The effects of age on neural correlates of recognition memory: an fMRI study

Mingzhu Hou[a], Tracy H. Wang[a], and Michael D. Rugg[a,b]

[a]Center for Vital Longevity and School of Behavioral and Brain Sciences,

The University of Texas at Dallas, Dallas, TX 75235, USA

[b]School of Psychology,

University of East Anglia, Norwich NR4 7TJ, UK

Corresponding author: Mingzhu Hou, 1600 Viceroy Drive, Suite 800, Dallas, TX 75235.

Email: mingzhu.hou@utdallas.edu

Abstract

Studies examining the effects of age on the neural correlates of recognition memory have yielded mixed results. In the present study, we employed a modified remember-know paradigm to compare the fMRI correlates of recollection and familiarity in samples of healthy young and older adults. After studying a series of words, participants underwent fMRI scanning during a test phase in which they responded "remember" to a test word if any qualitative information could be recollected about the study event. When recollection failed, participants signaled how confident they were that the test item had been studied. Young and older adults demonstrated statistically equivalent estimates of recollection and familiarity strength, while recognition memory accuracy was significantly lower in the older adults. Robust, age-invariant fMRI effects were evident in two sets of *a priori* defined brain regions consistently reported in prior studies to be sensitive to recollection and familiarity respectively. In addition, the magnitudes of 'familiarity-attenuation effects' in perirhinal cortex demonstrated age-invariant correlations with estimates of familiarity strength and memory accuracy, replicating prior findings. Together, the present findings add to the evidence that the neural correlates of recognition memory are largely stable across much of the healthy human adult lifespan.

Keywords: recollection, familiarity, perirhinal cortex, core recollection network

1. Introduction

It is widely held that recognition memory is supported by two functionally distinct memory signals, typically known as recollection and familiarity (see, for example, Yonelinas, 2002; Wixted & Mickes, 2010). Recollection refers to the retrieval of qualitative (usually contextual) information about a prior event, whereas familiarity refers to an acontextual sense of past occurrence. Compared to young adults, healthy older adults typically demonstrate a more prominent impairment in recollection than familiarity (for a review, see Koen & Yonelinas, 2014). As we discuss below, recollection and familiarity have been associated with dissociable neural correlates. Of importance here, studies examining the effects of age on the neural correlates of recollection and familiarity have yielded mixed findings.

Using functional magnetic resonance imaging (fMRI), the neural correlates of recollection and familiarity have been examined in multiple paradigms. For example, in studies employing the 'remember/know' procedure (Gardiner & Richardson-Klavehn 2000; Tulving, 1985), neural correlates of recollection (hereafter, 'recollection effects') have been operationalized by the contrast between the activity elicited by correctly recognized test items according to whether the items were endorsed as Remembered (recollected) or Know (familiar only). Neural correlates of familiarity ('familiarity effects') have been operationalized either by a contrast between items endorsed Know and correctly rejected new items (e.g. Duarte et al., 2010; Wheeler & Buckner, 2004), or by a linear contrast among unrecollected items according to the confidence rating assigned to each item in respect of its study status (on the assumption that confidence is a proxy for familiarity strength, e.g. Johnson et al., 2013; Yonelinas et al., 2005). In studies employing tests of associative recognition, recollection effects have been identified by

contrasting the activity elicited by studied pairs according to whether the pairs were correctly endorsed as 'intact' or incorrectly endorsed as 'rearranged', with the assumption that the 'intact' judgments depend on recollection of the inter-item associations, while 'rearranged' judgments reflect a combination of failed recollection and above criterion item familiarity. Familiarity effects in associative recognition tests have been operationalized by the contrast between studied items incorrectly endorsed as 'rearranged' and correctly rejected unstudied pairs (e.g. de Chastelaine et al., 2016, 2017; Kirwan & Stark, 2004). Finally, in studies using tests of source memory, recollection effects were operationalized by the contrast between activity elicited by recognized test items according to whether the items were associated with an accurate or inaccurate source judgment (e.g. Dulas & Duarte, 2014; Elward & Rugg, 2015; Thakral et al., 2015). The source memory paradigm does not, however, easily lend itself to the identification of familiarity effects since inaccurate source judgments might reflect recollection of information that is not diagnostic of source ('non-diagnostic recollection'; Yonelinas & Jacoby, 1996).

Numerous fMRI studies have linked distinct brain regions to recollection and familiarity (for reviews, see Horn et al., 2016; Kim 2010, 2013). Notably, regardless of the paradigm employed to elicit recollection effects, successful recollection has consistently been associated with enhanced BOLD activity in a set of regions (the 'core recollection' network) that includes medial prefrontal cortex (mPFC), posterior cingulate (PCC)/retrosplenial cortex, hippocampus, parahippocampal gyrus (PHC), left middle temporal gyrus (MTG) and left angular gyrus (AG) (Rugg & Vilberg, 2013). By contrast, familiarity has been consistently associated with enhanced activity in a different set of regions, including the dorsomedial and antero-lateral prefrontal cortex (dmPFC, alPFC), intra-parietal sulcus (IPS), precuneus and caudate nucleus (Kim, 2013; see also de Chastelaine et al., 2017). Familiarity has also been associated with *reduced* activity in

perirhinal cortex (PRC; e.g. de Chastelaine et al., 2017; Henson et al., 2003; Montaldi et al., 2006; Wang et al., 2014; for an early review, see Diana et al., 2007). Here, we examined the effects of age on the neural correlates of recognition memory by focusing on these recollection- and familiarity-sensitive regions.

The effects of age on fMRI correlates of recognition memory have been studied fairly extensively (for reviews, see Giovanello & Dew, 2015; Wang & Cabeza, 2016), but the findings with respect to both recollection and familiarity are inconsistent. In the case of recollection, while some studies reported null effects of age in regions belonging to the core recollection network (Folville et al., 2020; Dulas & Duarte, 2012; Wang et al., 2016), others reported either attenuated (Angel et al., 2013, 2016; Daselaar et al., 2006; de Chastelaine et al., 2016; Duarte et al., 2010; Kukolja et al., 2009) or enhanced (Duarte et al., 2008; Wang & Giovanello, 2016) effects in older adults in at least one region. Among the studies adopting the remember/know procedure, Wang et al. (2016) reported statistically equivalent recollection effects between young and older adults. Other studies however reported age-related attenuation of recollection effects in left AG (Angel et al., 2013, 2016; Duarte et al., 2010) and PHC (Angel et al., 2013). In the study of Daselaar et al. (2006), where recollection and familiarity were estimated as parametric variables derived from recognition confidence ratings, older adults exhibited smaller recollection effects in the hippocampus, AG and retrosplenial cortex. Findings of age effects in studies employing the associative recognition procedure are also mixed: in de Chastelaine et al. (2016), smaller recollection effects in older adults were evident in left mPFC, left and right hippocampus and left PCC (but see below). By contrast, Wang and Giovanello (2016) reported that recollection effects in left hippocampus were enhanced with increasing age. Similarly, in studies adopting a source memory procedure, Dulas and Duarte (2012) reported comparable

recollection effects across age groups, whereas attenuated effects in the left hippocampus of older adults were described by Kukolja et al. (2009). By contrast, Duarte et al. (2008) reported enhanced effects in the left AG in older adults.

In the case of familiarity, age effects have again included a mix of null findings (Wang & Giovanello, 2016), and both attenuated (Angel et al., 2013; Duarte et al., 2010) and enhanced effects (Daselaar et al., 2006) in one or more familiarity-sensitive regions in older adults. In two studies employing the remember-know procedure (Angel et al. 2013; Duarte et al. 2010), age-related attenuation of familiarity effects in dorsomedial and inferior lateral PFC were reported. In a study employing associative recognition (de Chastelaine et al., 2017), reduced familiarity effects in older adults were restricted to a small sub-region of familiarity-sensitive dmPFC. By contrast, Daselaar et al. (2006) reported age-invariant familiarity effects in precuneus/PCC and parieto-occipital cortex but larger effects in PRC (i.e. greater familiarity-related attenuation of activity) in older adults. The enhanced familiarity effects in PRC, which were accompanied by attenuated recollection effects in core recollection regions (see above), were accounted for by the proposal that older adults depend on familiarity more heavily than young adults to compensate for their impaired recollection.

Of importance, in the majority of the above-mentioned studies, older adults demonstrated lower memory performance than young adults, especially in the case of recollection. The presence of age differences in memory performance makes it difficult to determine whether corresponding age differences in task-related functional activity should be attributed to differences in performance or age (for discussion of this issue, see de Chastelaine et al., 2016; Rugg, 2017; Rugg & Morcom, 2005). With the exception of Angel et al. (2013), Daselaar et al. (2006) and Duarte et al. (2008), studies in which memory performance was matched or

statistically controlled for across age groups have tended to report limited or null effects of age (for recollection, see for example de Chastelaine et al., 2016; Dulas & Duarte, 2012; for familiarity, see for example de Chastelaine et al., 2017; Wang & Giovanello, 2016).

The present study sought to further examine the effects of age on the neural correlates of recollection and familiarity. Combining the approaches of Wang et al. (2016), Daselaar et al. (2006) and Angel et al. (2013), we adopted a modified remember-know procedure (Johnson et al., 2013; Yonelinas et al., 2005) that operationalized recollection in terms of R judgments, and familiarity in terms of recognition confidence. Participants responded 'R' to a test item if they could recollect any contextual information about its study event. In the absence of recollection, they rated their confidence that the item had been presented at study. fMRI effects were examined in two sets of regions consistently reported to be sensitive to recollection and familiarity respectively (Kim, 2013; Rugg & Vilberg, 2013). These regions of interest (ROIs) were defined by reference to an independent dataset described in two prior reports that examined the effects of age on the neural correlates of recollection (de Chastelaine et al., 2016) and familiarity (de Chastelaine et al., 2017). If the findings of Wang et al. (2016) generalize to the present study, we anticipated identifying equivalent recollection effects in the two age groups. According to the findings of Angel et al. (2013) and Daselaar et al. (2006), however, an age-related decline in recollection effects is expected. Similarly, we anticipated age-invariant or enhanced familiarity effects if the findings of Daselaar et al. (2006) generalize to the present study, but attenuated effects in older adults based on the findings of Angel et al. (2013).

2. Materials and methods

2.1 Participants

Twenty-four older adults aged between 63 and 77 years (mean = 69 years, 13 female) and nineteen young adults aged between 18 and 25 years (mean = 20 years, 8 female) participated in the experiment. Young participants were recruited from the University of California Irvine (UCI) undergraduate community. Older participants were recruited from the surrounding community through newspaper advertisements and flyers. All participants were right-handed fluent English speakers with negative neurological and psychiatric histories and normal or corrected-to-normal vision. Exclusion criteria included a history of cardiovascular disease other than treated hypertension, diabetes, psychiatric disorder, illness or trauma affecting the CNS, substance/alcohol abuse, or current or recent use of psychotropic medication. Additional exclusion criteria include a score on the Mini Mental State Exam (MMSE) < 26, a score on a standardized memory test >1.5 SD below the age-appropriate norm, or low performance (>1.5 SD below norm) on two or more of the non-memory tests on the neuropsychological test battery described in the following section. Informed consent was obtained in accordance with UC Irvine Institutional Review Board guidelines. Participants were compensated $25 for the first hour and $15 for each hour thereafter. Three participants (2 young and 1 older) were excluded due to poor memory performance (pR < .10). Two additional participants (1 young and 1 older) were excluded because they failed to spread their responses across all response options: specifically, old items were disproportionately endorsed as 'R' (> .57) at the expense of the 'confident old' option (< .07). Thus, a total of 22 older participants and 16 young participants are included in the analyses reported below.

2.2 Neuropsychological testing

A standardized neuropsychological battery was administered to participants on a separate day from the MRI session. The battery comprised the California Verbal learning Test – II

(CVLT; Norman et al., 2000), the Immediate and Delayed NYU paragraph recall test (Kluger et al., 1999), Digit Span Forward and Backward test of the WAIS-R (Wechsler, 2001), Digit/Symbol Coding test of the WAIS -R, Trail Making Tests A and B, letter and category fluency tests and the Wechsler Test of Adult Reading (WTAR) (Wechsler, 2001) (see Table 2 for a complete list). The Beck Depression Inventory was also administered (Beck et al., 1961).

2.3 Experimental items

The 240 critical words employed in the study and test lists were selected from the Medical Research Council Psycholinguistics Database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). All were concrete and ranged from three to nine letters in length (mean = 5.6, SD = 1.5) and from 0 -50 (mean = 14.1, SD = 12.4) occurrences per million (Kučera & Francis, 1967). The critical words were pseudo-randomly assigned into eighteen study-test lists, following the criterion that equal numbers of words denoting living and nonliving entities should be included in both study and test lists (see below). Study-test list pairs were yoked across young and older adults. Each study list comprised 120 words, which were pseudo-randomly ordered so that there were no more than three consecutive presentations of words denoting living or nonliving entities. A corresponding test list comprised the120 critical words from the study list and 120 unstudied words. The test list was subdivided into two sub-lists, each comprising 60 studied and 60 new words. The test words were pseudo-randomly ordered so that no more than three studied or unstudied words occurred in succession. Eighty null trials (40 for each sub-list) were intermixed with the critical test words. Two filler words presented at the beginning and at the end of the study list. Two fillers were also presented at the beginning of each test list, and another two fillers immediately followed each rest break during the test phase (see below). The stimuli were viewed via a mirror mounted on the scanner

head coil that reflected a back projected image displayed on a screen at the head of the scanner bore. Words were presented in central vision superimposed on a grey background square (subtending a visual angle of 5.72° x 5.72°) that was continuously displayed in the center of an otherwise dark screen.

