Title Page

Memory deficits following hippocampal versus parahippocampal damage double-dissociate according to both process and material type

Abbreviated title: Dissociating recollection and familiarity

Author Names

Georgios P. D. Argyropoulos PhD [1]

Carola Dell'Acqua MSc [1,2]

Emily Butler PhD [1]

Clare Loane PhD [1,3]

Adriana Roca-Fernandez MSc [1]

Azhaar Almozel BSc [1,4]

Nikolas Drummond MSc [1,5]

Carmen Lage-Martinez MD [1,6]

Elisa Cooper PhD [7]

Richard N. Henson PhD [7]

Christopher R. Butler FRCP, PhD [1,8,9]

Affiliations and Postal codes

[1] Memory Research Group, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK; Level 6, West Wing, John Radcliffe Hospital, Oxford, OX3 9DU, UK.

[2] Department of General Psychology & Padova Neuroscience Center, University of Padova, Padova, Italy

[3] Maurice Wohl Clinical Neuroscience Institute, Basic and Clinical Neuroscience Department, King's College London, London, UK

[4] Cardiff University, School of Biosciences, Cardiff, UK

[5] Department of Zoology, University of Cambridge, Cambridge, UK

[6] Valdecilla Biomedical Research Institute, University Hospital Marqués de Valdecilla.

[7] MRC Cognition & Brain Sciences Unit and Department of Psychiatry, University of Cambridge, UK

[8] Department of Brain Sciences, Imperial College London, London, UK.

[9]Departamento de Neurología, Pontificia Universidad Católica de Chile, Santiago, Chile.

Corresponding author

Georgios P. D. Argyropoulos; georgios.argyropoulos@ndcn.ox.ac.uk; Medical Sciences Division, Nuffield Department of Clinical Neurosciences, Level 6, West Wing, John Radcliffe Hospital, Oxford, OX3 9DU, UK.

Abstract

A central debate in neuroscience concerns whether different medial temporal lobe structures support different memory processes. Using three recognition memory paradigms for faces, scenes or words, we tested eight patients with hippocampal and varying parahippocampal damage, plus an extremely rare patient with focal perirhinal damage. Recollection impairment was uniquely related to hippocampal damage, regardless of material, whereas familiarity impairment was material-specific depending upon the damage locus within the parahippocampal gyrus.

Main

Ever since the first descriptions of the famous patient HM, individuals with medial temporal lobe (MTL) damage have been fundamental in delineating the brain regions supporting human memory. Patient studies offer crucial insights into causal brain-behavioural relationships, beyond the correlational information afforded by functional imaging in healthy participants. Nevertheless, many questions about the neural basis of amnesia remain unresolved[1].

In particular, competing accounts have been offered to explain the impact of MTL damage on recognition memory, i.e. the capacity to discriminate between previously encountered versus novel stimuli. A central question relates to 'process-specificity' - whether distinct MTL structures support different processes underlying recognition memory, such as "recollection" (remembering the context in which a stimulus occurred; fundamental for free recall) versus "familiarity" (a feeling that a stimulus was encountered, without retrieval of contextual information). According to one such 'dual-process' framework[2], recollection relies on the hippocampus (HPC), whereas the perirhinal (PRC) and parahippocampal cortices (PHC) in the parahippocampal gyrus, the two major sources of input to the HPC, support familiarity processes. This framework therefore predicts a double dissociation between memory processes following selective HPC vs. parahippocampal lesions. One opposing 'single-process theory'[3], however, posits that recollection and familiarity reflect subjective expressions of memory traces of varying strength, i.e. familiarity results when weak memory traces fail to re-activate the associated contextual information that characterises recollection. This theory further proposes that MTL regions function as a single, integrated memory system, such that MTL damage affects both recollection and familiarity. If anything, smaller lesions (e.g. those

selective to the PRC) should impair familiarity less than recollection, if stronger traces require a fully integrated MTL system[2].

Studies assessing these competing predictions face several challenges. Firstly, there is no universally accepted method of separating recollection and familiarity estimates, so multiple methods are recommended. Secondly, selective HPC damage is rare[4], and the term 'selective' or 'focal' is often used on the basis of visual ratings of structural MRI (as noted elsewhere[5]), which does not allow for brain-behavioural correlations. Indeed, conditions associated with HPC amnesia, e.g. ischemia/anoxia, often also cause extra-MTL damage[6]. Thirdly, selective PRC lesions are even rarer still; only two cases (patients NB[7] and IR[8]) have been reported. Across several experiments, NB showed deficits in familiarity but not recollection, supporting dual-process accounts. NB had damage in the left PRC, although it extended to the amygdala, entorhinal and anterolateral temporal cortices. Less is known regarding IR, who had right PRC damage and showed perceptual, in the absence of memory deficits. However, no MTL volumetry was reported, and the tasks were not designed to assess familiarity and recollection separately. Overall, given the potential confounds of case studies (impairment severity, measurement noise, individual differences[9]), more cases with focal PRC lesions are required. Importantly, direct comparison with HPC patients is needed to provide a full double-dissociation, where all patients are tested on the same paradigms.

A second question about the role of MTL in memory concerns material-specificity[10]. For example, it has been suggested that PRC is important for recognising objects and faces, while PHC is important for scene recognition[2]. The HPC has also been claimed to be important for processing scenes[11], though others have claimed that its role in memory is independent of material[12]. Since material-specific accounts may provide an alternative to

dual- and single-process frameworks in explaining memory deficits following MTL damage[13], and that previous dissociations between processes (e.g. recollection vs familiarity) could be confounded with material (e.g. faces vs scenes), it is important to examine recognition across different material types.

Here we present results from eight patients who suffered autoimmune limbic encephalitis. They all had HPC damage and varying degrees of parahippocampal damage. Only one of those 8 patients had damage in PRC ('HPC+PRC+'; rest of patients: 'HPC+PRC-'). We also present an exceptionally rare case ('HPC-PRC+') with a focal lesion in the right PRC (Figure 1). On standard neuropsychometry, all patients showed preserved semantic memory, language, visuospatial, motor, and executive function. Both PRC+ patients, but none of the 7 HPC+PRC- patients, showed impaired face recognition memory (Supplementary Table 1). Regarding HPC-PRC+, this was the only test in which he showed impaired performance.

None of our patients showed evidence of extra-MTL damage, acutely or post-acutely (see Methods). All patients, together with n=15 healthy controls ('HPC-PRC-'), were tested on three, more theoretically-focused memory paradigms. Each paradigm was designed to isolate recollection and familiarity in a different way, with the potential to provide convergent evidence. Furthermore, two of the paradigms were repeated on three material types, both verbal and nonverbal (words, unfamiliar faces and scenes). We could therefore test whether patients' memory performance dissociated according to process or material type.

