Extent of hippocampal atrophy predicts degree of deficit in recall

Eva Zita Patai[a,1], David G. Gadian[a], Janine M. Cooper[a,2], Anna M. Dzieciol[a], Mortimer Mishkin[b,3], and Faraneh Vargha-Khadem[a,c,3]

[a]University College London Institute of Child Health, London WC1N 1EH, United Kingdom; [b]National Institutes of Health, Bethesda, MD 20892; and [c]Great Ormond Street Hospital for Children National Health Service Foundation Trust, London WC1N 3JH, United Kingdom

Which specific memory functions are dependent on the hippocampus is still debated. The availability of a large cohort of patients who had sustained relatively selective hippocampal damage early in life enabled us to determine which type of mnemonic deficit showed a correlation with extent of hippocampal injury. We assessed our patient cohort on a test that provides measures of recognition and recall that are equated for difficulty and found that the patients' performance on the recall tests correlated significantly with their hippocampal volumes, whereas their performance on the equally difficult recognition tests did not and, indeed, was largely unaffected regardless of extent of hippocampal atrophy. The results provide new evidence in favor of the view that the hippocampus is essential for recall but not for recognition.

hippocampus | developmental amnesia | recall | recognition

A long-standing issue in memory research is the identity of the mnemonic process served by each of the components of the medial temporal lobe (MTL). Recall, which is the ability to retrieve from memory a stimulus or its context that is no longer present, is a more complex function than recognition, which is the ability to identify a current stimulus as old or new, or to choose a previously encountered stimulus among competing distractors. Over the years, a large number of studies have been conducted, aiming to tease apart the relative contribution of different MTL structures to recall and recognition. Many investigators have proposed that these two distinct memory processes rely on different MTL structures, with recall being dependent on the hippocampus and recognition being supported by parahippocampal structures, such as the perirhinal and entorhinal cortices (1–5, but see ref. 6). Also, studies examining the effects of damage to the mammillary bodies and fornices, structures within the hippocampal circuit, have shown that volume loss is correlated with deficits in recall but not in recognition (7, 8).

However, there is also evidence suggesting that both processes rely on the hippocampus (for a review, see ref. 6). Most of the literature on either side of this controversy is composed of studies with single cases or small sample sizes. Indeed, to date, none of the studies involving humans has been able to demonstrate a clear relationship between memory process and hippocampal volume (HV), possibly due to lack of variability in HV loss. Additionally, many of the contradictory findings could result from measures used to test recognition and recall not being equated for level of difficulty.

Previous work from our group demonstrated that patients with developmental amnesia due to severe hippocampal pathology sustained early in life are seriously impaired in their ability to recall visual or verbal stimuli but are relatively unimpaired in recognizing them (9). By identifying a group of patients with varying extents of HV reduction, and applying a standardized measure of memory [the Doors and People test (D&P test) (10)] in which the recognition and recall subtests are matched for difficulty, we were able to test whether or not these mnemonic processes depend equally on the magnitude of hippocampal damage. The outcome could help determine the specific mnemonic role of the hippocampus.

Methods

Participants. Using the D&P test (10), we assessed the memory ability of 29 patients (mean age, 16:2; range, 9–33 y; 12 females), who had sustained (*i*) hypoxic/ischemic episodes in the neonatal and/or perinatal period (see Table 1 for details) and (*ii*) HV reduction greater than 10% of the mean HV of a group of 65 healthy controls (11). (One female patient suffered respiratory failure at age 12, which is thought to be the cause of the hippocampal damage. We have included this patient, however, as the results do not change significantly when her data are removed.) (Three patients were excluded because they had full-scale IQ scores below 80, which is the lower cutoff for the normal range.) We compared the performance of the 29 patients with that of a subset of 26 members of the healthy control group selected to be comparable to the patient group in age and sex (mean age, 17:3; range, 9–35 y; 12 females). None of the patients had overt neurological impairment.

The study was approved by the Research Ethics Committee of University College London Hospital, and the participants and parents, where appropriate, gave informed consent.