2.4 Experimental Procedure

The experiment comprised the presentation of a study list followed by two consecutive test lists. The study and test phases were both administered while participants were in the MRI scanner, but fMRI data was only acquired during the test phase. A 30s rest break was inserted halfway through each test block. Prior to the study and test phases, participants were briefly reminded of the instructions and briefly practiced the upcoming task. Each study trial began with the presentation of a red fixation cross over a gray background for 500 ms. A word replaced the fixation cross and remained on the screen for 1000 ms. A black fixation cross then replaced the word for 1650 ms, and the next trial then followed. Participants were instructed to judge whether or not the object denoted by each word was animate (i.e. living or part of a living thing). Participants indicated their judgments with their right and left index fingers, with response assignment counterbalanced across participants.

The test phase - which began around five minutes after the completion of the study phase - was administered across two separate scanning sessions. The sequence of events for each test trial was the same as in the study phase, except that the black fixation cross that followed the offset of each word was presented for 2000 ms. Null trials consisted of the presentation of a black fixation cross for 3500 ms. Participants were instructed to make one of five different responses to each test word. If they judged the word as having been studied and could recollect specific contextual details about the study event, they were to press ‘R’ with the index finger of

one hand. When recollection failed, they were to rate their confidence about the study status of the test item using the hand opposite to the one assigned for the R judgment [‘confident old’ (index finger), ‘unconfident old’ (middle finger), ‘unconfident new’(ring finger) and ‘confident new’ (little finger)]. The response hand assignments were counterbalanced across participants. Participants were instructed to respond in a timely manner but with an emphasis on accuracy. After the completion of the test phase, a T1-weighted anatomical image (see below) was acquired.

Practice on the experimental tasks was administered prior to the experiment. The practice list for the study phase comprised 75 words, which were presented in the same manner as in the formal experiment. The practice list for the test phase comprised 75 studied words intermixed with 75 unstudied words. To ensure that participants adequately understood and utilized each response type appropriately at test, they were required to justify each of their responses in a self-paced version of the retrieval task during the first half of the test practice. In particular, if participants made a ‘R’ response, they were requested to describe the recollected contextual detail(s). The second half of the test practice was administered as in the experiment proper, so that participants could practice responding within the allotted time.

2.5 MRI acquisition and preprocessing

Functional and anatomical images were acquired with a Philips Achieva 3T MR scanner equipped with a transmit/receive radio frequency head coil. Functional scans were acquired with a T_2^*–weighted echo-planar image (EPI) sequence with the following parameters: TR 2 s, TE 30 ms, flip angle 70°, FOV 240x240, matrix size 80x79, in plane resolution 3mm x 3mm. Each EPI volume consisted of thirty 3 mm thick slices separated by a 1 mm interslice gap, acquired in ascending order, oriented parallel to the AC–PC line and positioned for full coverage of the

cerebrum and most of the cerebellum. Functional data were acquired during each of the two retrieval blocks (311 volumes each). The first five volumes of each block were discarded to allow tissue magnetization to achieve a steady state. A T1-weighted anatomical image was acquired using a 3D magnetization-prepared rapid gradient echo (MP-RAGE) pulse sequence (FOV= 240×240, matrix size 220x193, voxel size $1mm^3$, 150 slices, sagittal acquisition).

Data preprocessing was performed with Statistical Parametric Mapping software (SPM12; Wellcome Department of Imaging Neuroscience, London, UK: www.fil.ion.ucl.ac.uk/spm) in MATLAB R2018b (The MathWorks, Natick, MA). The functional volumes were realigned to the across-run mean image, slice-time corrected and spatially normalized to a study-specific template following the same procedure implemented by de Chastelaine et al. (2015). The normalized images were resampled to 3mm isotropic voxels and smoothed with an 8mm Gaussian kernel. Before being entered into subject-wise General Linear Models (GLMs), the functional data from the two test blocks were concatenated using the spm_concatenate.m function. Anatomical images were normalized to a study-specific T1 template following procedures analogous to those applied to the functional images.

2.6 Statistical analyses of MRI data

Analyses of the fMRI data were conducted in two stages. In the first stage, separate GLMs were employed to analyze the data from each participant. The design matrix consisted of five events of interest: (i) R responses to old test items, and (ii) confident old, (iii) unconfident old, (iv) unconfident new, and (v) confident new responses regardless of the study status of the items attracting the response (cf. Yonelinas et al., 2005). The average number of trials (and range) for each judgment were 35 (14-64), 44 (15-64), 47 (31-68), 67 (42-90), 44 (23-85) for young adults, and 32 (14-60), 57 (25-94), 45 (13-75), 53 (13-79), 49 (21-115) for the older

adults. Filler trials, trials with multiple or missed responses, and trials attracting 'R' false alarms were modeled as events of no interest, along with the 30-second rest breaks located in the middle of each test block. Stimulus-elicited neural activity was modeled with two delta functions convolved respectively with a canonical hemodynamic response function (HRF) (Friston et al., 1998) and a delayed HRF, which was generated by shifting the canonical HRF one TR (2s) later in time. The delayed HRF was orthogonalized with respect to the canonical function so that variance common to both functions was allocated to the canonical HRF (Andrade et al., 1999). Thus, loadings on the orthogonalized delayed function accounted only for variance unexplained by the canonical function. The results obtained with the delayed HRF added little of theoretical significance to the findings obtained with the canonical HRF and are not reported here (the results are available from the first author on request). The model also included as covariates six regressors modeling motion-related variance (three for rigid-body translation and three for rotation) and two constants for means across test blocks. Data from volumes showing a transient displacement of > 1mm or > 1 degree in any direction were treated as covariates of no interest.

Both whole brain and ROI analyses were conducted on the parameter estimates derived from the first stage GLMs. For the whole brain analyses, voxel-wise participant-specific parameter estimates corresponding to each event of interest were entered into a 2 (age group: young, older) x 5 (response type: R, confident old, unconfident old, unconfident new, confident new) mixed-design ANOVA model implemented in SPM12. The outcomes of these ANOVAs were evaluated at a height threshold of $p < .001$ and a cluster extent threshold FWE corrected to $p < .05$ ($k > 77$). Recollection effects were operationalized as greater BOLD activity for R responses than for confident old responses. Familiarity effects were operationalized as activity identified by a linear contrast across the four confidence judgments (weights of 3 1 -1 -3; cf.

Johnson et al. 2013; Yonelinas et al., 2005). To identify common effects between young and older adults, the across-group effects of recollection and familiarity were exclusively masked by the corresponding two-sided age group x fMRI effect interaction contrasts, each thresholded at $p < .05$. Note that the findings from these whole brain analyses are presented in the Results section largely for descriptive purposes.

For the primary analyses, we employed participant-specific parameter estimates for each event of interest derived from *a priori* defined regions of interest (ROI). The ROIs were defined on the basis of the analyses of an independent dataset described in two prior reports of the effects of age on the neural correlates of recollection (de Chastelaine et al., 2016) and familiarity (de Chastelaine et al., 2017), respectively.

In addition to the aforementioned ROIs previously reported to demonstrate familiarity-related enhancement of activity, we also included as additional ROIs left and right perirhinal cortex (PRC), employing the anatomical masks of these regions originally created by de Chastelaine et al. (2017; see Fig 1). The anatomically defined left and right PRC ROIs comprised 14 and 19 voxels respectively. The rationale for including the PRC derives from prior reports of the sensitivity of the region to familiarity and the reported role of age as a moderator of these familiarity effects (see Introduction).

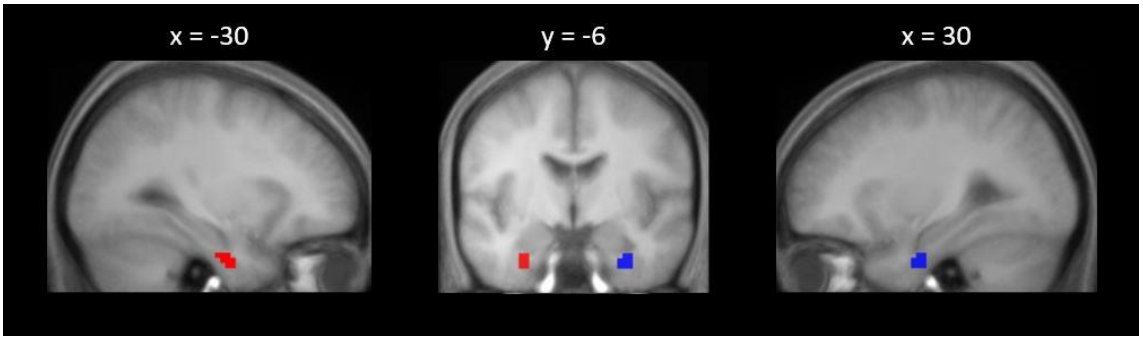

Fig 1. Anatomical masks of left and right PRC depicted on representative sagittal and coronal sections of the study-specific T1-weighted structure image.

Recollection- and familiarity-sensitive ROIs, and their corresponding peak MNI coordinates, are given in Table 1. For each ROI other than the PRC, parameter estimates were averaged within either a 3 mm (hippocampus, PHC, caudate) or 5 mm (all other ROIs) radius of the peak co-ordinates reported in the table. For left and right PRC, parameter estimates were averaged across all voxels falling within the anatomical masks.

We employed separate sets of ANOVAs to examine the effects of recollection and familiarity. Parameter estimates from the recollection ROIs were subjected to a 2 (age group: young, older) x 2 (response type: R, confident old) x 7 (region) mixed-design ANOVA and a 2 (age group: young, older) x 4 (response type: confident old, unconfident old, unconfident new, confident new) x 7 (region) mixed-design ANOVA, respectively. For the non-PRC familiarity ROIs and the PRC, analogous mixed effects ANOVAs (factors of age group, response type and region) were implemented.

Significant effects involving interactions between response type and age group or region were followed up with subsidiary pairwise comparisons between R and confident old judgments to examine the effects of recollection. Familiarity effects were examined using linear and quadratic trend analyses of the parameter estimates corresponding to the four confidence levels.

Table 1. Regions of interest for analyses of recollection and familiarity.

ROI	Peak_MNI (x, y, z)
Recollection	
Left_PHC	-21, -37, -17
Left_AG	-51, -70, 37
Left_MTG	-57, -55, 16
Left_mPFC	-3, 56, 13
Left_PCC	-6, -46, 37
Left_Hippocampus	-24, -13, -20
Right_Hippocampus	27, -16, -20
Familiarity	
Left_dmPFC	-6, 29, 43
Left_alPFC	-48, 29 22
Left_Caudate	-12, 8, 4
Left_IPS	-36, -58, 40
Left_Precuneus	-6, -76, 43
Right_Caudate	12, 11, 1
Left_PRC	–
Right_PRC	–

2.7 Behavioral measures

Recollection performance (pR) was indexed as the difference between the proportions of old words attracting an R judgment and new words incorrectly endorsed as R. The behavioral index of familiarity (pF) was estimated according to the independent Remember/Know procedure (Yonelinas & Jacoby, 1995). Familiarity was estimated after collapsing across confident old (CO) and unconfident old (UCO) judgments. Formulae for calculating pR and pF were as follows:

pR=(p(R|old)-p(R|new)

pF=[(p(CO+UCO|old)/(1-p(R|old)]-[(pCO+UCO|new)/(1-p(R|new)]

To supplement the metrics of pR and pF, we also computed the accuracy of R, confident old and unconfident old judgments respectively (cf. Wang et al., 2012; Wixted et al., 2010). For each judgment, accuracy was calculated as pHit/(pHit+pFalse alarm).

To test for relationships between fMRI effects and memory performance, and whether any such relationship differed with age, we conducted separate linear regression analyses predicting pR, pF and mean accuracy. In these analyses, the recollection fMRI effect was estimated as the effect (R - confident old) obtained in each of the recollection-sensitive ROIs. Because the confident old judgments seemingly reflected a mixture of familiarity and sub-criterial recollection signals (see results), the *familiarity-enhancement* effect was calculated as the difference between parameter estimates for unconfident old and confident new judgments in the non-PRC familiarity-sensitive ROIs (on the assumption that unconfident old judgments would provide a more 'process pure' assessment of familiarity strength). Following the same assumption, PRC *familiarity-attenuation* effects were calculated as the difference between parameter estimates associated with confident new and unconfident old judgments. In the models predicting pR, predictor variables included the functional effect, age group (young coded as 0, older coded as 1), and the functional effect x age group interaction term. The models predicting pF and memory accuracy were constructed in an analogous fashion.