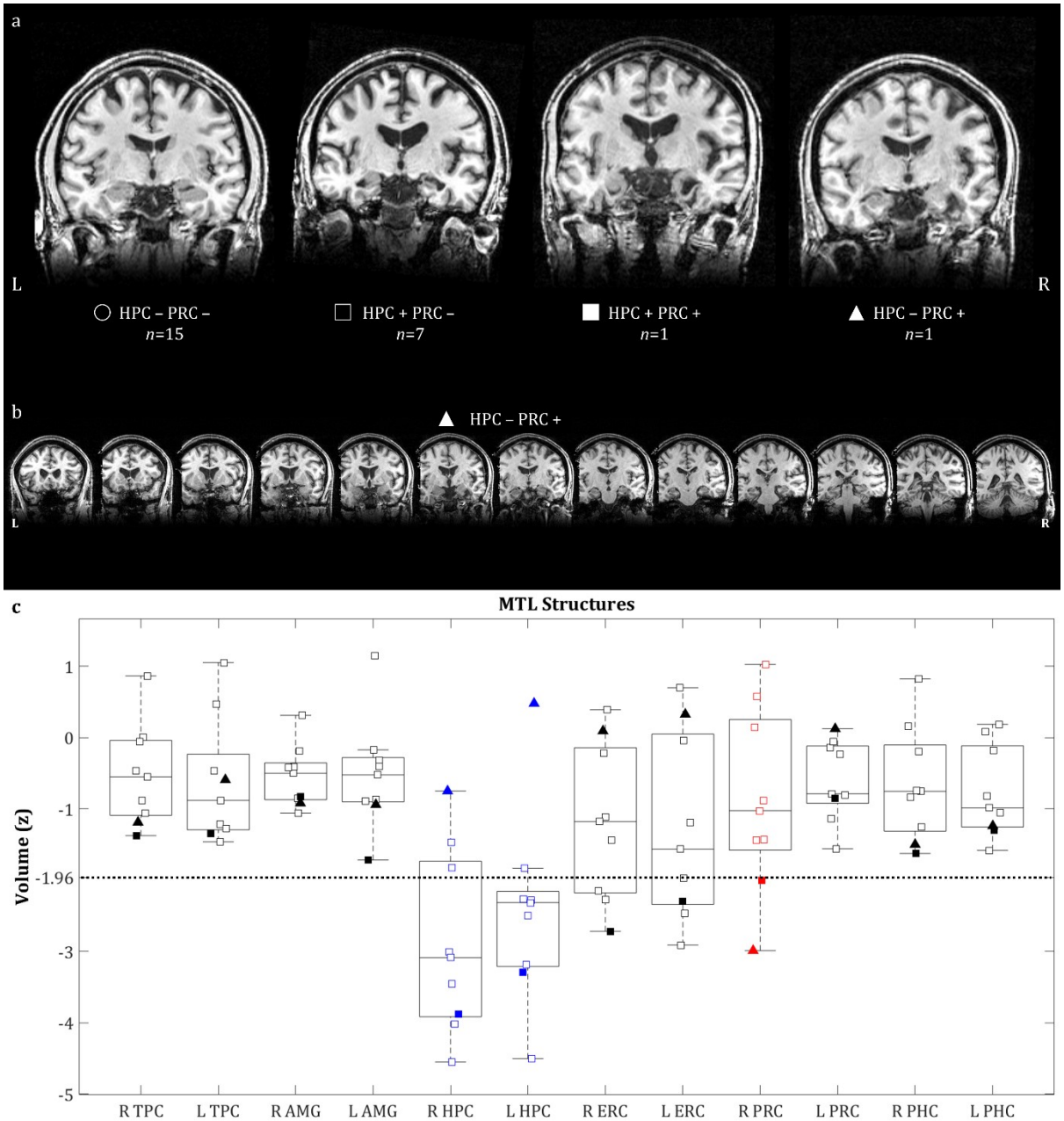

Figure 1. a: Coronal slices of structural MR images of one participant representative of each of the four groups of interest: HPC-PRC- : healthy controls (n=15); HPC+PRC-: MTL patients with HPC atrophy extending in various degrees to other MTL structures but not the PRC (n=7); HPC+PRC+: single MTL patient with bilateral HPC and R PRC atrophy (n=1); HPC-PRC+ : single patient with focal right PRC damage; **b:** A series of coronal slices for HPC-PRC+, highlighting his lesioned right PRC and his spared left PRC, along with the rest of his spared MTL structures; **c:** volumes of L / R MTL structures for the 9 patients; z: volumes are expressed as z-scores, based on the mean and standard deviation of the volumes of the 48 healthy controls whose MTL structures were manually delineated (see [14] for details); atrophy/damage defined as a z<-1.96 relative to 48 age-matched healthy controls, reflecting significant volume reduction ($p<0.05$) for the manually delineated volume of the corresponding structure (dotted line). The line within each boxplot indicates the median value; bottom of box: 25th %ile; top of box: 75th %ile; upper and lower whiskers: scores outside the middle 50%. Patients generally have reduced HPC volume (blue squares; patient HPC+PRC-1 showed marginal atrophy: L HPC(z)=-1.83) relative to controls, except for HPC-PRC+ (filled triangle). Patients have normal (right) PRC volume (highlighted in red) except for HPC-PRC+ and HPC+PRC+. **Key:** L, R: left, right hemisphere; TPC: temporopolar cortex; AMG: amygdala; HPC: hippocampus; ERC: entorhinal cortex; PRC: perirhinal cortex; PHC: parahippocampal cortex.

Paradigm 1 was based on receiver operating characteristics (ROC) derived from confidence ratings on recognition memory (Supplementary Figure 1). Paradigm 2 was based on a response deadline paradigm (RDP) and predicated on the selective reliance of recognition memory on familiarity at short response deadlines, in contrast with long response deadlines (Supplementary Figure 2). Paradigms 1 and 2 were repeated for unfamiliar human faces, unfamiliar natural scenes with high feature overlap, and visually presented, high-frequency words. Paradigm 3 was a (continuous) "source" recognition paradigm (SRP), in which participants saw objects superimposed on scenes, and assessed whether they had seen the object before (familiarity) and in which left-right position on the scene it had occurred previously (recollection)[15] (Supplementary Figure 3).

In order to address the question of process-specificity, we started with composite measures of recollection and familiarity, collapsed across material types and paradigms (figure 2a). Relative to controls (HPC-PRC-), all the patients tended to have impaired recollection, except for the HPC-PRC+ case, consistent with reports on patient NB[7]. Indeed, the amount of (spared) HPC volume in patients correlated with the degree of recollection (figure 2b), but not with the degree of familiarity (figure 2c). This impairment of recollection in HPC+ patients and selective relationship of HPC volumes to recollection was replicated separately for each paradigm (figure 2e,h,k) and material type (figure 3a-c).

While the selective role of HPC in recollection was clear, the case of familiarity was less so. The two PRC+ cases (HPC+PRC+ and HPC-PRC+) showed impaired familiarity relative to healthy controls (figure 2a), supporting a role for PRC in familiarity. However, this was driven by the differences in Paradigm 3 (figure 2j). Since familiarity in Paradigm 3 pertained exclusively to objects, whereas Paradigms 1 and 2 involved different material

types, we next combined data across Paradigms 1 and 2 but split by material-type. We examined the relationship of volume reduction in different structures in the parahippocampal gyrus (ERC, PRC, PHC) with familiarity/recollection for different material types (faces, scenes, words).