Procedures. The D&P test was administered to all participants according to the instructions in the published manual (10) and as described in detail in ref. 9. The test consists of four subtests, two assessing recognition and two assessing recall, and, within each of these pairs, one assessing visual ability and the other, verbal ability. The four subtests, which are labeled People, Shapes, Names, and Doors, measure verbal recall, visual recall, verbal recognition, and visual recognition, respectively. For verbal recall (People test), four photographs each depicting an individual together with their printed name and occupation were presented on separate cards. After viewing the fourth picture, participants were asked to recall each name cued by their profession. This procedure was repeated until all four names were correctly

Significance

Which memory processes are dependent on the hippocampus is still being debated. Using data obtained from a large cohort of patients who had sustained hippocampal damage early in life enabled us to determine which type of mnemonic deficit showed a correlation with extent of hippocampal injury. Scores on tests of recall memory correlated significantly with hippocampal volumes, but scores on equally difficult tests of recognition memory did not. The results provide strong support for the view that the hippocampus is essential for recall but not for recognition.

Table 1. Number and sex of patients in each etiological group

Etiology	N (male:female)
Prematurity	3 (1:2)
Perinatal asphyxia	4 (3:1)
Acute respiratory failure (neonatal)	11 (7:4)
TGA	7 (4:3)
Congenital heart disease, respiratory failure, and epilepsy	1 (1:0)
Epilepsy	2 (1:1)
Respiratory failure due to surgical complications	1 (0:1)

TGA, transposition of the great arteries.

recalled or for a maximum of three presentations. Similarly, for Visual recall (Shapes test), participants copied each of four simple line drawings. They then tried to draw the four shapes from memory. This procedure was repeated until all four shapes were correctly recalled or for a maximum of three presentations. The Verbal recognition (Names test) subtest consisted of two study-test blocks. In the study phase of the first block, 12 female first names and surnames were presented on separate cards for 3 s each, and the experimenter read them aloud. Immediately thereafter, participants saw 12 lists of four names, each list presented on a separate card, and asked in each case to select the name from the study list. The same procedure was repeated in a second block, this time consisting of male names, but with the foils and the names on each test list differing from the study list in only one syllable of the surname. Finally, the Visual recognition (Doors test) subtest also consisted of two study-test blocks. In the study phase of the first block, participants viewed photographs of 12 doors, each presented on separate sheets accompanied by an appropriate label. Immediately thereafter, participants viewed 12 arrays of four doors, each on a separate sheet, and tried to identify the door from the study list. This same subtest was repeated with a second block consisting of 12 photographs of doors presented in exactly the same way as the first study-test block, but with foils that are more similar to the doors on the study list than is the case on the first block. As all four subtests were designed to be equally difficult based on the performance of a large group of healthy participants, the scores of our patients on all four subtests could be directly compared.

Scoring. Given that both patient and control groups included individuals under the age of 16, which is the first age band at which standard scores are available on the D&P test, we calculated standard scores based on the scores of our control group ($n = 26$). There was no significant difference in performance between our young controls (<16 y of age, $n = 16$) and our adult controls (>16 y of age; $n = 10$) (t test, all P values > 0.1). To confirm that our control sample was representative of the original control group used for the standardization of the D&P test, we compared the mean scores of our control group to the mean scores ($\pm$SD) available in the D&P test manual and found that the two scores on each test were nearly identical [raw scores from our sample: People, 27.9 (6.7); Shapes, 34.6 (2.5); Names, 18.4 (3.5); Doors, 18.6 (3.5); and sample means estimated from Fig. 1 in the D&P manual: People, 27.5 (5); Shapes, 34 (5); Names, 19 (4); Doors, 19 (5)]. Using

the data obtained from the group of 26 controls, we derived standard scores by calculating the means and SDs of the control group and converting the patients' raw scores to z scores relative to the control group's scores. We used nonparametric tests in all of the following analyses as the data were not normally distributed (Shapiro–Wilk test of normality was violated for the z scores of the following: Shapes ($P < 0.002$), and Names ($P = 0.027$) subtests; for the process of recall ($P = 0.049$); and for the retention of visual ($P = 0.01$) and verbal ($P < 0.001$) material. In addition, the z scores for visual retention showed a trend toward a nonnormal distribution ($P = 0.052$).