2.8 Statistical Analyses

Statistical analyses were conducted with SPSS 27.0 and JASP 0.14.0.0. Nonsphericity between levels of repeated measures factors in the ANOVAs was corrected with the

Greenhouse–Geisser procedure (Greenhouse & Geisser, 1959). Significance levels for all tests were set at $p < .05$. For comparisons conducted at each ROI to examine fMRI effects, the significance levels were corrected with the Holm-Bonferroni procedure to give a family-wise error rate of $p < .05$ across the ROIs. Unless noted otherwise, all results reported as significant remained so after correction.

In a complementary set of analyses, Bayes factors were estimated to examine the evidence supporting null findings arising from the null hypothesis significance testing approach described above. We report BF_{10} values for t-tests and BF inclusion (BF_{incl}) values for ANOVAs and linear regressions for the relevant analyses. BF_{10} values estimate how many times more likely the data are to favor the alternative hypothesis over the null hypothesis, with $BF_{10} > 1$ indicating support for the alternative hypothesis and $BF_{10} < 1$ indicating support of the null hypothesis. BF_{incl} values are computed by comparing models containing the effects of interest with models in which the effect has been omitted (Mathôt, 2017). As a result, they provide an estimation of how many times more likely a particular effect is to account for variance in the data than not to do so. BF_{incl} values > 1 favor the existence of the effect, while BF_{incl} values < 1 indicate support of the absence of the effect. For both BF_{10} and BF_{incl}, values between .33 and 1 are conventionally considered as anecdotal evidence, between .1 and .33 as substantial evidence, and between 0 and .10 as strong evidence in favor of the null hypothesis (Jeffreys, 1961).

3. Results

3.1 Neuropsychological data

Demographic information and neuropsychological test scores are shown in Table 2. As is evident from the table, and as is typical for studies such as this (cf. de Chastelaine et al., 2015), older adults had completed more years of education than the younger group of college-age adults. Compared with the young adults, older adults scored significantly lower on CVLT long-delay cued recall, delayed recognition and false positives, and the SDMT and Trails tests.

Table 2. Demographic and neuropsychological data for young and older adults.

	Young		Old	
	Mean (SD)	Range	Mean (SD)	Range
Age	20.3 (1.8)	18-25	69.1(4.1)	63-77
Years of education[a]	14.8 (2.2)	12-19	16.7 (1.8)	14-20
Mini Mental State Examination	29.3 (.7)	28-30	29.2 (1.0)	27-30
CVLT short-delayed free recall	12.6 (2.0)	8-16	11.6 (3.2)	7-16
CVLT short-delayed cued recall	13.2 (2.1)	9-16	12.5 (2.3)	9-16
CVLT long-delayed free recall	13.1 (2.2)	8-16	11.6 (3.1)	6-16
CVLT long-delayed cued recall[a]	13.8 (1.9)	10-16	12.0 (2.9)	7-16
CVLT recall composite[b]	13.1 (2.0)	9.3-16	11.9 (2.8)	7.5-16
CVLT delayed recognition[a]	15.7 (.7)	14-16	15.0 (1.1)	12-16
CVLT recognition false positives[a]	.44 (.73)	0-2	1.9 (2.3)	0-7
NYU paragraph immediate recall	7.5 (2.1)	3-10.5	7.8 (2.6)	4.5-13.5
NYU paragraph delayed recall	10.7 (2.6)	5-16	10.4 (3.1)	5-16
Forward/backward Digit Span	18.9 (4.3)	13-25	18.7 (2.8)	16-26
SDMT[a]	62.3 (6.7)	53-79	51.8 (7.6)	39-70
Trail Making test A (seconds)[a]	23.4 (9.3)	14-48	29.0 (6.1)	21-42
Trail Making test B (seconds)[a]	42.0 (11.8)	28-72	62.9 (18.2)	31-94
Letter Fluency	46.1 (12.0)	29-75	45.4 (10.1)	24-61
Category fluency	24.1 (5.5)	17-33	23.5 (4.8)	16-33
WTAR	40.9 (4.0)	34-48	43.1 (5.5)	33-50

Note. a: difference between young and older adults, $p < .05$. b: average number of words recalled on the short- and long-delayed free- and cued-recall tests.

3.2 Behavioral results

Table 3 shows the recognition memory performance for young and older adults. We conducted a 5 (response type: R, confident old, unconfident old, unconfident new, confident new) x 2 (age group: young, older) x 2 (study status: old, new) mixed effects ANOVA on the proportions of trials endorsed with each class of response. Besides a significant main effect of response type [$F(2.78, 100.05) = 8.03$, $p < .001$, partial $\eta^2 = .182$], the ANOVA revealed significant response type x study status [$F(4, 144) = 192.43$, $p < .001$, partial $\eta^2 = .842$] and response type x age group x study status interactions [$F(4, 144) = 3.22$, $p = .014$, partial $\eta^2 = .082$]. Follow-up simple effects analyses revealed that for old items, young and older adults demonstrated statistically equivalent response rates for each of the recognition judgments [$ts < 1.30$, $ps > .133$, $BF_{10}s < .80$). For the new items, however, young adults were more likely to use the unconfident new option than were older adults [$t(36) = 3.39$, $p = .002$]. By contrast, older participants exhibited higher false alarm rates for R [$t(36) = 2.10$, $p = .043$] and confident old [$t(33.29) = 3.17$, $p = .001$] judgments.

We went on to contrast recollection (pR) and familiarity (pF) estimates between age groups. Young and older adults did not significantly differ on either pR [for young, $M = .29$, $SD = .10$; for older, $M = .25$, $SD = .11$, $t(36) = 1.19$, $p = .243$, $BF_{10} = .55$] or pF [for young, $M = .60$, $SD = .12$; for older, $M = .53$, $SD = .14$, $t(36) = 1.50$, $p = .142$, $BF_{10} = .76$].

We also examined potential age group differences in the accuracy of R and old judgments. For R judgments, mean accuracy was .99 for young adults (SD = .03) and .93 for

older adults (SD = .08). Numerically lower accuracy in the older adults was also observed for the confident old and unconfident old judgments [respectively: M_{young_co} = .92, SD = .09, M_{older_co} = .83, SD = .11; M_{young_uco} = .59, SD = .14, M_{older_uco} = .53, SD = .15]. A 3 (response type: R, confident old, unconfident old) x 2 (age group: young, older) mixed effects ANOVA revealed a significant main effect of response type [$F(1.54, 55.38) = 225.34$, $p < .001$, partial $\eta^2 = .862$]. Follow-up analyses identified a significant decrease of accuracy going from R, to confident old, to the unconfident old judgments (for all three pairwise comparisons, $ts > 5.87$, $ps < .001$). Of importance, the main effect of age group was also significant, $F(1, 36) = 6.10$, $p = .018$, partial $\eta^2 = .145$, indicative of lower overall accuracy in older adults than young adults.

Turning to the RT data (see Table 3), a 5 (response type: R, confident old, unconfident old, unconfident new, confident new) x 2 (age group: young, older) mixed effects ANOVA revealed a significant main effect of response type [$F(3.14, 113.18) = 50.55$, $p < .001$, partial $\eta^2 = .584$], reflecting the inverted U-shape of the relationship between RTs and confidence that is typically observed for this paradigm (Johnson et al., 2013; Wang et al., 2012; Yonelinas et al., 2005). The response type x group interaction was also significant [$F(3.14, 113.18) = 4.83$, $p = .003$, partial $\eta^2 = .118$]. However, follow-up analyses did not identify a group difference for any of the response types ($ts < 1.60$, $ps > .089$, $BF_{10}s < 1.03$). Rather, the interaction was driven by shorter RTs for R judgments than for confident new judgments in young but not in older adults [respectively: $t(15) = 5.28$, $p < .001$; $t(21) = 1.76$, $p = .093$, $BF_{10} = .83$].

Table 3. Recognition memory performance for young and older adults (standard deviations are in parenthesis).

	R	Confident old	Unconfident old	Unconfident new	Confident new
Young					
Proportion_Old	.30 (.11)	.33 (.11)	.23 (.07)	.10 (.04)	.04 (.03)
Proportion_New	.01 (.01)	.03 (.03)	.16 (.07)	.46 (.12)	.33 (.11)
RT	1968 (280)	2160 (296)	2363 (343)	2324 (336)	2231 (346)
Older					
Proportion_Old	.27 (.11)	.39 (.12)	.19 (.09)	.10 (.05)	.03 (.02)
Proportion_New	.02 (.02)	.08 (.06)	.18 (.09)	.33 (.11)	.38 (.19)
RT	2003 (246)	2155 (256)	2482 (280)	2440 (277)	2079 (192)

Note. RTs for R responses are presented for old items only due to the very few new items endorsed as R. RTs for confidence judgments are presented collapsed across old/new status.

3.3 Exploratory whole brain fMRI analyses

At our pre-experimentally determined thresholds, no clusters demonstrated an age group x response type interaction indicative of group differences in recollection or familiarity effects. As detailed in the Methods, the main effects of recollection and familiarity were exclusively masked by the corresponding (and liberally thresholded) age group x response type interactions to identify effects common to the two age groups. As is evident from Fig 2A, significant age-invariant recollection effects overlapped each of the *a priori* determined ROIs with the exception of the right hippocampus. Similarly, age-invariant familiarity effects overlapped each of the non-PRC familiarity-sensitive ROIs (Fig 2B).

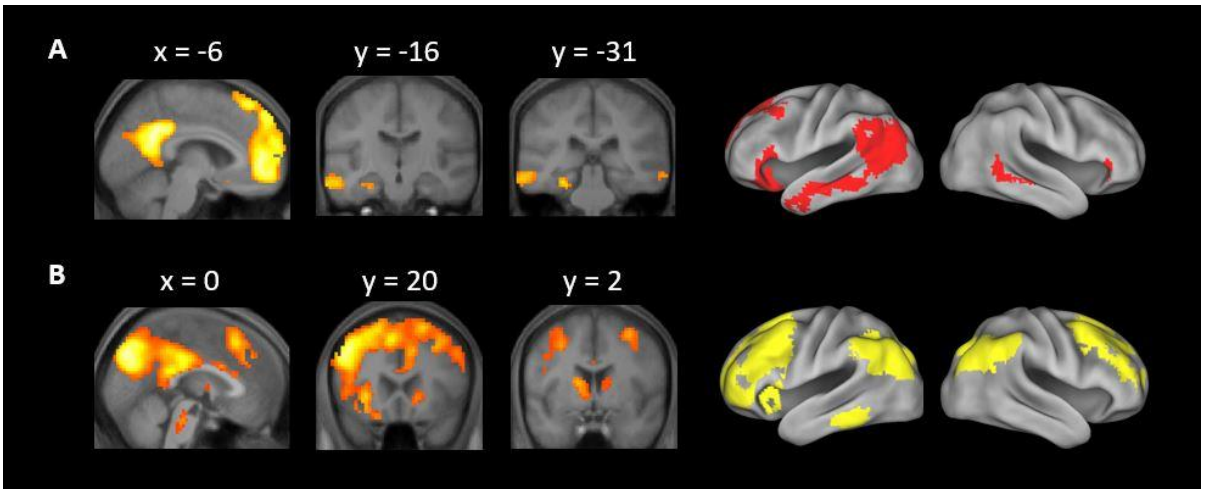

Fig 2. Whole brain analyses. A: Clusters demonstrating age-invariant recollection effects. B: Clusters demonstrating age-invariant effects of familiarity. All effects are height thresholded at $p < .001$ and combined with a $p < .05$, FWE-corrected cluster extent threshold.

3.4 ROI analyses

3.4.1 Recollection ROIs

As already detailed (see Methods), recollection effects were evaluated with a mixed effects ANOVA on parameter estimates extracted from *a priori* determined recollection-sensitive ROIs. As is evident from the first panel of Table 4 (see also Fig 3), all three main effects were significant. These effects reflected differential levels of BOLD activity according to region, greater mean BOLD activity for R than for confident old judgments, and greater activity overall for older than for young adults. Of importance, while the response type x region interaction attained significance, no significant interactions involving the factors of both age group and response type were identified.

To follow up the response type x region interaction, planned comparisons were conducted on data from each ROI after collapsing across age groups. Pairwise comparisons between R and confident old judgments revealed significant recollection effects in all regions

except for the right hippocampus [for right hippocampus, $t(37) = 1.98$, $p = .055$, $BF_{10} = 1.01$, for the other regions, $ts > 3.45$, $ps < .002$].

Table 4. Results of mixed-design ANOVAs across the recollection-sensitive ROIs (bold values denote significance at the $p < .05$ level).