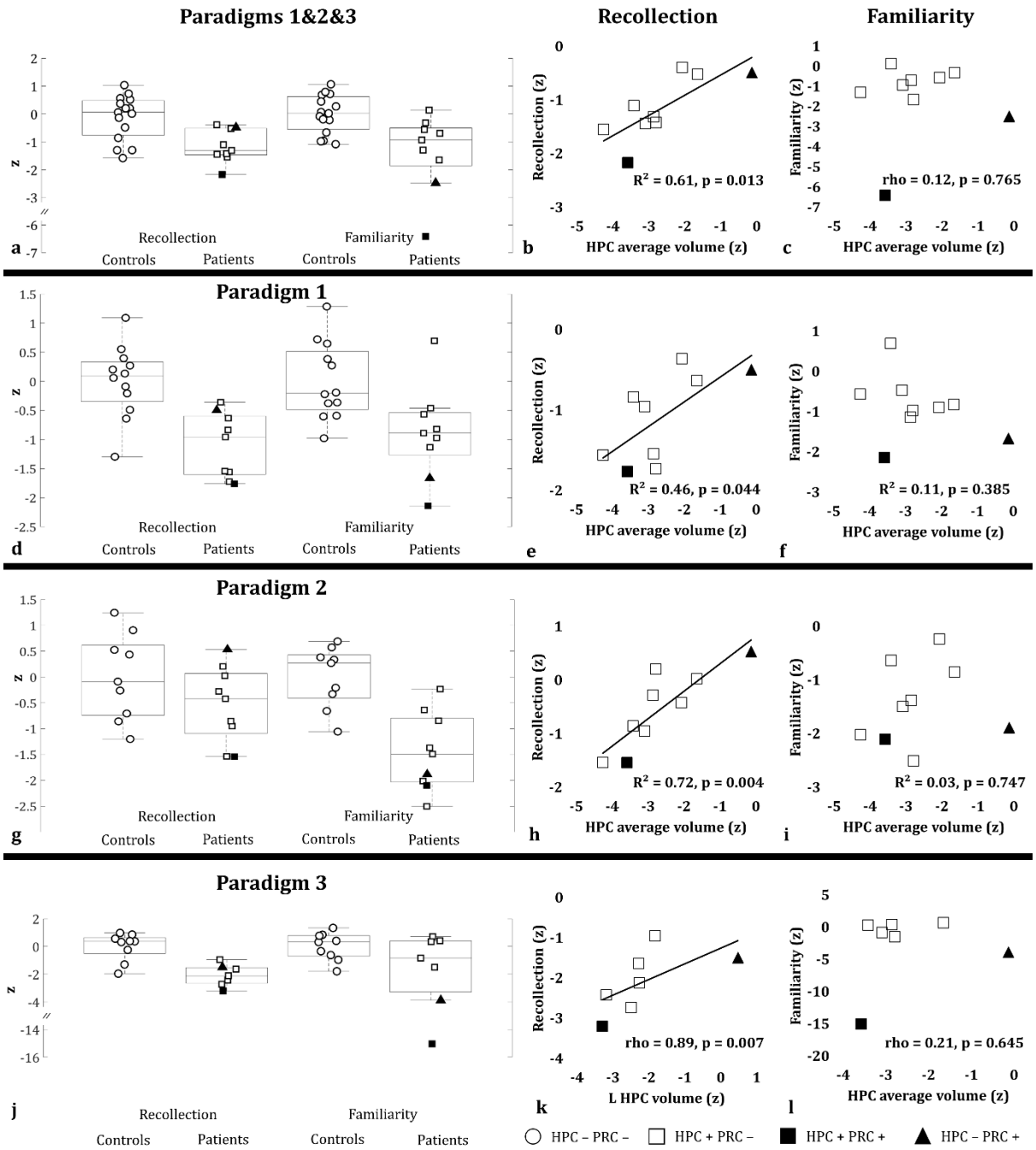

Figure 2: familiarity and recollection estimates for healthy controls (HPC-PRC-) and patients collapsing across material types; **a:** collapsing across paradigms; Both PRC+ patients showed familiarity impairment relative to healthy controls (HPC-PRC+: t=3.48, p=0.004; HPC+PRC+: t=9.04, p=0000003), which was more pronounced than that of HPC+PRC- patients (HPC-PRC+: t=2.69, p=0.036; HPC+PRC+: t=8.86, p=0.0001). HPC-PRC+ showed preserved recollection (vs HPC-PRC-: t=0.43, p=0.675); HPC+PRC+ showed impaired recollection (t=2.46, p=0.028); HPC+PRC- patients showed impaired recollection (vs HPC-PRC-: t=2.97, p=0.008, d=1.49) and familiarity (t=2.48, p=0.022; d=1.17; two-way ANOVA: Group*Process: F=0.23, p=0.635, η^2_p=0.01); **b:** recollection correlated with HPC (L HPC: rho=0.80, p=0.010; R HPC: r=0.78, p=0.014), but not with L/R PHC/PRC/ERC volumes (all rs, |r|≤0.47, p≥0.203); **c:** familiarity did not correlate with L/R HPC volumes (all rhos, |rho|≤0.350, p≥0.356); **d,g,j:** HPC-PRC+ and HPC+PRC+ showed impaired familiarity in each paradigm. HPC+PRC+ showed impaired recollection (Paradigm 1: t=2.75, p=0.019; Paradigm 2: t=1.75, p=0.118; Paradigm 3: t=3.03, p=0.016; HPC-PRC+: all ts, |t|≤1.41; all ps, p≥0.195); **d:** Paradigm 1: Relative to healthy controls, HPC-PRC+ (t=2.40, p=0.035) and HPC+PRC+ showed impaired familiarity (t=3.09, p=0.010), but did not differ from HPC+PRC- patients (HPC-PRC+: t=1.63, p=0.153; HPC+PRC+: t=2.36, p=0.056); familiarity was not preserved in HPC+PRC- patients relative to recollection (Group*Process: F=1.16, p=0.296, η^2_p=0.06; Group: F=21.08, p=0.0003, η^2_p=0.55); **g:** Paradigm 2: Relative to healthy controls, HPC-PRC+ (t=3.00, p=0.017) and HPC+PRC+ showed impaired familiarity (t=3.35, p=0.010), but did not differ from HPC+PRC- patients (HPC-PRC+: t=0.69, p=0.517; HPC+PRC+: t=0.94, p=0.382); familiarity was not preserved in HPC+PRC- patients relative to recollection (Group*Process: F=1.87, p=0.193, η^2_p=0.12; Group: F=15.55, p=0.001, η^2_p=0.53); **j:** Paradigm 3: Relative to healthy controls, HPC-PRC+ (t=3.68, p=0.006) and HPC+PRC+ showed impaired familiarity (t=14.25, p=0.0000006) and larger impairment than HPC+PRC- patients (HPC-PRC+: t=3.60, p=0.023; HPC+PRC+: t=14.38, p=0.0001); familiarity was selectively preserved in HPC+PRC- patients (vs. HPC-PRC-: recollection: t=3.87, p=0.002, d=2.28; familiarity: t=0.30, p=0.771, d=0.17; two-way ANOVA: Group*Process: F=8.31, p=0.014; η^2_p=0.41); line in boxplots=median; bottom of box=25th %ile; top=75th %ile; whiskers=scores outside the middle 50%; **e,h,k:** the relationship of recollection with HPC volumes in (b) was replicated for each paradigm; **f,i,l:** no relationship between L/R HPC volumes and familiarity in any paradigm (all rhos/rs, |rho/r| ≤0.47, p≥0.205); **key:** HPC: hippocampus; L, R: left/right hemisphere; ERC: entorhinal cortex; PHC: parahippocampal cortex; PRC: perirhinal cortex.

Collapsing across material types, familiarity selectively correlated with the volume of the parahippocampal gyrus and not the HPC (figure 3a,e). Consistent with their preserved object familiarity in Paradigm 3, HPC+PRC- patients showed spared face familiarity in Paradigms 1+2 (figure 3f). By contrast, both PRC+ patients showed impaired face familiarity, consistent with their impaired face recognition memory in neuropsychological assessment (Supplementary Table 1). This pattern dovetails with meta-analytical findings[4] on HPC patients' performance in neuropsychological tests of recognition memory for faces, as well as with our previous findings on a larger cohort of HPC patients, who showed group-level sparing of face recognition.[14] It also supports the idea that the PRC is engaged in processing faces and objects (see [10] for discussion).

For scene familiarity (figure 3g), HPC+PRC- patients were marginally impaired, and PRC+ patients did not differ from HPC+PRC- patients. Scene familiarity correlated with PHC volume across patients, whereas scene recollection correlated with HPC volume. Likewise, for word familiarity (figure 3h), HPC+PRC- patients were impaired, and PRC+ patients did not differ from the HPC+PRC- patients (figure 3; Supplementary Figures 4,6). However in this case, familiarity correlated with entorhinal (ERC) volume rather than PHC volume, consistent with findings from fMRI studies on healthy young adults[16], structural MRI studies on healthy elderly adults[17], as well as with single-case studies on ERC damage (patient MR[18]). These relationships were replicated separately in Paradigms 1 and 2 (Supplementary figures 5+6). Thus the pattern that emerged was that familiarity deficits show material-specificity, likely because the deficit depends on the MTL structure that processes a certain material-type, in accordance with the differential connectivity of the PRC and PHC with the ventral and dorsal visual processing streams, respectively[19]. By contrast, recollection correlates with HPC volume regardless of material-type, perhaps since recollection requires binding any type of material to a context representation[20].

