Academic Press. p. 61–94.

Rothmond DA, Weickert CS, Webster MJ. 2012. Developmental changes in human dopamine neurotransmission: cortical receptors and terminators. BMC Neurosci. 13:18.

Shenhav A, Botvinick MM, Cohen JD. 2013. The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron. 79:217–240.

Sheslow D, Adams W. 2003. Wide range assessment of memory and learning. In: Administration and technical manual. 2nd ed. Lutz (FL): Psychological Assessment Resources.

Siegler RS. 1996. Emerging minds: the process of change in children's thinking. New York: Oxford University Press.

Singer T, Critchley HD, Preuschoff K. 2009. A common role of insula in feelings, empathy and uncertainty. Trends Cogn Sci. 13:334–340.

Son LK, Metcalfe J. 2000. Metacognitive and control strategies in study-time allocation. J Exp Psychol Learn Mem Cogn. 26: 204–221.

Stuss DT, Alexander MP. 2007. Is there a dysexecutive syndrome? Philos Trans R Soc Lond B Biol Sci. 362:901–915.

Stuss DT, Knight RT. 2012. Principles of frontal lobe function. New York: Oxford University Press.

Ullsperger M, Harsay HA, Wessel JR, Ridderinkhof KR. 2010. Conscious perception of errors and its relation to the anterior insula. Brain Struct Funct. 214:629–643.

Uncapher MR, Wagner AD. 2008. Posterior parietal cortex and episodic encoding: insights from fMRI subsequent memory effects and dual-attention theory. Neurobiol Learn Mem. 91: 139–154.

Vo VA, Li R, Kornell N, Pouget A, Cantlon JF. 2014. Young children bet on their numerical skills: metacognition in the numerical domain. Psychol Sci. 25:1712–1721.

Xia M, Wang J, He Y. 2013. BrainNet Viewer: a network visualization tool for human brain connectomics. PLoS ONE. 8:e68910.

Yonelinas AP. 2002. Components of episodic memory: the contribution of recollection and familiarity. Philos Trans R Soc Lond B Biol Sci. 356:1363–1374.