2 (age group) x 2 (response type: R, confident old) x 7 (region) ANOVA	
Response type	**$F(1, 36) = 48.27$, $p < .001$, partial $\eta^2 = .573$**
Group	**$F(1, 36) = 5.49$, $p = .025$, partial $\eta^2 = .132$**
Region	**$F(4.25, 153.18) = 13.16$, $p < .001$, partial $\eta^2 = .268$**
Response type x group	$F(1, 36) = .13$, $p = .719$, partial $\eta^2 = .004$, $BF_{incl} = .16$
Region x group	$F(4.25, 153.18) = 1.05$, $p = .387$, partial $\eta^2 = .028$, $BF_{incl} = .09$
Response type x region	**$F(2.37, 85.21) = 7.56$, $p < .001$, partial $\eta^2 = .174$**
Response type x region x group	$F(2.37, 85.21) = 1.55$, $p = .215$, partial $\eta^2 = .041$, $BF_{incl} = .05$
2 (age group) x 4 (response type: four confidence ratings) x 7 (region) ANOVA	
Response type	**$F(3, 108) = 25.56$, $p < .001$, partial $\eta^2 = .415$**
Group	**$F(1, 36) = 14.85$, $p < .001$, partial $\eta^2 = .292$**
Region	**$F(4.08, 146.98) = 22.27$, $p < .001$, partial $\eta^2 = .382$**
Response type x group	$F(3, 108) = .55$, $p = .649$, partial $\eta^2 = .015$, $BF_{incl} = .02$
Region x group	**$F(4.08, 146.98) = 3.33$, $p = .012$, partial $\eta^2 = .085$**
Response type x region	**$F(7.81, 281.15) = 5.13$, $p < .001$, partial $\eta^2 = .125$**
Response type x region x group	$F(7.81, 281.15) = 1.48$, $p = .165$, partial $\eta^2 = .040$, $BF_{incl} = .002$

An additional 2 (age group: young, older) x 4 (response type: confident old, unconfident old, unconfident new, confident new) x 7 (ROI) mixed-design ANOVA was conducted to examine familiarity effects in the core recollection ROIs. As is evident from the second panel of Table 4, each of the main effects achieved significance, reflecting differential mean BOLD

activity according to response type, group and region. While the interactions including the factors of both response type and age group failed to attain significance, the region x group and response type x region interactions were significant. In the analyses following up the region x group interaction, greater BOLD activity in older adults relative to young adults was evident in left PHC, mPFC, MTG and PCC (ts > 2.69, ps < .012) but not in other regions (ts < 1.06, ps > .297, BF_{10}s < .50).

To unpack the response type x region interaction, a one-way ANOVA was conducted for each ROI, with the BOLD activity of the four confidence judgments as the dependent variable. Linear and quadratic polynomial trends were then examined across the confidence judgments. As is evident from Table 5, the main effect of response type was significant in all ROIs. A significant quadratic trend was consistently evident across ROIs, indicating that the BOLD activity decreased from confident old to the unconfident judgments, then increased for the confident new judgments (see Fig 3). In addition, a linear trend was also identified in left hippocampus, AG, MTG and PCC.

Table 5. Results of one-way ANOVA and trend analysis across confidence judgments in each recollection-sensitive ROI (bold values denote significance at the p < .05 level).

Region	Main effect of response type	Linear trend	Quadratic trend
Left_PHC	**F(3, 111) = 5.87, p = .001, partial η^2 = .137**	F(1, 37) = 1.67, p = .204, partial η^2 = .043	**F(1, 37) = 15.65, p < .001, partial η^2 = .297**
Left_Hippocampus	**F(2.23, 82.37) = 7.55, p = .001, partial η^2 = .169**	**F(1, 37) = 4.23, p = .047, partial η^2 = .103**	**F(1, 37) = 20.30, p < .001, partial η^2 = .354**
Left_mPFC	**F(3, 111) = 7.12, p < .001, partial η^2 = .161**	F(1, 37) = .01, p = .917, partial η^2 < .001	**F(1, 37) = 16.04, p < .001, partial η^2 = .302**
Left_AG	**F(2.48, 91.86) = 32.07, p < .001, partial η^2 = .464**	**F(1, 37) = 37.08, p < .001, partial η^2 = .501**	**F(1, 37) = 38.04, p < .001, partial η^2 = .507**

Left_MTG	**F(3, 111) = 12.51, p < .001, partial η^2 = .253**	**F(1, 37) = 4.79, p = .035, partial η^2 = .115**	**F(1, 37) = 45.18, p < .001, partial η^2 = .550**
Left_PCC	**F(3, 111) = 22.13, p < .001, partial η^2 = .374**	**F(1, 37) = 25.20, p < .001, partial η^2 = .405**	**F(1, 37) = 41.69, p < .001, partial η^2 = .530**
Right_Hippocampus	**F(3, 111) = 6.81, p < .001, partial η^2 = .156**	F(1, 37) = .14, p = .707, partial η^2 = .004	**F(1, 37) = 16.71, p < .001, partial η^2 = .311**

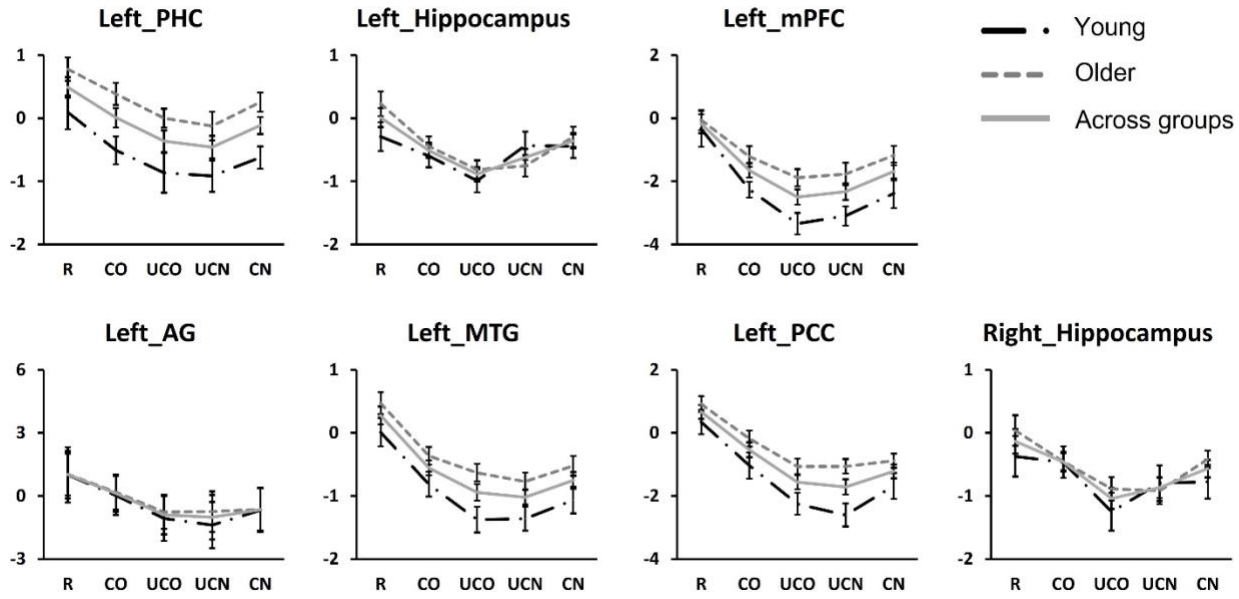

Fig 3. BOLD activity associated with each class of recognition judgement (R, confident old, unconfident old, confident new, unconfident new) for the different recollection ROIs. Error bars represent standard error of means. CO: confident old; UCO: unconfident old; UCN: unconfident new; CN: confident new.

3.4.2 Non-PRC familiarity ROIs

The outcome of the ANOVA conducted to examine the familiarity effects across the non-PRC ROIs is shown in Table 6, first panel. As is evident from the table, the main effects of response type and region were significant. In addition, region significantly interacted with both

age group and response type, but the interactions between group and response type, and the three-way interaction, were both far from significant. Analyses following up the region x group interaction identified greater mean BOLD activity for young than for older adults in dmPFC [$t(36) = 2.53$, $p = .016$; other regions, $ts < 1.27$, $ps > .214$, $BF_{10}s < .60$]. However, the difference in dmPFC did not survive Holm-Bonferroni correction.

Table 6. Results of mixed-design ANOVA at non-PRC familiarity-sensitive ROIs (bold values denote significance at the $p < .05$ level).

2 (age group) x 4 (response type: four confidence ratings) x 6 (region) ANOVA	
Response type	**$F(3, 108) = 35.08$, $p < .001$, partial $\eta^2 = .494$**
Group	$F(1, 36) = .34$, $p = .566$, partial $\eta^2 = .009$, $BF_{incl} = .28$
Region	**$F(3.90, 140.23) = 21.68$, $p < .001$, partial $\eta^2 = .376$**
Response type x group	$F(3, 108) = .98$, $p = .405$, partial $\eta^2 = .027$, $BF_{incl} = .04$
Region x group	**$F(3.90, 140.23) = 2.82$, $p = .028$, partial $\eta^2 = .073$**
Response type x region	**$F(8.43, 303.37) = 5.04$, $p < .001$, partial $\eta^2 = .123$**
Response type x region x group	$F(8.43, 303.37) = .21$, $p = .992$, partial $\eta^2 = .006$, $BF_{incl} = 8.76e^{-4}$
2 (age group) x 2 (response type: R, confident old) x 6 (region) ANOVA	
Response type	$F(1, 36) = .96$, $p = .333$, partial $\eta^2 = .026$, $BF_{incl} = .35$
Group	$F(1, 36) = .15$, $p = .702$, partial $\eta^2 = .004$, , $BF_{incl} = .26$
Region	**$F(3.44, 123.77) = 23.59$, $p < .001$, partial $\eta^2 = .396$**
Response type x group	$F(1, 36) = 2.16$, $p = .150$, partial $\eta^2 = .057$, $BF_{incl} = .87$
Region x group	$F(3.44, 123.77) = 1.73$, $p = .156$, partial $\eta^2 = .046$, $BF_{incl} = 2.01$
Response type x region	$F(3.74, 134.47) = 2.28$, $p = .068$, partial $\eta^2 = .060$, $BF_{incl} = .02$
Response type x region x group	**$F(3.74, 134.47) = 2.52$, $p = .048$, partial $\eta^2 = .065$**

To unpack the response type x region interaction, for each ROI we conducted a one-way ANOVA across the four confidence judgments and performed linear and quadratic trend

analyses. As is evident from Table 7 (see also Fig 4), a main effect of response type was identified in all ROIs. Trend analyses reveal a significant linear tread in each ROI, manifesting as a step-wise decrease in activity from confident old to confident new judgments in all regions other than left and right caudate and left alPFC. In these latter three regions, the linear effect was accompanied by a significant quadratic trend.

Table 7. Results of one-way ANOVA and trend analysis across confidence judgments in each non-PRC familiarity-sensitive ROI (bold values denote significance at the $p < .05$ level).

Region	Main effect of response type	Linear trend	Quadratic trend
Left_Caudatae	**$F(3, 111) = 12.36$, $p < .001$, partial $\eta^2 = .250$**	**$F(1, 37) = 22.43$, $p < .001$, partial $\eta^2 = .377$**	**$F(1, 37) = 8.98$, $p = .005$, partial $\eta^2 = .195$**
Left_IPS	**$F(2.38, 88.18) = 27.83$, $p < .001$, partial $\eta^2 = .429$**	**$F(1, 37) = 69.54$, $p < .001$, partial $\eta^2 = .653$**	$F(1, 37) = .56$, $p = .459$, partial $\eta^2 = .015$
Left_alPFC	**$F(2.30, 84.95) = 25.07$, $p < .001$, partial $\eta^2 = .404$**	**$F(1, 37) = .40.35$, $p < .001$, partial $\eta^2 = .522$**	**$F(1, 37) = 10.66$, $p = .002$, partial $\eta^2 = .224$**
Left_Precuneus	**$F(3, 111) = 34.23$, $p < .001$, partial $\eta^2 = .481$**	**$F(1, 37) = 114.26$, $p < .001$, partial $\eta^2 = .755$**	$F(1, 37) = 1.08$, $p = .306$, partial $\eta^2 = .028$
Left_dmPFC	**$F(3, 111) = 17.63$, $p < .001$, partial $\eta^2 = .323$**	**$F(1, 37) = 41.13$, $p < .001$, partial $\eta^2 = .526$**	$F(1, 37) = 2.42$, $p = .129$, partial $\eta^2 = .061$
Right_Caudate	**$F(3, 111) = 10.55$, $p < .001$, partial $\eta^2 = .222$**	**$F(1, 37) = 21.97$, $p < .001$, partial $\eta^2 = .373$**	**$F(1, 37) = 9.03$, $p = .005$, partial $\eta^2 = .196$**

Results of the ANOVA employed to examine recollection effects in the non-PRC familiarity ROIs are given in the second panel of Table 6. As is evident from the table, both the main effect of region and the response type x region x group interaction were significant. To follow up the three-way interaction, a 2 (response type) x 2 (age group) mixed ANOVA was conducted for each ROI. The main effect of response type was significant for the right caudate only, albeit only before correction for multiple comparisons [for right caudate, $F(1, 36) = 5.63$, $p = .023$, partial $\eta^2 = .135$; for other regions, $Fs < 1.50$, $ps > .229$, partial $\eta^2 s < .041$, $BF_{incl}s$

< .60]. In addition, a significant response type x age group interaction was identified in the alPFC before correction for multiple comparisons [$F(1, 36) = 6.73$, $p = .014$, partial $\eta^2 = .157$, for other regions, $Fs < 2.80$, $ps > .104$, partial η^2s < .073, BF_{incl}s < .97]. However, follow-up analyses revealed no evidence of a recollection effect in this region in either age group [for young, $t(15) = 1.82$, $p = .090$, $BF_{10} = .97$; for older, $t(21) = 1.86$, $p = .077$, $BF_{10} = .97$].