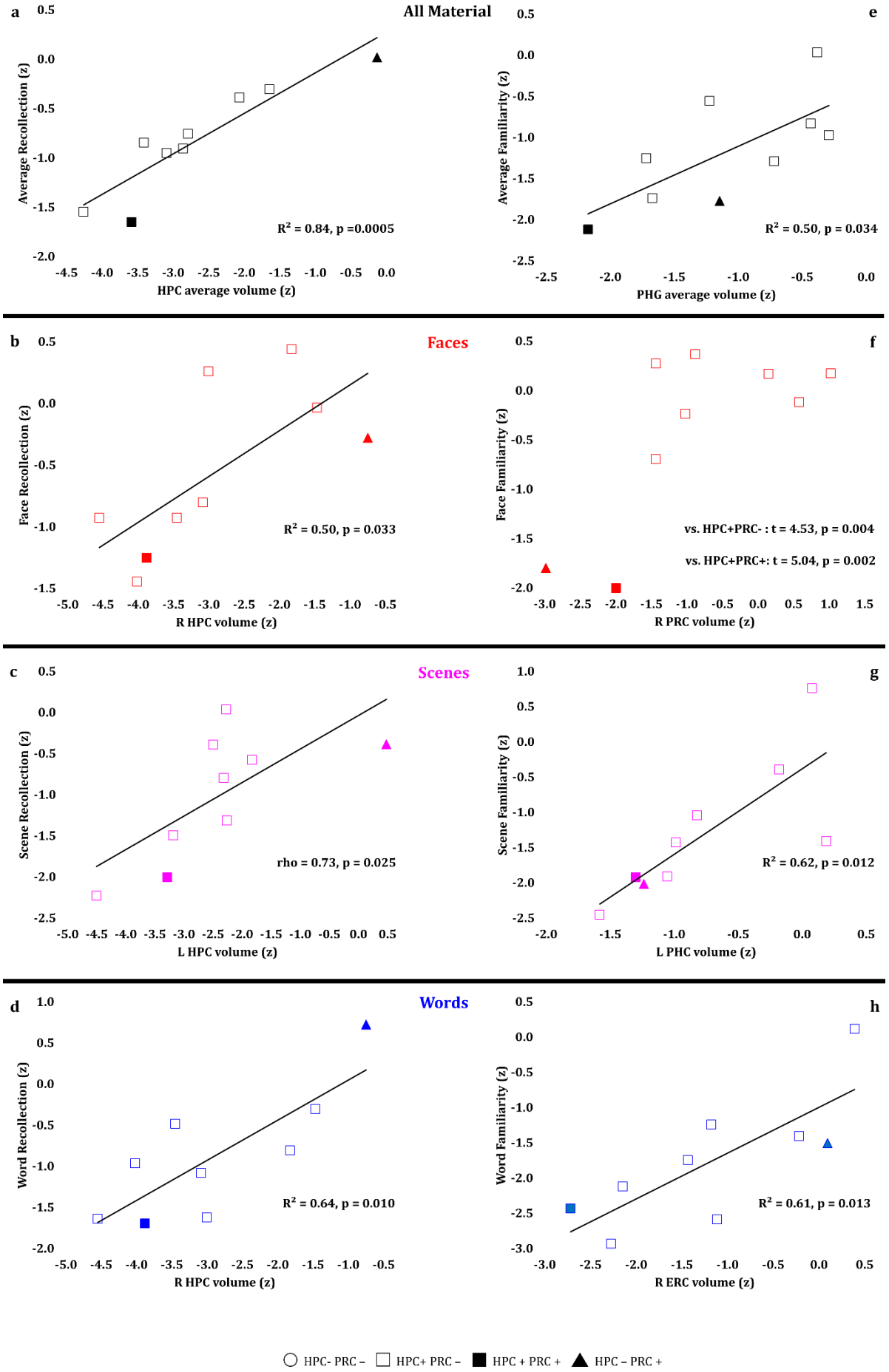
a
All Material
Average Recollection (z)
HPC average volume (z)
R² = 0.84, p =0.0005
e
Average Familiarity (z)
PHG average volume (z)
R² = 0.50, p = 0.034
b
Faces
Face Recollection (z)
R HPC volume (z)
R² = 0.50, p = 0.033
f
Face Familiarity (z)
R PRC volume (z)
vs. HPC+PRC- : t = 4.53, p = 0.004
vs. HPC+PRC+: t = 5.04, p = 0.002
c
Scenes
Scene Recollection (z)
L HPC volume (z)
rho = 0.73, p = 0.025
g
Scene Familiarity (z)
L PHC volume (z)
R² = 0.62, p = 0.012
d
Words
Word Recollection (z)
R HPC volume (z)
R² = 0.64, p = 0.010
h
Word Familiarity (z)
R ERC volume (z)
R² = 0.61, p = 0.013
HPC- PRC –
HPC+ PRC –
HPC + PRC +
HPC – PRC +

Figure 3. patients' z-transformed recollection (a-d) and familiarity (e-h) estimates for the different material types collapsing across Paradigms 1 and 2; **a,e:** collapsing across material types and the two Paradigms, recollection selectively correlated with HPC volume (PHG volume: r=0.26, p=0.503; HPC vs PHG volume: z=2.13, p=0.017), whereas familiarity selectively correlated with PHG volume (HPC volume: r=-0.10, p=0.799; PHG vs HPC volume: z=1.68, p=0.047); **b-d:** across material types, recollection estimates correlated with HPC volumes, but not with L/R ERC/PRC/PHC volumes (all rs, |r|≤ 0.48; ps, p ≥0.194); **f:** face familiarity - both PRC+ patients underperformed relative to HPC+PRC- patients, none of whom showed impairment (z > -1); **g:** scene familiarity: the two PRC+ patients did not differ from the HPC+PRC- patients (both ts, t ≤0.79 ; ps, p ≥0.459), whereas the HPC+PRC- patients showed impaired performance relative to healthy controls (Supplementary Figure 4); scene familiarity correlated with L PHC volumes (R PHC: r = 0.58, p = 0.101), but not with L/R HPC (both rs/rhos, |r/rho| ≤0.18; ps, p ≥0.637) or with L/R ERC/PRC volumes (all rs, |r| ≤0.31; ps, p ≥0.411); **h:** the two PRC+ patients did not differ from the HPC+PRC- patients in word familiarity (both ts, t≤0.68; ps, p ≥0.522), whereas the HPC+PRC- patients showed impairment relative to healthy controls (Supplementary Figure 4); word familiarity correlated with R ERC volumes, but not with L/R HPC (all rs/rhos, |r/rho| ≤ 0.55; ps, p ≥ 0.125) or L ERC or L/R PHC/PRC volumes (all rs, |r| ≤ 0.63; all ps, p ≥ 0.068); familiarity/recollection estimates are expressed as z-scores based on the mean and SD of healthy controls; these relationships were seen in both Paradigm 1 and 2 (Supplementary Figures 5,6); **key:** HPC: hippocampus; L, R: left/right hemisphere; ERC: entorhinal cortex; PHC: parahippocampal cortex; PHG: parahippocampal gyrus (sum of ERC, PRC, and PHC volumes); PRC: perirhinal cortex

Overall, our data provide the most compelling support yet for dual-process models of recognition memory, in which recollection and familiarity depend on different brain regions. But they go further by supporting material-specific models to the extent that, while recollection always depends on HPC, familiarity depends on distinct regions within the parahippocampal gyrus as a function of material type.

Methods

Participants

4 groups of participants were included in the study: patient HPC-PRC+, a markedly rare case of focal PRC damage and spared HPC (n=1); patients with HPC atrophy not extending to the PRC (n=7; HPC+PRC-); a patient with HPC atrophy extending to the PRC (n=1; HPC+PRC+); healthy controls (HPC-PRC-; n=15).

Patients

Those 9 patients (7M:2F; age at behavioural assessment: mean = 60.40; SD = 6.26 years; education: mean = 12.22, SD = 1.09 years) were recruited within the context of the Memory and Amnesia Project (https://www.ndcn.ox.ac.uk/research/memory-research-group/projects-1/memory-and-amnesia-project).

HPC+PRC- and HPC+PRC+ patients

8/9 patients showed HPC atrophy and volume reduction extending to various degrees in the parahippocampal gyrus, due to autoimmune limbic encephalitis, which was diagnosed according to consensus criteria [21]. These patients were representative of the clinical and neuropsychological group-level profile of the autoimmune limbic encephalitis cohort presented in [14]: i) they were all native speakers of English; ii) they were all recruited in the post-acute stable chronic phase of the disease (delay from symptom onset range: 1.77-14.92 years); iii) in their acute clinical T2-weighted MRI scans, all 8/8 patients showed abnormalities in the HPC with respect to volume, T2 signal intensity, and/or diffusion; in the HPC+PRC- patients, these abnormalities did not extend to the parahippocampal gyrus (in one patient, there was also high T2-signal and swelling noted in the right AMG), whereas in the case of the HPC+PRC+ patient, abnormalities had

been also noted in both the HPC and the parahippocampal gyrus, and he was the only HPC patient whose atrophy extended to the (right) PRC; iv) No abnormalities were detected beyond the MTL in the research scan that patients underwent post-acutely (delay from symptom onset range: 1.72-12.93 years; see Brain Imaging section below); v) in their post-acute neuropsychological assessment (delay from symptom onset range: 1.69-12.93 years), they all showed average to above-average premorbid intelligence [National Adult Reading Test [22]], along with vi) preserved (post-morbid) intelligence, semantic memory and language [Wechsler Abbreviated Scales of Intelligence: Vocabulary, Similarities, Matrices [23]; Graded Naming Test [24]; Camel and Cactus test [25]]; vii) executive function [Delis-Kaplan Executive Function System – Trails: Number-Letter Switching [26]] including working memory [Wechsler Memory Scale: Digit Span forward and backward [27]] (individual impairment on a test was defined as an age-scaled standardised score of ≤−1.67, corresponding to the 5th %ile, in line with standard neuropsychological practice [e.g. [28]]), and viii) visuospatial perception [Visual Object and Space Perception battery: cube analysis, dot counting, position discrimination [29]] (all scores above the 5% cut-off point employed in this test); ix) however, 7/8 of the patients showed impaired performance in at least one test of anterograde memory [Wechsler Memory Scale [30]; Rey-Osterrieth Complex Figure Test [31]; the Warrington Recognition Memory Tests for faces and words [32] and the Warrington Topographical Memory test for scenes [33]; the Doors and People test [34]], (Supplementary Table 1); x) None of the patients had a history of pre-morbid psychiatric or neurological disorder that could have resulted in cognitive impairment. None had any contraindication to MRI at the time of entry into the study.

Patient HPC-PRC+

HPC-PRC+ was a 51-year-old (at the time of study participation) male with 12 years of education. At the age of 21, while working, he collapsed on the floor, and was hospitalised, where a clinical MRI showed a cerebral abscess in his right PRC, sparing the HPC and other MTL structures. The lesion is illustrated in figure 1 as a hypointensity in the structural T1-weighted MRI that he underwent at the age of 48 as part of our research study. A few years after the incident, he was diagnosed with epilepsy, and an EEG disclosed slow activity in the right anterior temporal region. He was treated with carbamazepine and the seizures remitted completely. Neuropsychological assessment (conducted at the age of 48) demonstrated normal levels of intelligence, language, executive function, visuospatial perception, visual and verbal recall, as well as verbal and visual recognition memory (all test scores: $z > -1.67$), with the striking exception of recognition memory for faces ($z = -2.33$) (Supplementary Table 1).

All 9/9 patients participated in Paradigms 1 and 2, and 7/9 patients participated in Paradigm 3 due to scheduling conflicts.

Healthy Controls

15 healthy controls (8M:7F; age at behavioural assessment: mean = 62.11; SD = 6.02 years; education: mean = 13.13, SD =1.77 years) matched for age, sex and years of education with the patient group were recruited through local advertisement (patients vs controls: age: $t = 0.66$, $p = 0.515$; M:F ratio: $\chi^2 = 1.43$, $p = 0.231$; education: $t = 1.39$, $p = 0.179$). They were all native speakers of English, with no known psychiatric or neurological disorders.

Due to scheduling conflicts and technical errors, 12/15 healthy control datasets were available for Paradigm 1; 9/15 for Paradigm 2, and 9/15 for Paradigm 3.

All participants provided written informed consent according to the Declaration of Helsinki. Ethical approval was received from South Central Oxford Research Ethics Committee (REC no: 08/H0606/133).

Brain Imaging

Scanning procedures

We acquired 3D T1-weighted MRIs for all 9 patients (Siemens 3T Trio system; 32-channel head coil; University of Oxford Centre for Clinical Magnetic Resonance Research) using a Magnetisation Prepared Rapid Gradient Echo (MPRAGE) sequence (echo time = 4.7ms; repetition time = 2040ms; flip angle = 8°; field of view = 192mm; voxel size = 1 × 1 × 1mm) for all patients.

Manual volumetry

Manual segmentation of MTL structures (left / right TPC, AMG, ERC, PRC, and PHC) was conducted in native space (using ITK-SNAP [35]) by a trained researcher (ARF) according to segmentation procedures based on published atlases and protocols [36,37], described in [38]. We also calculated the volume of the left and right parahippocampal gyri by summing the left and right ERC, PRC, and PHC, respectively. The volumes of all structures were corrected for total intra-cranial volume (TIV), calculated from the unified segmentation procedure in SPM12 and expressed as z-scores based on the mean volume and SD of a group of 48 healthy controls (age: median = 64.85; IQR = 15.56 years; sex: 23M:25F). Overall, the patient group did not differ from the group of 48 healthy controls in terms of M:F ratio (7M:2F; χ^2=2.71, p = 0.100) or age at research scan (median = 56.93; IQR = 11.78; U = 148, p = 0.142).

Whole-brain Voxel-based Morphometry (modulated grey matter)

In order to ensure that none of our patients presented with GM volume reduction beyond the MTL, we conducted a series of VBM analyses, contrasting each patient's whole-brain modulated GM tissue maps (reflecting GM volume) against those of 67 datasets of healthy controls, previously presented in [14]. VBM was conducted using the Statistical Parametric Mapping software (SPM12 v. 7219; http://www.fil.ion.ucl.ac.uk/spm/software/spm12) in Matlab R2018b. Images were examined for scanner artefacts and reoriented to have the same point of origin (anterior commissure) and spatial orientation. They were then bias-corrected to remove intensity non-uniformities, and segmented into GM, white matter (WM), and cerebrospinal fluid (CSF) with the unified segmentation procedure. The diffeomorphic anatomical registration through the exponentiated lie algebra (DARTEL) toolbox was applied to participants' GM, WM, and CSF to refine inter-subject registration, and study-specific GM templates were generated [39]. After affine registration of the GM DARTEL templates to the tissue probability maps in MNI (Montreal Neurological Institute, Quebec, Canada) space, non-linear warping of GM images was performed to this template in MNI space. Voxel values in the tissue maps were modulated by the Jacobian determinant that was calculated during spatial normalisation, with modulated GM images reflecting tissue volume. These images (voxel size: 1 mm^3 isotropic) were smoothed using a large Gaussian filter of 12 mm FWHM, as recommended for unbalanced designs [40]. In a series of analyses, we compared GM volume between each single patient and the group of 67 healthy controls (contrast: 'healthy controls > patient'; second-level between-subject covariates: age, sex, TIV, study). As appropriate for case-controls designs, equality of variance was assumed [41]. We examined clusters surviving whole-brain FWE-correction ($p<0.05$) at peak-voxel level over $p<0.001$ (uncorrected). Volume reduction was not detected in any patient beyond the MTL.

Behavioural Paradigms

Paradigm 1: ROC

The first paradigm was based on a paradigm that examines receiver-operating characteristics (ROC) derived from the distribution of confidence responses across previously and newly encountered items, enabling the dissociation of recollection and familiarity processes [see [42,43] for methods]. It has been employed in several studies that examine the impact of MTL damage on recognition memory [e.g. [44–46]]. We examined whether impairment in recollection and familiarity for different memoranda is associated with damage in the HPC and the parahippocampal gyrus.

Stimulus Materials

Three stimulus types were used: faces, scenes, and words.

Faces

We used 160 pictures (targets: n=80;foils:n=80) of unknown Caucasian faces (front view) from the Face Database [47], presented in the centre of the display (17 cm wide, 11 cm tall). These were faces of individuals from a broad age range (18-91 years of age). The pictures involved a neutral grey background provided by a portable projection screen. All photos were taken under natural lighting. The target and foil faces were matched for age (targets: M = 61.50, IQR = 48.00; foils: M = 58.00, IQR = 45.50; U = 3196.50, p = 0.991) and for M:F ratio (targets: 24:56; foils: 26:54; $\chi^2 = 0.12$, p = 0.73).

Scenes

We also included 160 pictures of natural landscapes (targets: n = 80; foils: n = 80), presented in the centre of the display (17 cm wide, 11 cm tall) that were i) not identifiable

/ known; ii) with no sign of manmade features (buildings, objects), or of people or animals. These images were taken from the royalty-free platform Shutterstock (https://www.shutterstock.com). The set of 80 target scenes that were presented in the encoding phase did not differ from the set of 80 foils with respect to the composition of the different themes (Autumn: 4:4; Beach: 6:5; Clouds: 4:4; Desert: 7:7; Forest: 5:6; Hills: 7:6; Lake: 5:6; Mountains: 8:7; River: 6:4; Rocks: 10:11; Sea: 4:6; Waterfalls: 4:6; Winter: 10:8; $\chi^2 = 1.89$, $p > 0.999$).