3.4.3 PRC

Results of the ANOVAs conducted on the parameter estimates derived from the PRC are shown in Table 8. As is evident from the first panel of the table, the main effect of response type was significant but no effect involving the factors of group or hemisphere attained significance. A significant quadratic trend was identified in bilateral PRC [$F(1, 36) = 18.24$, $p < .001$, partial $\eta^2 = .336$; for linear trend, $F(1, 36) = .18$, $p = .677$, partial $\eta^2 = .005$], indicating that the impression given by Fig 4 for BOLD activity to decline between confident old and unconfident old judgments, and then to increase with confidence for new judgments, was reliable.

The outcome of the ANOVA examining recollection effects in the PRC is given in the second panel of Table 8. As is evident from the table, there was a main effect of response type that was modified by a response type x hemisphere interaction. To follow up the interaction, pairwise comparisons were conducted between R and confident old judgments in each hemisphere. The comparisons revealed significant a recollection effect for left but not right PRC [for left PRC, $t(37) = 2.88$, $p = .007$, for right PRC, $t(37) = .95$, $p = .350$, $BF_{10} = .27$].

Table 8. Results of mixed-design ANOVA across left and right PRC (bold values denote significance at the $p < .05$ level).

2 (age group) x 4 (response type: four confidence ratings) x 2 (hemisphere) ANOVA	
Response type	**$F(3, 108) = 6.42$, $p < .001$, partial $\eta^2 = .151$**
Group	$F(1, 36) = .01$, $p = .942$, partial $\eta^2 < .001$, $BF_{incl} = .35$
Hemisphere	$F(1, 36) = 2.35$, $p = .134$, partial $\eta^2 = .061$, $BF_{incl} = 1.65$
Response type x group	$F(3, 108) = .68$, $p = .567$, partial $\eta^2 = .019$, $BF_{incl} = .11$
Hemisphere x group	$F(1, 36) = .70$, $p = .410$, partial $\eta^2 = .019$, $BF_{incl} = .32$
Response type x hemisphere	$F(3, 108) = .26$, $p = .852$, partial $\eta^2 = .007$, $BF_{incl} = .04$
Response type x hemisphere x group	$F(3, 108) = 1.53$, $p = .211$, partial $\eta^2 = .041$, $BF_{incl} = .09$
2 (age group) x 2 (response type: R, confident old) x 2 (hemisphere) ANOVA	
Response type	**$F(1, 36) = 4.23$, $p = .047$, partial $\eta^2 = .105$**
Group	$F(1, 36) = .02$, $p = .891$, partial $\eta^2 = .001$, $BF_{incl} = .36$
Hemisphere	$F(1, 36) = .03$, $p = .876$, partial $\eta^2 = .001$, $BF_{incl} = .19$
Response type x group	$F(1, 36) = .79$, $p = .380$, partial $\eta^2 = .021$, $BF_{incl} = .40$
Hemisphere x group	$F(1, 36) = .62$, $p = .437$, partial $\eta^2 = .017$, $BF_{incl} = .32$
Response type x hemisphere	**$F(1, 36) = 4.85$, $p = .034$, partial $\eta^2 = .119$**
Response type x hemisphere x group	$F(1, 36) = .15$, $p = .705$, partial $\eta^2 = .004$, $BF_{incl} = .31$

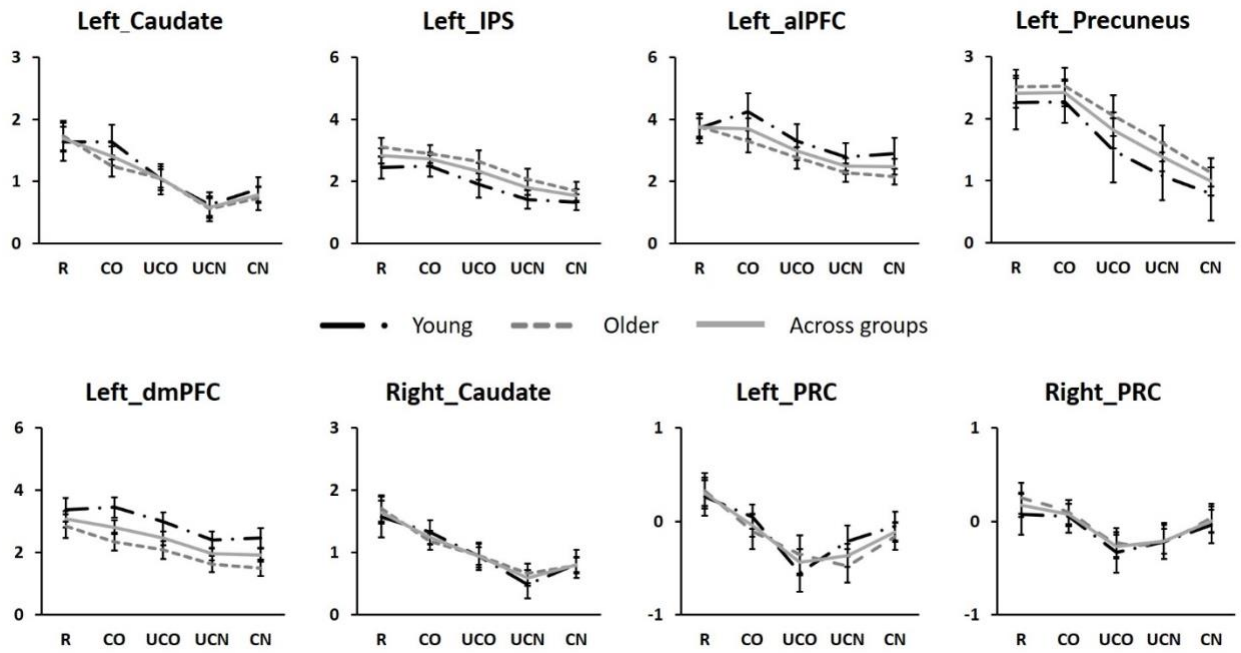

Fig 4. Activity associated with each class of recognition judgment (R, confident old, unconfident old, confident new, unconfident new) in each familiarity ROI. Error bars represent standard error of means. CO: confident old; UCO: unconfident old; UCN: unconfident new; CN: confident new.

To complement the trend analyses of the PRC and non-PRC familiarity effects reported above, pairwise comparisons between the different confidence ratings were also conducted. The findings, which strongly converge with the more parsimonious analyses presented here, can be found in the supplemental materials.

3.4.4 Relationships between fMRI effects and recognition performance

Full details of the regression analyses examining relationships between the different fMRI effects and memory performance are reported in supplemental materials (Supplemental Tables 4-6). As is evident from those tables, the only analyses to reveal evidence of such a

relationship were those for the left and right PRC familiarity-attenuation effects, which demonstrated age-invariant correlations with both pF (left: β = .51, partial r = .52, p = .001; right: β = .41, partial r = .42, p = .010; for the functional effect x age group interaction: left: β = .16, p = .653, BF_{incl} = .84; right: β = .41, p = .196, BF_{incl} = 1.92) and mean memory accuracy (left: β = .51, partial r = .54, p = .001, right: β = .37, partial r = .40, p = .015; for the functional effect x age group interaction, left: β = .08, p = .816, BF_{incl} = .97; right: β = .23, p = .462, BF_{incl} = 1.47. Scatter plots illustrating these relationships (after controlling for the effects of age group) can be found in Fig 5.

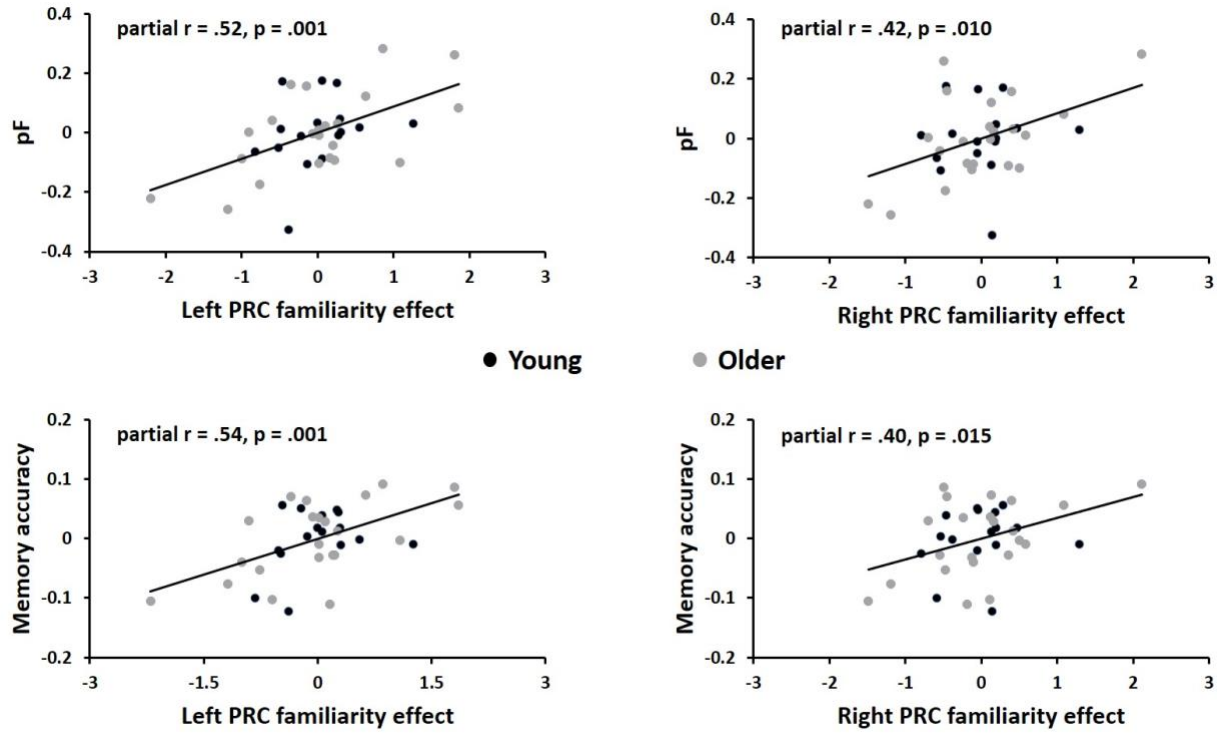

Fig 5. Scatter plots of the relationships (controlling for age group) between left and right PRC familiarity-attenuation effects (confident new – unconfident old) and estimates of familiarity (top) and mean memory accuracy (bottom).

4. Discussion

We employed a modified remember-know paradigm to investigate the effects of age on fMRI correlates of recognition memory. On the experimental memory test, young and older adults demonstrated comparable memory performance as indexed by estimates of recollection and familiarity. However, older adults demonstrated significantly lower recognition memory accuracy than young adults. Age-invariant functional effects were evident in two sets of brain regions which have consistently been reported to demonstrate sensitivity to recollection and familiarity respectively. Additionally, the magnitudes of familiarity-attenuation effects in PRC demonstrated age-invariant correlations (although see below) with estimates of familiarity and memory accuracy. These findings are discussed in more detail below.

4.1 Behavioral findings

Although age-related decline in recollection and, less consistently, familiarity estimates have been reported in studies employing the remember-know procedure (e.g. Alghamdi & Rugg, 2020; Duarte et al., 2010; Howard et al., 2006; Wang et al., 2012, 2016; for a review, see Koen & Yonelinas, 2014), here we did not identify significant age differences for either estimate. As suggested by a reviewer, the relatively short study-test delay that was employed in the present study might have contributed to these null findings, although they are not without precedent. For example, in two prior studies (Duarte et al., 2006, 2008), statistically equivalent recollection estimates were reported in young adults and a subgroup of older adults who demonstrated relatively high item memory on the experimental memory test (see Mark & Rugg, 1998 for seemingly similar findings). The present findings for familiarity are arguably more consistent with numerous prior studies that reported little or no decline in familiarity with advancing age (e.g. Alghamdi & Rugg, 2020; Caldwell & Masson, 2001; Koen & Yonelinas, 2016; see Koen &

Yonelinas, 2014 for review). Of importance, although not reflected in the recollection and familiarity estimates, higher false alarm rates in the older adults were evident for both R and confident old judgments, resulting in significantly lower overall memory accuracy in this group (cf. Wang et al., 2012). This finding suggests that the memory accuracy metric might be a more sensitive measure of age differences in recognition memory than theoretically motivated estimates of recollection and familiarity and highlights the utility of examining the effects of age on memory performance with multiple measures.