Words

Our set of word stimuli comprised 160 words (targets: n=80; foils: n = 80), presented in the centre of the display (font size: 28). The words were common nouns in singular number; targets and foils did not differ in i) corpus frequency [SUBTLEXUS word frequency [48]] (targets: M = 15.58, IQR = 28.32; foils: M = 15.14; IQR = 15.78 occurrences per million words; targets vs. foils: U = 3039.50, p = 0.585); ii) length (targets: M = 5; IQR = 2; foils: M = 5, IQR = 2; targets vs. foils: U = 3074.50, p = 0.658); iii) mean concreteness ratings (targets: M = 5.81, IQR = 1.54; foils: M = 5.80; IQR = 1.87; targets vs. foils: U = 1684.00, p = 0.631); iv) mean imageability ratings (targets: M = 5.91; IQR = 1.39; foils: M = 5.97, IQR = 1.74; targets vs. foils: U = 1688.50, p = 0.648); v) mean familiarity ratings (targets: M = 5.65, IQR = 0.86; foils: M = 5.71, IQR = 0.77; targets vs. foils: U = 1701.50, p = 0.698); vi) age of acquisition (targets: mean = 3.43; SD = 1.01; foils: mean = 3.34 ; SD = 0.91; targets vs. foils: t = 0.49, p =0.63); vii) mean ratings of arousal levels (targets: mean = 4.49; SD = 0.94; foils: mean = 4.29; SD = 0.98; targets vs. foils: t = 1.12, p = 0.26); viii) mean valence ratings (targets: M = 5.29; IQR = 0.77; foils: M = 5.20; IQR = 1.06; targets vs.

foils: U = 1653.50, p = 0.521) [see [49] for details on ratings of concreteness, imageability, familiarity, age of acquisition, arousal, and valence].

Procedure

The experiment was written in Matlab, using the Psychophysics Toolbox (v.3) extensions [50–52]. Each participant was tested in a quiet room. The session lasted approximately 45 minutes. The order of trial blocks is illustrated in Supplementary Figure 1. Each stimulus was first presented to participants in the encoding phase, before testing in the recognition memory phase. In both the encoding and test phases, each trial started with the presentation of a fixation cross for 0.5 seconds at the centre of the display, replaced by a stimulus. In the encoding phase, participants were asked to judge if each stimulus was "pleasant", "neutral" or "unpleasant". Word stimuli were presented for 3 seconds, and face and scene stimuli for 4.5 seconds, irrespective of participants' response latencies. In the test phase, participants were asked to judge whether the presented item had been previously encountered in the encoding phase on a 6-point confidence scale (1=definitely new; 2=probably new; 3=maybe new; 4=maybe old; 5=probably old; 6=definitely old), in a self-paced fashion. This phase included all of the stimuli that had been previously presented in the encoding phase (targets), along with an equal number of novel stimuli (foils). Participants were asked to make full use of the confidence scale. Based upon extensive piloting, we equated levels of difficulty across material types, using two encoding and, correspondingly, two test phases for scenes and faces, but one encoding and one test phase for words. Moreover, the encoding phase for words was positioned at the beginning of the session, and the recognition phase for words at the end, similar to other studies [e.g. [5]]. The order of blocks was kept constant across participants, in order

to enable the comparison of the single cases (HPC-PRC+, HPC+PRC+) with healthy controls and HPC+PRC- patients. As mentioned above, faces and scenes also remained on the display 1.5 seconds longer than words. Indeed, parameter estimates for healthy controls did not differ across the three material types for either recollection (one-way repeated-measures ANOVA; independent variable: Material(3); $F=0.322$, $p=0.728$, $\eta^2=0.028$; pair-wise t-tests: all ts, $|t| \leq 0.759$; all ps, $p \geq 0.464$) or familiarity ($F=0.773$, $p=0.474$, $\eta^2=0.066$; pair-wise t-tests: all ts, $|t| \leq 1.23$; all ps, $p \geq 0.244$), suggesting that the three material types did not differ with respect to difficulty.

A filler task was also introduced in a series of blocks interspersed within the session, in order to minimise the influence of working memory, as well as to amplify forgetting between encoding and test phases. In each trial, two numbers were presented side-by-side at the centre of the screen. Participants were required to answer a question below those two numbers, asking participants to decide which of the two numbers was higher or lower. Participants selected '1' for the number on the left, '2' for the number on the right, or '3' if the two numbers were equal. Participants were given 3 seconds to respond, before the new trial started.

Behavioural Data Analysis

The behavioural results were analysed with a dual-process dissociation algorithm, which takes into account the distribution of responses (confidence judgments) to derive a single value of recollection and familiarity for each participant [53]. We used an algorithm based on the original one of Yonelinas and colleagues (available at http://psychology.ucdavis.edu/labs/Yonelinas/DPSDSSE.xls), implemented in Matlab

code (http://www.ruhr-unibochum.de/neuropsy/tests/memorysolve.zip), and reported in [54].

Paradigm 2: RDP

A second paradigm was also used to provide estimates for recollection and familiarity for faces, scenes, and words. This paradigm was based on a response deadline paradigm (RDP) and predicated on the selective reliance of recognition memory on familiarity in short response deadlines, in contrast with long response deadlines.[46]

The paradigm was administered in two separate sessions, one including a short response deadline (800 ms), and another involving a long response deadline (2,400 ms). The session including the long response deadline was administered first across participants on different days, with a minimum of a 5 days' delay between the two sessions, so as to prevent interference from the first session in the second session. Patients and healthy controls did not differ in the delay between the two sessions (Patients: M = 14; IQR = 227 days; healthy controls: M = 14; IQR = 122.50 days; U = 38, p = 0.861). Moreover, we ensured that the first session of the RDP was administered on a different day from the ROC, with a minimum of a 1 day's delay across participants. Healthy controls and patients did not differ on the length of the delay between the ROC and the first RDP session (Patients: M = 30; IQR = 196 days; healthy controls: M = 30; IQR = 141.50 days; U = 36, p = 0.712).

Stimulus Materials

Faces

We used a total of 120 faces (Short Deadline Session: targets: n=30;foils:n=30; Long Deadline Session: targets: n = 30; foils: n = 30) of unknown Caucasian faces (front view) from the Face Database [47], presented in the centre of the display. These were faces of individuals from a broad age range (18-92 years of age). The pictures involved a neutral grey background provided by a portable projection screen. All photos were taken under natural lighting. The face stimuli used in the short deadline session did not differ from those in the long deadline session in either age (Short Deadline Session: M = 63.00; SD = 45.75 years of age; Long Deadline Session: M = 61.00; SD = 47.75 years of age; Short vs. Long Deadline Session: U = 1747.50, p = 0.785) or M:F ratio (Short Deadline Session: 20M:40F; Long Deadline Session: 19M:41F; Short vs. Long Deadline Session: $\chi^2 = 0.038$, p = 0.845). Targets and foils did not differ with respect to either age or M:F ratio in either the Short (age: targets vs. foils: U = 444.5, p = 0.939; M:F ratio: targets vs. foils: $\chi^2 < 0.0005$, p > 0.999) or in the Long Deadline Session (age: targets vs. foils: U = 429, p = 0.761; M:F ratio: targets vs. foils: $\chi^2 = 0.077$, p = 0.781). The face stimuli used in the RDP were different from those used in the ROC paradigm.

Scenes

Scene-stimuli involved 120 pictures of natural landscapes (Short deadline session: targets: n = 30; foils: n = 30; Long deadline session: targets: n = 30; foils: n = 30), presented in the centre of the display that were i) not identifiable / known; ii) with no sign of manmade features (buildings, objects), or of people or animals. These images were taken from the royalty-free platform Shutterstock (https://www.shutterstock.com). The scene

stimuli used in the RDP were different from those used in the ROC paradigm. The Short deadline session did not differ from the Long deadline session with respect to the composition of the different themes across the scenes presented [Short deadline session: autumn(n=3), beach (n=4), clouds (n=3), desert (n=5), forest (n=5), hills (n=5), lakes (n=4), mountains (n=5), rocks (n=7), rivers (n=4), sea (n=5), winter (n=7), waterfalls (n=3); Long deadline session: autumn(n=3), beach (n=5), clouds (n=2), desert (n=5), forest (n=4), hills (n=5), lakes (n=5), mountains (n=4), rocks (n=8), rivers (n=4), sea (n=5), winter (n=7), waterfalls (n=3); Short vs. Long Deadline session: $\chi^2 = 0.711$, $p > 0.999$]. Likewise, no such differences were noted between target and foil items in either the Short ($\chi^2 = 2.286$, $p > 0.999$) or the Long Deadline session ($\chi^2 = 1.810$, $p > 0.999$).