4.2 fMRI findings

4.2.1 Recollection effects

One of the primary goals of the present study was to examine potential age differences in the fMRI correlates of recollection. Employing an approach that focused on ROIs defined *a priori*, we failed to detect any age differences in these correlates. These age-invariant effects were consistent with the results of an exploratory whole brain analyses. The present null findings stand in contrast to the age differences in fMRI recollection effects reported by Daselaar et al. (2006) and Angel et al. (2013). While it is tempting to attribute the attenuated recollection effects in older adults reported by Daselaar et al. (2006) to the lower recollection performance of these participants relative to the young group, this explanation does not readily apply to the findings of Angel et al. (2013), where estimates of recollection were statistically matched across age groups. However, the older participants in Angel et al. (2013) adopted a more liberal criterion for 'remember' responses than young adults. Thus, it is possible that a higher proportion of trials supported primarily by familiarity found their way into the remember judgments of older than young participants, disproportionately diluting their recollection effects (and, presumably, impacting familiarity effects also, see below). In contrast to Angel et al. (2013), we observed

only a non-significant trend towards an age difference in response bias for R judgments [$t(36) = 1.78$, $p = .083$, $BF_{10} = 1.08$]. The present findings are however consistent with another report of age-invariance in the fMRI correlates of recollection (Wang et al., 2016; for null findings after controlling for age differences in memory performance, see de Chastelaine et al., 2016). Together, the present and these more recent findings suggest that fMRI recollection effects are little affected by advancing age.

As was just mentioned, we examined recollection effect in *a priori* defined ROIs. Consistent with prior studies (e.g. de Chastelaine et al., 2016; Johnson et al., 2013), robust effects, indexed by higher activity for R than confident old judgments, were evident in all members of core recollection network except for the right hippocampus. These findings converged with the outcomes of the exploratory whole brain analyses. The reason for the lack of recollection effects in the right hippocampus is unclear, especially in light of findings from prior studies that identified robust recollection effects in this region for verbal materials (e.g. de Chastelaine et al., 2016; Johnson et al., 2013; Wang et al., 2016: Yonelinas et al., 2005).

The only familiarity-sensitive ROI to demonstrate a significant recollection effect (operationalized by the R > confident old contrast) was the left PRC, echoing two prior reports of recollection effects in this region (Staresina et al., 2012; Wang et al., 2014). Together with these prior reports, the present findings challenge the proposal that the medial temporal lobe (MTL) demonstrates a strong functional dissociation between a familiarity-sensitive PRC and a recollection-sensitive hippocampus (e.g. Aggleton & Brown, 2006; Diana et al., 2007; Eichenbaum et al., 2007). However, the finding that both the PRC and the hippocampus manifest recollection effects does not mean that these regions share the same functional role. For example, using time-resolved BOLD data and direct intra-cranial recordings, Staresina et al. (2012)

reported that while novelty effects preceded recollection effects in the PRC, the temporal ordering of the two effects was reversed in the hippocampus. Furthermore, the onset of novelty effects in PRC preceded the onset of hippocampal recollection effects, which in turn preceded the onset of the PRC recollection effects. These findings are consistent with the proposals that PRC acts as a 'gatekeeper' to the hippocampus (Fernández & Tendolkar, 2006), and that PRC recollection effects reflect a re-entrant signal that originates in the hippocampus (Staresina et al., 2012).

4.2.2 Familiarity effects

As in the case of recollection effects, we were unable to identify any effects of age on fMRI familiarity effects. The finding of age-invariant non-PRC familiarity effects is consistent with the report of Daselaar et al. (2006). By contrast, Angel et al. (2013) identified attenuated such effects in older adults. Equally discordant, Daselaar et al. (2006) reported enhanced familiarity-attenuation effects in the PRC of their older adults, whereas we could find no evidence of an age difference in these effects. Behavioral estimates of familiarity strength did not differ significantly between age groups in either Daselaar et al. (2006), Angel et al. (2013) or the present study, suggesting that the disparate findings across these studies is not attributable to differing patterns of performance across age groups [echoing our prior comments in respect of divergent findings for recollection effects, the disparities between our findings and those from Angel et al (2013) might arise from the more liberal response bias for 'remember' judgments in the older participants in that study]. The present findings are consistent with those of other recent studies in which age-invariant fMRI correlates of familiarity were reported (de Chastelaine et al., 2017; Wang & Giovanello, 2016). Together with these prior findings, the present results suggest that, like correlates of recollection, neural correlates of familiarity can also be relatively

insensitive to increasing age, at least up to the mid-70s (see Wang et al., 2009 for a study of the effects of very advanced age on the neural correlates of recognition memory).

Consistent with prior findings (Johnson et al., 2013; Montaldi et al., 2006; Yonelinas et al., 2005), we identified a significant positive linear trend in BOLD activity that tracked familiarity strength in each of the non-PRC familiarity ROIs (see Table 7 and Fig 4). By contrast, BOLD activity in the PRC demonstrated a quadratic (U shaped) pattern, such that it decreased between confident and unconfident old judgments and showed a graded increase thereafter (see Fig 4, and Supplemental Table 3 for convergent pairwise tests). The higher BOLD activity identified for confident new than unconfident old judgments is consistent with prior reports of familiarity-related attenuation of activity in PRC (e.g. de Chastelaine et al., 2017; Henson et al., 2003; Montaldi et al., 2006; Staresina et al., 2012; for a review, see Brown & Banks, 2015). The higher BOLD activity for confident old than the unconfident judgments is however decidedly inconsistent with the proposal that PRC activity declines with increasing familiarity strength. We interpret this finding as a reflection of the fact that confident old judgments were supported by both familiarity *and* recollection, as is discussed in the following paragraph.

We identified a significant quadratic trend across confidence ratings in every core recollection ROI. An obvious, if *post-hoc*, explanation for the finding of greater BOLD activity for confident than unconfident old judgments is that confident old judgments were supported by recollection as well as familiarity, that is, these judgments were not a 'process pure' index of strong familiarity. For instance, participants might have adopted a relatively strict criterion for endorsing test items as remembered, such that some proportion of confident old judgments were supported by a sub-criterial recollection signal. Of course, an alternative possibility is that the core recollection network does not honor the distinction between recollection and familiarity but,

rather, the distinction between 'strong' and 'weak' memories, as has already been argued to be the case for the hippocampus (Squire, Wixted, & Clark, 2007; Wixted et al., 2010). While this possibility cannot be ruled out on the basis of the current findings, evidence from prior studies strongly opposes it (for reviews, see Rugg et al., 2012; Rugg & Vilberg, 2013). Notably, both the hippocampus and cortical members of the core recollection network have been reported to respond in a graded manner specifically to the amount or fidelity of the contextual information retrieved about a prior event (Diana et al., 2010; Thakral et al., 2015; Yu et al., 2012).

Relative to the activity associated with unconfident old and new judgments, activity elicited by items attracting confident new judgments was also elevated in recollection ROIs. This finding is reminiscent of prior reports of hippocampal 'novelty effects', which have been assumed to reflect memory encoding operations engaged by situationally novel information (e.g. Bowman & Dennis, 2015; de Chastelaine et al., 2017; Kafkas & Montaldi, 2014; Stark & Okado, 2003, for reviews, see Kafkas & Montaldi, 2018; Nyberg, 2005). Whereas novelty effects have been examined most frequently in the hippocampus, it has been proposed that novelty processing engages a broadly distributed brain network. For example, Kafkas and Montaldi (2014) described a novelty network that included bilateral hippocampus, PRC, PHC, fusiform cortex, left MTG, and right orbitofrontal and middle occipital cortex - regions that demonstrate substantial overlap with the core recollection network. It remains to be established what the functional roles of these regions might be in novelty processing, and how these roles complement their potential roles in the support of recollection [but see Cabeza et al. (2012) for a proposal relevant specifically to the AG].

4.2.3 Relationships between functional effects and memory performance

Contrary to the findings of prior studies (de Chastelaine et al., 2016, 2017; Hou et al., 2020), we did not find evidence of a relationship between either fMRI recollection or non-PRC familiarity effects and memory performance. These null findings might result from the combination of our modest sample size and, perhaps, relatively low variance in memory performance across the participants. However, as we discuss below, significant relationships between PRC familiarity effects and memory performance were evident.

Consistent with the report of de Chastelaine et al. (2017), we identified robust, seemingly age-invariant relationships between PRC familiarity-attenuation effects and behavioral estimates of familiarity strength and recognition accuracy (note that these two memory metrics were strongly correlated, $r = .80$, $p < .001$, while the correlations of each metric with pR was small and far from significance; $rs < .19$, $ps > .265$). By contrast, we were unable to identify a significant relationship between PRC familiarity effects and estimates of recollection (see supplemental materials). Overall, these findings highlight the specificity of the relationship between PRC familiarity-attenuation effects and familiarity-driven memory performance and suggest that this relationship is stable in the face of advancing age. We note however that caution is necessary in concluding that the present findings are strongly indicative that this relationship is age-invariant: although the interaction terms in the relevant regression models were far from significant (min $p = .196$), the accompanying Bayes factors were close to or exceeded 1.

4.2.4 Item-related BOLD activity

We note that in several core recollection regions item-related activity (activity elicited by all classes of test item relative to baseline) was higher in the older than the young age group (see. Fig 3). This finding is reminiscent of prior reports of enhanced BOLD activity in older relative to young adults (e.g. Grady et al., 2010; Persson et al., 2007; Wang et al., 2016; see also Langnes et

al., 2019 for findings in participants aged from 20-70 yrs old). Of importance, this global effect of age on item-related BOLD signals was orthogonal to the magnitude of the recollection effects in the same regions which, as already discussed, were age-invariant (very similar findings were reported by Wang et al. 2016). The functional significance of this 'global' effect of age is obscure. One possibility is that the effect reflects an overall enhancement of activity in older adults in the face of neural inefficiency during the memory task (Duverne et al., 2008). However, it should be noted that the higher item-related activity in older adults appears not to be maintained over later life. In a recent study, Langnes et al. (2019) reported that the magnitude of hippocampal activity decreased with increasing age from around 70 years. Future studies are needed to examine whether decreases in item-related hippocampal activity in adults in the seventh decade and beyond presage weaker recollection effects.

4.3 Limitations

There are some limitations to the present study. First, the modest sample size could have limited our ability to detect the potential brain-behavior relationships and subtle but theoretically important age differences. Of importance, some caution is required in respect of the null findings for age that we consistently report. To buttress these findings we estimated the corresponding Bayes factors in each of these cases. With the exception of the results for the relationships with memory performance discussed above, in the cases of most theoretical importance (interactions involving the factors of response and age group in the analyses of recollection and familiarity effects in the corresponding ROIs) the factors uniformly provide substantial or strong evidence in favor of the null hypothesis. A second limitation of the present study (shared with almost all fMRI studies of cognitive aging) is that the functional memory measures employed here might have been confounded with age differences in vascular factors mediating between neural activity

and the BOLD signal (e.g. Liu et al., 2013; Lu et al., 2011; Tsvetanov et al., 2015). Since we did not control for these factors, their influence on the functional findings cannot be ruled out. Lastly, our employment of a cross-sectional rather than a longitudinal design means that we cannot distinguish between effects of age and confounding factors such as cohort effects or selection bias (Rugg, 2017).

5. Conclusion

In conclusion, using an experimental paradigm that, to our knowledge, has not previously been employed in fMRI studies of cognitive aging, we identified robust, age-invariant neural correlates of recollection and familiarity-driven recognition memory. Together with prior reports (de Chastelaine et al., 2016, 2017; Wang et al., 2016), these findings suggest that the neural correlates of recollection- and familiarity-based memory judgments are largely stable across much of the healthy adult lifespan.

Acknowledgments

This work was supported by the National Institute on Aging (grant numbers 5P50AG016573, 1RF1AG039103). We thank our experimental participants for their participation.

References

Aggleton, J.P., & Brown, M.W. (2006). Interleaving brain systems for episodic and recognition memory. Trends in cognitive sciences, 10, 455–463.

Alghamdi, S.A., & Rugg, M.D. (2020). The effect of age on recollection is not moderated by differential estimation methods. Memory, 28, 1067–1077.

Andrade, A., Paradis, A.L., Rouquette, S., & Poline, J.B. (1999). Ambiguous results in functional neuroimaging data analysis due to covariate correlation. Neuroimage 10, 483–486.

Angel, L., Bastin, C., Genon, S., Balteau, E., Phillips, C., Luxen, A., Maquet, P., Salmon, E., & Collette, F. (2013). Differential effects of aging on the neural correlates of recollection and familiarity. Cortex, 49, 1585–1597.

Angel, L., Bastin, C., Genon, S., Salmon, E., Fay, S., Balteau, E., Maquet, P., Luxen, A., Isingrini, M., & Collette, F. (2016). Neural correlates of successful memory retrieval in aging: Do executive functioning and task difficulty matter? Brain Research, 1631, 53–71.

Beck, A. T., Ward, C.H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. Archives of general psychiatry, 4, 561–571.

Bowman, C.R., & Dennis, N.A. (2015). Age differences in the neural correlates of novelty processing: The effects of item-relatedness. Brain research, 1612, 2–15.

Brown, M.W., & Banks, P.J. (2015). In search of a recognition memory engram. Neuroscience & Biobehavioral Reviews, 50, 12–28.

Cabeza, R., Ciaramelli, E., & Moscovitch, M. (2012). Cognitive contributions of the ventral parietal cortex: an integrative theoretical account. Trends in cognitive sciences, 16, 338–352.

Caldwell, J.I., & Masson, M.E. (2001). Conscious and unconscious influences of memory for object location. Memory & Cognition, 29, 285–295.

Daselaar, S.M., Fleck, M.S., Dobbins, I.G., Madden, D.J., & Cabeza, R. (2006). Effects of healthy aging on hippocampal and rhinal memory functions: An event-related fMRI study. Cerebral Cortex, 16, 1771–1782.

de Chastelaine, M., Mattson, J.T., Wang, T.H., Donley, B.E., & Rugg, M.D. (2015). Sensitivity of negative subsequent memory and task-negative effects to age and associative memory performance. Brain research, 1612, 16–29.

de Chastelaine, M., Mattson, J.T., Wang, T.H., Donley, B.E., & Rugg, M.D. (2016). The neural correlates of recollection and retrieval monitoring: Relationships with age and recollection performance. NeuroImage, 138, 164–175.

de Chastelaine, M., Mattson, J.T., Wang, T.H., Donley, B.E., & Rugg, M.D. (2017). Independent contributions of fMRI familiarity and novelty effects to recognition memory and their stability across the adult lifespan. NeuroImage, 156, 340–351.

Diana, R.A., Yonelinas, A.P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. Trends in cognitive sciences, 11, 379–386.

Diana, R.A., Yonelinas, A.P., & Ranganath, C. (2010). Medial temporal lobe activity during source retrieval reflects information type, not memory strength. Journal of cognitive neuroscience, 22, 1808–1818.

Duarte, A., Graham, K.S., & Henson, R.N. (2010). Age-related changes in neural activity associated with familiarity, recollection and false recognition. Neurobiology of Aging, 31, 1814–1830.

Duarte, A., Henson, R.N., & Graham, K.S. (2008). The effects of aging on the neural correlates of subjective and objective recollection. Cerebral Cortex, 18, 2169–2180.

Duarte, A., Ranganath, C., Trujillo, C., & Knight, R.T. (2006). Intact recollection memory in high-performing older adults: ERP and behavioral evidence. Journal of cognitive neuroscience, 18, 33–47.

Dulas, M. R., & Duarte, A. (2012). The effects of aging on material-independent and material-dependent neural correlates of source memory retrieval. Cerebral Cortex, 22, 37-50.

Dulas, M.R., & Duarte, A. (2014). Aging affects the interaction between attentional control and source memory: an fMRI study. Journal of cognitive neuroscience, 26, 2653–2669.

Duverne, S., Habibi, A., & Rugg, M. D. (2008). Regional specificity of age effects on the neural correlates of episodic retrieval. Neurobiology of aging, 29, 1902-1916.

Eichenbaum, H., Yonelinas, A.P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. Annual Review of Neuroscience, 30, 123–152.

Elward, R.L., & Rugg, M.D. (2015). Retrieval goal modulates memory for context. Journal of cognitive neuroscience, 27, 2529–2540.

Fernández, G., & Tendolkar, I. (2006). The rhinal cortex: ‘gatekeeper’of the declarative memory system. Trends in cognitive sciences, 10, 358–362.

Folville, A., Bahri, M.A., Delhaye, E., Salmon, E., D’Argembeau, A., & Bastin, C. (2020). Age-related differences in the neural correlates of vivid remembering. NeuroImage, 206, 116336.

Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-related fMRI: characterizing differential responses. Neuroimage, 7, 30–40.

Gardiner, J.M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E.E. Tulving, E. Fergus & I.M. Craik (Eds.), The Oxford handbook of memory (pp. 229–244). New York: Oxford University Press.

Giovanello, K.S., & Dew, I.T.Z. (2015). Relational memory and its relevance to aging. In: D. R. Addis, M. Barense, & A. Duarte (Eds.), The Wiley handbook on the cognitive neuroscience of memory (pp. 371–392). West Sussex: John Wiley.

Grady, C.L., Protzner, A.B., Kovacevic, N., Strother, S.C., Afshin-Pour, B., Wojtowicz, M., Anderson, J.A., Churchill, N., & McIntosh, A. R. (2010). A multivariate analysis of age-related differences in default mode and task-positive networks across multiple cognitive domains. Cerebral cortex, 20, 1432–1447.

Greenhouse, S.W., & Geisser, S. (1959). On methods in the analysis of profile data. Psychometrika, 24, 95–112.

Henson, R.N.A., Cansino, S., Herron, J. E., Robb, W.G.K., & Rugg, M.D. (2003). A familiarity signal in human anterior medial temporal cortex?. Hippocampus, 13, 301–304.

Horn, M., Jardri, R., D'Hondt, F., Vaiva, G., Thomas, P., & Pins, D. (2016). The multiple neural networks of familiarity: A meta-analysis of functional imaging studies. Cognitive, Affective and Behavioral Neuroscience, 16, 176–190.

Hou, M., de Chastelaine, M., Jayakumar, M., Donley, B.E., & Rugg, M.D. (2020). Recollection-related hippocampal fMRI effects predict longitudinal memory change in healthy older adults. Neuropsychologia, 146, 107537.

Howard, M.W., Bessette-Symons, B., Zhang, Y., & Hoyer, W.J. (2006). Aging selectively impairs recollection in recognition memory for pictures: evidence from modeling and receiver operating characteristic curves. Psychology and aging, 21, 96–106.

Jeffreys, H. (1961). Theory of probability (3rd ed.). Oxford: Oxford University Press, Clarendon Press.

Johnson, J.D., Suzuki, M., & Rugg, M.D. (2013). Recollection, familiarity, and content-sensitivity in lateral parietal cortex: a high-resolution fMRI study. Frontiers in Human Neuroscience, 7, 1–15.

Kafkas, A., & Montaldi, D. (2014). Two separate, but interacting, neural systems for familiarity and novelty detection: A dual-route mechanism. Hippocampus, 24, 516–527.

Kafkas, A., & Montaldi, D. (2018). How do memory systems detect and respond to novelty? Neuroscience letters, 680, 60–68.

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

Kim, H. (2013). Differential neural activity in the recognition of old versus new events: An activation likelihood estimation meta-analysis. Human brain mapping, 34, 814–836.

Kirwan, C.B., & Stark, C.E. (2004). Medial temporal lobe activation during encoding and retrieval of novel face-name pairs. Hippocampus, 14, 919–930.

Kluger, A., Ferris, S.H., Golomb, J., Mittelman, M.S., & Reisberg, B. (1999). Neuropsychological prediction of decline to dementia in nondemented elderly. Journal of geriatric psychiatry and neurology, 12, 168–179.

Koen, J.D., & Yonelinas, A.P. (2014). The effects of healthy aging, amnestic mild cognitive impairment, and Alzheimer's disease on recollection and familiarity: A meta-analytic review. Neuropsychology review, 24, 332–354.

Koen, J.D., & Yonelinas, A.P. (2016). Recollection, not familiarity, decreases in healthy ageing: Converging evidence from four estimation methods. Memory, 24, 75–88.

Kučera, H., & Francis, W.N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.

Kukolja, J., Thiel, C.M., Wilms, M., Mirzazade, S., & Fink, G.R. (2009). Ageing-related changes of neural activity associated with spatial contextual memory. Neurobiology of Aging, 30, 630–645.

Langnes, E., Vidal-Piñeiro, D., Sneve, M. H., Amlien, I. K., Walhovd, K. B., & Fjell, A. M. (2019). Development and decline of the hippocampal long-axis specialization and differentiation during encoding and retrieval of episodic memories. Cerebral Cortex, 29, 3398-3414.

Liu, P., Hebrank, A.C., Rodrigue, K.M., Kennedy, K.M., Section, J., Park, D.C., & Lu, H. (2013). Age-related differences in memory-encoding fMRI responses after accounting for decline in vascular reactivity. Neuroimage, 78, 415–425.

Lu, H., Xu, F., Rodrigue, K.M., Kennedy, K.M., Cheng, Y., Flicker, B., Hebrank, A.C., Uh, J., & Park, D.C. (2011). Alterations in cerebral metabolic rate and blood supply across the adult lifespan. Cerebral Cortex 21, 1426–1434.

Mark, R.E., & Rugg, M.D. (1998). Age effects on brain activity associated with episodic memory retrieval. An electrophysiological study. Brain. 121, 861–873.

Mathôt, S. (2017). Bayes like a Baws: Interpreting Bayesian repeated measures in JASP. Cognitive Science and More. Retrieved from https://www. cogsci. nl/blog/interpreting-bayesian-repeated-measures-in-jasp.

Montaldi, D., Spencer, T.J., Roberts, N., & Mayes, A.R. (2006). The neural system that mediates familiarity memory. Hippocampus, 16, 504–520.

Norman, M.A., Evans, J.D., Miller, W.S., & Heaton, R.K. (2000). Demographically corrected norms for the California verbal learning test. Journal of Clinical and Experimental Neuropsychology, 22, 80–94.

Nyberg, L. (2005). Any novelty in hippocampal formation and memory? Current opinion in neurology, 18, 424–428.

Persson, J., Lustig, C., Nelson, J.K., & Reuter-Lorenz, P.A. (2007). Age differences in deactivation: a link to cognitive control? Journal of cognitive neuroscience, 19, 1021–1032.

Rugg, M.D., & Morcom, A.M. (2005). The relationship between brain activity, cognitive performance, and aging: the case of memory. Cognitive neuroscience of aging: Linking cognitive and cerebral aging, 132–154.

Rugg, M.D., & Vilberg, K.L. (2013). Brain networks underlying episodic memory retrieval. Current Opinion in Neurobiology, 23, 255–260.

Rugg, M.D., Vilberg, K.L., Mattson, J.T., Sarah, S.Y., Johnson, J.D., & Suzuki,M. (2012). Item memory, context memory and the hippocampus: fMRI evidence. Neuropsychologia, 50, 3070–3079.

Rugg, M.D. (2017). Interpreting age-related differences in memory-related neural activity. In: Cabeza, R., Nyberg, L., Park, D.C. (Eds.), Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging (pp. 183–203), second ed. Oxford University Press, New York.

Squire, L.R., Wixted, J.T., & Clark, R.E. (2007). Recognition memory and the medial temporal lobe: a new perspective. Nature Reviews Neuroscience, 8, 872–883.

Staresina, B.P., Fell, J., Do Lam, A.T., Axmacher, N., & Henson, R. N. (2012). Memory signals are temporally dissociated in and across human hippocampus and perirhinal cortex. Nature neuroscience, 15, 1167–1173.

Stark, C.E., & Okado, Y. (2003). Making memories without trying: medial temporal lobe activity associated with incidental memory formation during recognition. Journal of Neuroscience, 23, 6748–6753.

Thakral, P.P., Wang, T.H., & Rugg, M.D. (2015). Cortical reinstatement and the confidence and accuracy of source memory. Neuroimage, 109, 118–129.

Tsvetanov, K.A., Henson, R.N., Tyler, L.K., Davis, S.W., Shafto, M.A., Taylor, J.R., Williams, N., Cam, C., & Rowe, J.B., 2015. The effect of ageing on f MRI: correction for the confounding effects of vascular reactivity evaluated by joint fMRI and MEG in 335 adults. Human Brain Mapping. 36, 2248–2269.

Tulving, E. (1985). Memory and consciousness. Canadian Psychology, 26, 1–12.

Wang, T.H., de Chastelaine, M., Minton, B., & Rugg, M.D. (2012). Effects of age on the neural correlates of familiarity as indexed by ERPs. Journal of Cognitive Neuroscience, 24, 1055–1068.

Wang, T.H., Johnson, J.D., de Chastelaine, M., Donley, B. E., & Rugg, M. D. (2016). The Effects of Age on the Neural Correlates of Recollection Success, Recollection-Related Cortical Reinstatement, and Post-Retrieval Monitoring. Cerebral Cortex, 26, 1698–1714.

Wang, T.H., Kruggel, F., & Rugg, M.D. (2009). Effects of advanced aging on the neural correlates of successful recognition memory. Neuropsychologia, 47, 1352–1361.

Wang, W.C., Cabeza, R., 2016. Episodic memory encoding and retrieval in the aging brain. In: Cabeza, R., Nyberg, L., Park, D.C. (Eds.), Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging (pp. 301–335), second ed. Oxford University Press, New York.

Wang, W.C., & Giovanello, K.S. (2016). The role of medial temporal lobe regions in incidental and intentional retrieval of item and relational information in aging. Hippocampus, 26, 693–699.

Wang, W.C., Ranganath, C., & Yonelinas, A.P. (2014). Activity reductions in perirhinal cortex predict conceptual priming and familiarity-based recognition. Neuropsychologia, 52, 19–26.

Wechsler, D. (2001). Wechsler Test of Adult Reading. The Psychological Corporation, San Antonio, TX.

Wheeler, M.E., & Buckner, R.L. (2004). Functional-anatomic correlates of remembering and knowing. Neuroimage, 21, 1337–1349.

Wixted, J.T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. Psychological review, 117, 1025–1054.