Words

Our set of word stimuli comprised 60 words (targets: n=30; foils: n = 30), presented in the centre of the display. The words were i) common nouns; ii) in singular number; iii) Targets and foils did not differ in corpus frequency [SUBTLEXUS word frequency [48]] (targets: M = 17.36; IQR = 31.52 occurrences per million words; foils: M = 17.39; IQR= 26.43 occurrences per million words; targets vs. foils: U = 437, p = 0.854); iv) length (targets: M= 5; IQR = 2; foils: M = 5; IQR = 2; targets vs. foils: U = 427, p = 0.741); v) mean concreteness ratings (targets: M = 6.00, IQR = 1.98; foils: M = 5.68; IQR = 1.76; targets vs. foils: U = 252.5, p = 0.995); vi) mean imageability ratings (targets: M = 6.06, IQR = 1.34; foils: M = 6.13, IQR = 1.29; targets vs. foils: U =241, p =0.792); vii) mean familiarity ratings (targets: M = 5.94, IQR = 0.76; foils: M = 5.94, IQR = 0.73; targets vs. foils: U = 239.5, p = 0.766); viii) age of acquisition (targets: M = 2.87; IQR = 1.48; foils: M = 2.60; IQR = 1.43; targets vs. foils: U = 220.5, p = 0.468) ; ix) mean ratings of arousal levels (targets: mean =

4.54, SD =1.05; foils: mean = 4.31; SD = 0.94; targets vs. foils: t = 0.77, p = 0.443); x) mean valence ratings (targets: M = 5.47, IQR = 1.04; foils: M = 5.49; IQR = 0.81; targets vs. foils: U = 242.0, p = 0.809) [see [49] for details on ratings of concreteness, imageability, familiarity, age of acquisition, arousal, and valence]. The same words were presented in the two sessions.

Procedure

Stimulus presentation and data logging were programmed using the Psychophysics Toolbox (v.3) extensions [50–52]. The encoding phase of the paradigm involved 3 blocks (faces, scenes, words) of 30 trials each. Participants were asked to rate each stimulus according to pleasantness ('Unpleasant', 'Neutral' or 'Pleasant'). They had 3 seconds to rate pleasantness of words, and 4.5 seconds to rate pleasantness of faces and scenes. In the retrieval/test phase, participants were required to judge if the item presented on the screen was previously encountered in the encoding phase (pressing '1' for 'Old') or not (pressing '9' for 'New'). The items were presented over 60 trials, broken down into 6 blocks of 10 trials with breaks after each block.

In each trial, a fixation cross was first presented, followed by the item, which was presented for either 400ms (short response deadline) or 2000 ms (long response deadline). The participant was required to observe the item (face, scene, word) without responding. The item was then bordered in a blue square for 400 ms, during which time the participant was required to provide their response by pressing the 'OLD' or the 'NEW' button. An error noise was triggered for responses generated before the onset or after the offset of the response window.

For the same reasons as those described for the ROC above, a series of blocks of filler trials were interspersed within the session, comprising 20 trials each, with a response window of 3 seconds per trial. Participants were presented with two numbers on the screen, and were asked to select which number was the highest or the lowest. They pressed the 'left' (arrow to select the number presented on the left side of the display, and the 'right' arrow for the number presented on the right side of the display. Participants pressed the 'down' arrow to respond that the numbers were equal. The session structure is presented in Supplementary Figure 2.

Behavioural Data Analysis

Participants' sensitivity indices (d') in the short deadline session were derived, reflecting a familiarity estimate. The increase in d' in the longer deadline relative to the short deadline was assumed to rely on recollection, similar to other studies using the RDP [46].

Paradigm 3: SRP

We used the Source Recognition Paradigm (SRP) as reported in [15]. The SRP is a continuous source monitoring paradigm, in which foreground objects are presented on background scenes. Each object-scene pair is presented twice: on one half of the repetitions, the object switched its left-right location on the scene. On each trial, participants make a three-way decision of: "new" (first time object-scene pair seen), "stay" (second time pair seen, with same object location) and "move" (second time pair seen, but object location switched). This paradigm may be seen as providing more direct, objective measures of recollection and familiarity, since it does not rely on inferences based on subjective confidence or response speed.

Stimulus Materials

A colour photograph of an object was superimposed on the left / right side of a colour photograph of a background scene (Supplementary Figure 3). 246 object-scene pairs were randomly divided into two blocks with unique stimuli. The 123 stimuli were shown twice within a block. In their first presentation, half of the objects were randomly assigned to start on the left, and the other half on the right. In their second presentation, 61 of the object-scene stimuli were randomly assigned to the "stay" condition and the other 62 were assigned to be in the "move" condition. See [15] for further details.

Procedure

E-Prime 2.0 was used for stimulus presentation and response data logging. Participants were instructed to respond "new", "stay" or "move" using their left index finger, right index finger, or right middle finger, respectively. Each object-scene pair was presented in the centre of a grey screen (800 ms), with instructions for the button-to-finger mapping displayed at the bottom of the screen. A green fixation cross was overlaid on the centre of the screen. A button-press response was required for each trial before the onset of the following trial. If a response was made during the object-scene presentation, then the following object-scene was presented prior to a random interval (50-100 ms) after stimulus offset, during which the fixation changed to a red circle to prepare participants for the following trial. If no response was detected, a grey screen with red fixation circle remained until the response was given, followed by the same random interval.

Behavioural Data Analysis

Multinomial processing tree (MPT) models were used to parameterise Item and Source memory, based on the 9 response categories involved in fitting those MPTs: i) Stay trials called "Stay" (Correct source); ii) "Move" (Incorrect source); iii) "New" (misses); iv) Move

trials called “Stay” (Incorrect source); v) “Move” (Correct source); vi) or “New” (misses); vii) New trials called “Stay” (False Alarm); viii) “Move” (False Alarm); ix) or “New” (Correct Rejection). We derived separate estimates for Item (Di) and Source memory (Ds) based on the Source-Item model, which makes the assumption that two memory processes contribute to memory in an exclusive fashion. (see [15] for further details). Between-groups differences in Di and Ds values did not differ substantially when using the Item-Source model, which makes the assumption that Source Memory is a subset of Item Memory, possibly subserved by a single memory process.

Statistical Analysis

For each paradigm and condition, patients’ familiarity and recollection estimates were z-transformed on the basis of the mean and SD of the corresponding estimates of healthy controls. We subsequently examined whether i) PRC+ patients showed lower familiarity estimates relative to healthy controls and the HPC+PRC- patients by separately comparing HPC-PRC+ and HPC+PRC+ with healthy controls as well as with HPC+PRC- patients using appropriate comparisons for case-controls designs [55–57]; ii) HPC+PRC- patients showed lower recollection and/or familiarity estimates relative to healthy controls; iii) recollection and familiarity estimates were a function of HPC, or parahippocampal volumes across patients. Variance homogeneity was assessed using Levene’s test, and normal distribution using the Shapiro-Wilk test (not normally distributed if $p < 0.05$) and skewness and excess kurtosis values (not normally distributed if $|z| > 1.96$). Parametric (Student t-test; Welch t-test used when the assumption of homogeneity of variances was violated) and non-parametric tests between groups (Mann-Whitney U employed when the assumption of normal distribution was not met in a group) were used appropriately, as were bivariate correlations (Pearson’s r and

Spearman's rho). Comparisons of correlations between dependent samples were conducted using [58], based on [59] (single-sided testing). For not normally distributed data, medians (M) and interquartile intervals (IQR) are reported instead of means and standard deviations (SD). The following effect size estimates are also reported: η_p^2 for main effects and interactions in ANOVAs; η^2 for non-parametric comparisons; Cohen's d for between-groups t-tests. In paradigm 3, MPTs were fit using the "MPTinR" package (version 1.8.0) [60], implemented in R (version 3.2.3)(R Core Team, 2015). See [15] for further details. The rest of the statistical analyses were performed using SPSS (version 25.0, SPSS Inc).

Data availability

Behavioural data will be made publicly available post-acceptance at: https://osf.io/a82ht/?view_only=3e3c4a4d84b545fab18e79ec604a0c36

References

1. Squire, L.R. & Wixted, J.T. *Annu. Rev. Neurosci.* **34**, 259–88 (2011).

2. Montaldi, D. & Mayes, A.R. *Hippocampus* **20**, 1291–1314 (2010).

3. Wixted, J.T. & Squire, L.R. *Trends Cogn. Sci.* **15**, 210–217 (2011).

4. Bird, C.M. & Burgess, N. *Curr. Biol.* **18**, 1932–1936 (2008).

5. Cipolotti, L. et al. *Neuropsychologia* **44**, 489–506 (2006).

6. Huang, B.Y. & Castillo, M. *RadioGraphics* **28**, 417–439 (2008).

7. Köhler, S. & Martin, C.B. *Neuropsychologia* 107339 (2020).doi:10.1016/J.NEUROPSYCHOLOGIA.2020.107339

8. Inhoff, M.C. et al. *Neuropsychologia* **124**, 9–18 (2019).

9. Lambon Ralph, M.A., Patterson, K. & Plaut, D.C. *Cogn. Neuropsychol.* **28**, 466–474 (2011).

10. Robin, J., Rai, Y., Valli, M. & Olsen, R.K. *Hippocampus* **29**, 313–339 (2019).

11. Zeidman, P., Mullally, S.L. & Maguire, E.A. *Cereb. Cortex* **25**, 3836–55 (2015).

12. Kim, S., Dede, A.J.O., Hopkins, R.O. & Squire, L.R. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 4767–72 (2015).

13. Lacot, E. et al. *Neuropsychologia* **104**, 76–91 (2017).

14. Argyropoulos, G.P.D. et al. *Elife* **8**, (2019).

15. Cooper, E., Greve, A. & Henson, R.N. *Cortex* **91**, 297–315 (2017).

16. de Vanssay-Maigne, A. et al. *Neuroimage* **58**, 1131–1138 (2011).

17. Yonelinas, A.P. et al. *Hippocampus* **17**, 1134–1140 (2007).

18. Brandt, K.R., Eysenck, M.W., Nielsen, M.K. & von Oertzen, T.J. *Brain Cogn.* **104**, 82–92 (2016).

19. Suzuki, W.L. & Amaral, D.G. *J. Comp. Neurol.* **350**, 497–533 (1994).

20. Kafkas, A. et al. *Hippocampus* **209**, 1–20 (2016).

21. Graus, F. et al. *Lancet Neurol.* **15**, 391–404 (2016).

22. Nelson, H.E. & Willison, J. (1991).

23. Wechsler, D. *WASI* (Psychological Corporation: 2011).

24. Mckenna, P. & Warrington, E.K. *J. Neurol. Neurosurgery, Psychiatry* **43**, 781–788 (1980).

25. Bozeat, S., Lambon Ralph, M.A., Patterson, K., Garrard, P. & Hodges, J.R. *Neuropsychologia* **38**, 1207–1215 (2000).

26. Delis, D.C.D. et al. *Can. J. Sch. Psychol.* (Psychological Corporation: 2001).

27. Wechsler, D. *Psychol. Corp.* (1997).

28. Butler, C.R. et al. *J. Neurol. Neurosurg. Psychiatry* **85**, 387–91 (2014).

29. Warrington, E.K. & James, M. *Thames Val. Test Co.* (1991).

30. Wechsler, D. (The Psychological Corporation: San Antonio, TX, 1997).

31. Rey, A. *Paris Les Ed. du Cent. Psychol. Appliquée* (1959).

32. Warrington, E. (NFER-Nelson: Windsor, UK, 1984).

33. Warrington, E.K. (Psychology Press: 1996).

34. Baddeley, A., Emslie, H. & Nimmo-Smith, I. (Thames Valley Test Co.: Bury St. Edmunds, England, 1994).

35. Yushkevich, P.A. et al. *Neuroimage* **31**, 1116–1128 (2006).

36. Insausti, R. et al. *Am. J. Neuroradiol.* **19**, 659–671 (1998).

37. Pruessner, J.C. et al. *Cereb. Cortex* **12**, 1342–1353 (2002).

38. Loane, C. et al. *J. Neurol. Neurosurg. Psychiatry* jnnp-2018-320168 (2019).doi:10.1136/jnnp-2018-320168

39. Ashburner, J. *Neuroimage* **38**, 95–113 (2007).

40. Salmond, C.H. et al. *Neuroimage* **17**, 1027–1030 (2002).

41. Mühlau, M. et al. *AJNR. Am. J. Neuroradiol.* **30**, 539–43 (2009).

42. Yonelinas, A.P. *J. Exp. Psychol. Learn. Mem. Cogn.* **20**, 1341–1354 (1994).

43. Yonelinas, A.P. & Parks, C.M. *Psychol. Bull.* **133**, 800–832 (2007).

44. Aggleton, J.P. et al. *Neuropsychologia* **43**, 1810–1823 (2005).

45. Yonelinas, A.P. et al. *Nat. Neurosci.* **5**, 1236–1241 (2002).

46. Bowles, B. et al. *Proc. Natl. Acad. Sci.* **104**, 16382–16387 (2007).

47. Minear, M. & Park, D.C. *Behav. Res. Methods, Instruments, Comput.* **36**, 630–633 (2004).

48. Brysbaert, M. & New, B. *Behav. Res. Methods* **41**, 977–90 (2009).

49. Scott, G.G., Keitel, A., Becirspahic, M., Yao, B. & Sereno, S.C. *Behav. Res. Methods* 1–13 (2018).doi:10.3758/s13428-018-1099-3

50. Brainard, D.H. *Spat. Vis.* **10**, 433–436 (1997).

51. Pelli, D.G. *Spat. Vis.* **10**, 437–442 (1997).

52. Kleiner, M., Brainard, D. & Pelli, D. *Perception* (2007).

53. Yonelinas, A.P., Dobbins, I., Szymanski, M.D., Dhaliwal, H.S. & King, L. *Conscious. Cogn.* **5**, 418–441 (1996).

54. Pustina, D., Gizewski, E., Forsting, M., Daum, I. & Suchan, B. *Behav. Brain Res.* **229**, 57–67 (2012).

55. Crawford, J.R., Garthwaite, P.H. & Porter, S. *Cogn. Neuropsychol.* **27**, 245–260 (2010).

56. Crawford, J.R. & Garthwaite, P.H. *Neuropsychologia* **40**, 1196–208 (2002).

57. Crawford, J.R. & Howell, D.C. *Clin. Neuropsychol.* **12**, 482–486 (1998).

58. Lenhard, W. & Lenhard, A. (2014).doi:10.13140/RG.2.1.2954.1367

59. Eid, M., Gollwitzer, M. & Schmitt, M. *Lehrb. ; mit Online-Materialien* (2010).

60. Singmann, H. & Kellen, D. *Behav. Res. Methods* **45**, 560–575 (2013).

Acknowledgments

We are very grateful to the participants who took part in this study. This research was supported by a Medical Research Council Clinician Scientist Fellowship to CRB (MR/K010395/1).

Ethics declarations

Competing interests

The authors declare no competing interests.

Author information

Affiliations

Memory Research Group, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK; Level 6, West Wing, John Radcliffe Hospital, Oxford, OX3 9DU, UK.

Georgios P. D. Argyropoulos, Carola Dell'Acqua, Emily Butler, Clare Loane, Adriana Roca-Fernandez, Azhaar Almozel, Nikolas Drummond, Carmen Lage-Martinez, Christopher R. Butler

Department of General Psychology & Padova Neuroscience Center, University of Padova, Padova, Italy

Carola Dell'Acqua

Maurice Wohl Clinical Neuroscience Institute, Basic and Clinical Neuroscience Department, King's College London, London, UK

Clare Loane

Cardiff University, School of Biosciences, Cardiff, UK

Azhaar Almozel

Department of Zoology, University of Cambridge, Cambridge, UK

Nikolas Drummond

Valdecilla Biomedical Research Institute, University Hospital Marqués de Valdecilla.

Carmen Lage-Martinez

MRC Cognition & Brain Sciences Unit, and Department of Psychiatry, Cambridge, UK

Elisa Cooper, Richard N. Henson

Department of Brain Sciences, Imperial College London, London, UK

Christopher R. Butler

Departamento de Neurología, Pontificia Universidad Católica de Chile, Santiago, Chile.

Christopher R. Butler

Contributions

CRB and RNH conceptualised the study. EB, ND, EC, CRB and RNH designed the experiments. GPDA, CDA, EB, CL, ARF, AA, ND, CLM, EC, and CRB collected the data. GPDA,

CDA, EB, CL, ARF, AA, EC, and RNH analysed the data. GPDA and CRB wrote the paper. CRB provided clinical assessment. CRB supervised the study.

Competing interests

The authors declare no competing financial interests.

Corresponding author

Correspondence to Georgios P. D. Argyropoulos.