Wixted, J.T., Mickes, L., & Squire, L.R. (2010). Measuring recollection and familiarity in the medial temporal lobe. Hippocampus, 20, 1195–1205.

Yonelinas, A.P. (2002). The nature of recollection and familiarity: A review of 30 years of research. Journal of memory and language, 46, 441–517.

Yonelinas, A.P., & Jacoby, L.L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. Journal of Memory and Language, 34, 622–643.

Yonelinas, A.P., & Jacoby, L.L. (1996). Noncriterial recollection: Familiarity as automatic, irrelevant recollection. Consciousness and Cognition, 5, 131–141.

Yonelinas, A.P., Otten, L.J., Shaw, R.N., & Rugg, M.D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. Journal of Neuroscience, 25, 3002–3008.

Yu, S.S., Johnson, J.D., & Rugg, M.D. (2012). Hippocampal activity during recognition memory co-varies with the accuracy and confidence of source memory judgments. Hippocampus, 22, 1429–1437.

Title: The effects of age on neural correlates of recognition memory: an fMRI study

Authors: Mingzhu Hou, Tracy H. Wang, and Michael D. Rugg

Supplemental Materials

1. Pairwise comparisons among confidence judgments at each set of ROIs
2. Relationships between fMRI effects and recognition performance

1. Pairwise comparisons among confidence judgments at each set of ROIs

Results for the pairwise comparisons among confidence judgments for the recollection ROIs are shown in Supplemental Table 1. As is evident from the table, among the significant findings that survived the Holm-Bonferroni correction, significantly higher BOLD activity was evident for confident old than unconfident old judgments in all ROIs with the exception of the right hippocampus. Confident old judgments also demonstrated higher BOLD activity than new judgments in all ROIs except for the left and right hippocampus. While comparable levels of activity were evident between unconfident old and unconfident new judgements across all the ROIs, in the left hippocampus, mPFC, MTG and PCC, unconfident old judgments

demonstrated significantly *lower* activity than confident new judgments. Lower BOLD activity was also evident for the unconfident than the confident new judgments in the left PHC, AG, MTG and PCC.

Supplemental Table 1. p values for pairwise comparisons among confidence judgments for each recollection ROI (bold values denote significance at the $p < .05$ level, comparisons that did not survive multiple comparison correction at each ROI are shown in italics).

	Left PHC		
	UCO	UCN	CN
CO	**CO > UCO, p = .004**	**CO > UCN, p < .001**	CO = CN, p = .260, BF_{10} =
UCO		UCO = UCN, p = 507, BF_{10} = .22	UCO = CN, p = .063, BF_{10}
UCN			**UCN < CN, p = .012**
	Left Hipp		
	UCO	UCN	CN
CO	**CO > UCO, p < .001**	CO = UCN, p = .344, BF_{10} = .27	CO = CN, p = 217, BF_{10} =
UCO		*UCO < UCN, p = .036*	**UCO < CN, p < .001**
UCN			UCN = CN, p = .085, BF_{10}
	Left mPFC		
	UCO	UCN	CN
CO	**CO > UCO, p < .001**	**CO > UCN, p = .002**	CO = CN, p = .895, BF_{10} =
UCO		UCO = UCN, p = .378, BF_{10} = .25	**UCO < CN, p = .007**
UCN			*UCN < CN, p = .019*
	Left AG		
	UCO	UCN	CN
CO	**CO > UCO, p < .001**	**CO > UCN, p < .001**	**CO > CN, p < .001**
UCO		UCO = UCN, p = .204, BF_{10} = .38	UCO = CN, p = .078, BF_{10}
UCN			**UCN < CN, p = .011**
	Left MTG		
	UCO	UCN	CN
CO	**CO > UCO, p < .001**	**CO > UCN, p < .001**	CO = CN, p = .050, BF_{10} =

UCO		UCO = UCN, p = .367, BF_{10} = .26	**UCO < CN, p = .006**
UCN			**UCN < CN, p = .001**
		Left PCC	
	UCO	UCN	CN
CO	**CO > UCO, p < .001**	**CO > UCN, p < .001**	**CO > CN, p < .001**
UCO		UCO = UCN, p = .374, BF_{10} = .25	**UCO < CN, p = .015**
UCN			**UCN < CN, *p* = 010**
		Right Hipp	
	UCO	UCN	CN
CO	**CO > UCO, p < .001**	**CO > UCN, p = .007**	CO = CN, p = .383, BF_{10} = .25
UCO		UCO = UCN, p = .276, BF_{10} = .31	**UCO < CN, p = .004**
UCN			UCN = CN, p = .071, BF_{10} = .83

Note: CO: confident old; UCO: unconfident old; UCN: unconfident new; CN: confident new.

Supplemental Table 2 shows the findings of pairwise comparisons among the confidence judgments conducted for the non-PRC familiarity ROIs. As is evident from the table, each ROI demonstrated a similar pattern of decreasing activity from confident old to unconfident new judgments. After correction for multiple comparisons, BOLD activity remained significantly higher for confident old than all the other judgments in all the ROIs except for left and right caudate. Decreased activity was evident from unconfident old to unconfident new judgments across all regions. With the exceptions of the left and right caudate, unconfident old judgments also demonstrated higher activity than confident new judgments in all the ROIs. In the left IPS and Precuneus, BOLD activity was higher for unconfident new than confident new judgements.

Supplemental Table 2. p values for pairwise comparisons among confidence judgments for each non-PRC familiarity ROI (bold values denote significance at the p < .05 level,

comparisons that did not survive multiple comparison correction at each ROI are shown in italics).

	Left Caudate		
	UCO	UCN	CN
CO	*CO > UCO, p = .029*	**CO > UCN, p < .001**	**CO > CN, p < .001**
UCO		**UCO > UCN, p = .003**	UCO = CN, p = .054, BF_{10} =
UCN			UCN = CN, p = .069, BF_{10} =
	Left IPS		
	UCO	UCN	CN
CO	**CO > UCO, p = .007**	**CO > UCN, p < .001**	**CO > CN, p < .001**
UCO		**UCO > UCN, p = .002**	**UCO > CN, p < .001**
UCN			**UCN > CN, p = .013**
	Left alPFC		
	UCO	UCN	CN
CO	**CO > UCO, p < .001**	**CO > UCN, p < .001**	**CO > CN, p < .001**
UCO		**UCO > UCN, p = .002**	**UCO > CN, p = .002**
UCN			UCN = CN, p = .829, BF_{10} =
	Left Precuneus		
	UCO	UCN	CN
CO	**CO > UCO, p < .001**	**CO > UCN, p < .001**	**CO > CN, p < .001**
UCO		**UCO > UCN, p = .010**	**UCO > CN, p < .001**
UCN			**UCN > CN, p = .009**
	Left dmPFC		
	UCO	UCN	CN
CO	**CO > UCO, p = .008**	**CO > UCN, p < .001**	**CO > CN, p < .001**
UCO		**UCO > UCN, p = .002**	**UCO > CN, p < .001**
UCN			UCN = CN, p = .755, BF_{10} =
	Right caudate		
	UCO	UCN	CN

CO	*CO > UCO, p = .018*	**CO > UCN, p < .001**	**CO > CN, p < .001**
UCO		**UCO > UCN, p = .007**	UCO = CN, p = .270, BF_{10} = .31
UCN			UCN = CN, p = .085, BF_{10} = .72

Note: CO: confident old; UCO: unconfident old; UCN: unconfident new; CN: confident new.

Analogous comparisons were also conducted for the mean BOLD activity averaged across left and right PRC (given that there were no significant interactions involving hemisphere in the ANOVA, see the first panel of Table 8 in the main text). As is evident from Supplemental Table 3, confident old judgments demonstrated significantly higher BOLD activity than unconfident judgments. By contrast, lower BOLD activity was evident for unconfident old judgments compared to confident new judgments.

Supplemental Table 3. p values for pairwise comparisons among confidence judgments for left and right PRC (bold values denote significance at the p < .05 level, comparisons that did not survive multiple comparison correction are shown in italics).

	Left and right PRC		
	UCO	UCN	CN
CO	**CO > UCO, p < .001**	**CO > UCN, p = .004**	CO = CN, p = .465, BF_{10} = .2
UCO		UCO = UCN, p = .550, BF_{10} = .21	**UCO < CN, p = .007**
UCN			*UCN < CN, p = .024*

Note: CO: confident old; UCO: unconfident old; UCN: unconfident new; CN: confident new.

2. Relationships between fMRI effects and recognition performance

Regression analyses were conducted to examine the relationships between the functional effects and memory recognition performance, and whether they varied with age group. In the model examining the relationships between the recollection effects in the core recollection ROIs (see Methods) and pR, the recollection effect x age group interaction failed to achieve significance in all ROIs except for left MPFC (for left mPFC, β = .54, partial r = .39, p = .020, for other ROIs, absolute βs < .28, absolute partial rs < .16, ps > .364, BF_{incl} < .47). The finding for the mPFC did not survive Holm-Bonferroni correction, however. After the removal of the interaction term, none of the recollection effects in the core recollection regions significantly predicted pR, BF_{incl}s < .85 (see Supplemental Table 4). Similarly, fMRI

recollection effects x age group interactions were not predictive of either pF or mean recognition accuracy (absolute βs < .40, absolute partial rs < .23, ps > .185, BF_{incl} < 1.57). As is also evident from Supplemental Table 4, in two further sets of analyses, recollection effects did not significantly predict either pF or mean accuracy.

Supplemental Table 4. Results of regression models examining the relationships between recollection effects and memory performance, with age group as a covariate.

ROI	β	partial r	p	BF_{incl}
Predicting pR				
Left_PHC	.11	.11	.519	.37
Left_Hipp	.21	.21	.213	.51
Left_mPFC	-.18	-.18	.286	.45
Left_AG	-.27	-.27	.102	.84
Left_MTG	-.08	-.08	.647	.33
Left_PCC	-.07	-.07	.688	.32
Right_Hipp	-.10	-.10	.556	.37
Predicting pF				
Left_PHC	.11	.11	.513	.42
Left_Hipp	.17	.17	.304	.48
Left_mPFC	.18	.18	.284	.58
Left_AG	.23	.24	.153	.79
Left_MTG	.15	.15	.370	.46
Left_PCC	.27	.27	.106	1.07
Right_Hipp	.09	.09	.579	.37
Predicting memory accuracy				
Left_PHC	.28	.30	.073	1.83
Left_Hipp	.21	.22	.189	.89
Left_mPFC	.21	.22	.186	1.02
Left_AG	.08	.08	.631	.52
Left_MTG	.16	.18	.300	.72
Left_PCC	.20	.21	.206	.94

Right_Hipp	.12	.13	.447	.58

In the regression models examining relationships between the familiarity-enhancement effects and memory recognition performance, the age group x functional effect did not significantly predict pR, pF or memory accuracy (absolute βs < .70, absolute partial rs < .33, ps > .052, BF_{incl}s < .96) in any case. As is evident in Supplemental Table 5, in the follow-up models, none of familiarity-enhancement effects significantly predicted any measure of memory performance.

Supplemental Table 5. Results of regression models examining the relationships between familiarity-enhancement effects in the non-PRC ROIs and memory performance, with age group as a covariate.

ROI	β	partial r	p	BF_{incl}
Predicting pR				
Left_Caudate	-.01	-.01	.949	.31
Left_IPS	.08	.08	.633	.33
Left_alPFC	.19	.19	.262	.48
Left_Precuneus	-.04	-.04	.797	.32
Left_dmPFC	.23	.24	.160	.65
Right_Caudate	.05	.05	.761	.32
Predicting pF				
Left_Caudate	.01	.01	.959	.34
Left_IPS	.22	.22	.190	.62
Left_alPFC	.14	.14	.409	.43
Left_Precuneus	.07	.07	.677	.36
Left_dmPFC	.19	.20	.246	.57
Right_Caudate	-.14	-.14	.403	.45
Predicting memory accuracy				
Left_Caudatae	.01	.01	.940	.47
Left_IPS	.14	.15	.370	.63
Left_alPFC	.14	.15	.384	.62

Left_Precuneus	.02	.02	.924	.47
Left_dmPFC	.18	.19	.263	.77
Right_Caudate	-.02	-.02	.909	.47

In an analogous set of regression models examining the relationship between the familiarity-attenuation effects in PRC and memory performance, we again failed to identify significant functional effect x age group interaction terms (absolute βs < .62, absolute rs < .27, ps > .122, BF_{incl}s < 1.92). As is evident from Supplemental Table 6, after the removal of the interaction term, familiarity-attenuation effects in both left and right PRC were predictive of pF and memory accuracy.

Supplemental Table 6. Results of regression models examining the relationships between familiarity-attenuation effects in PRC and memory performance, with age group as a covariate (bold values denote significance at the p < .05 level).

ROI	β	partial r	p	BF_{incl}
Predicting pR				
Left PRC	.21	.21	.221	.61
Right PRC	-.01	-.01	.947	.31
Predicting pF				
Left PRC	**.51**	**.52**	**.001**	
Right PRC	**.41**	**.42**	**.010**	
Predicting memory accuracy				
Left PRC	**.51**	**.54**	**.001**	
Right PRC	**.37**	**.40**	**.015**	