Imaging Procedures. Whole-brain MRI scans were obtained using a 1.5-T Siemens Avanto scanner, with a T1-weighted 3D fast low angle shot sequence: repetition time, 11 ms; echo time, 4.94 ms; flip angle, 15°; matrix size, 224 × 256; field of view, 250 mm; partition thickness, 1 mm; 176 sagittal partitions in the third dimension; acquisition time, 5.34 min. For the measurement of HVs, the datasets were reformatted into 1-mm-thick contiguous slices in a tilted coronal plane perpendicular to the caudorostral length of the hippocampus using MEDx 3.43 (Medical Numerics, Inc.). Hippocampal cross-sectional areas were measured by one of the authors (D.G.G.), using every slice. The volumes were calculated by summing the cross-sectional areas and multiplying by the distance between the measured slices. They were then corrected for intracranial volume and, unless otherwise stated, are presented as the mean of left and right (corrected) HVs for each individual.

The hippocampus (H) was defined as a composite of the following regions: cornu ammonis subfields (CA1–4), dentate gyrus, subiculum and presubiculum, amygdalo-hippocampal transition area, and uncus. The rostral boundary, marking the division between hippocampus and amygdala, was set at the alveus and anterior tip of the temporal horn of the lateral ventricle with the ventricle serving also as the lateral boundary. The medial boundary, marking the division between hippocampus and entorhinal cortex, was placed at the dorsomedial edge of the temporal lobe, except at the rostral and caudal tips of the hippocampus. The caudal boundary was set to include the last slice in which the hippocampus could be distinguished from the fornix.

All measurements were carried out blind to the behavioral data and group membership. For more detailed information on the procedures that were used for identifying and outlining the boundaries of the hippocampus, see ref. 11.
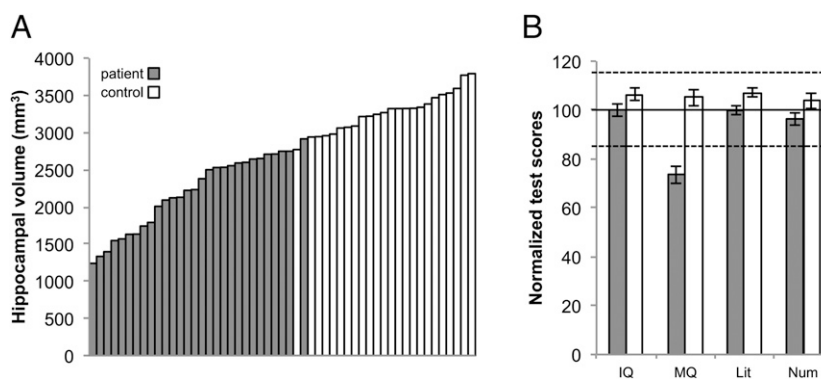
Fig. 1. (*A*) HVs and (*B*) neuropsychological profiles of patients and controls. MQ: memory quotient based on General Memory standard scores. Literacy: average of Reading Comprehension, Spelling, and Word Reading standard scores. Numeracy: average of Mathematical Reasoning and Numerical Operations standard scores. Solid line, mean score; dashed lines, ±1 SD.

NEUROSCIENCE

Results

Compared with the mean HV of a sample of 65 healthy individuals (hereafter referred to as the HI), the HVs of the patients ranged from 38% to 90% of the mean volume of the HI, whereas the HVs of our control group ranged from 90% to +115% of the mean HV of HI (Fig. 1A). (HVs were unavailable for two of the controls, as they did not undergo an MRI scan at the time of their visit.) There was no difference between left and right HVs ($t = -0.2$, $P = 0.9$). These reductions in HV in the patient group were not accompanied by either overall gray-matter loss ($t = -0.7$, $P = 0.5$), or an increase in the volume of cerebrospinal fluid ($t = -1.2$, $P = 0.2$), although overall white-matter density was reduced ($t = 2.3$, $P = 0.02$), as indexed by voxel-based morphometry (12).

The mean General Memory score or Memory Quotient (MQ) of the patient group also was below that of both our healthy control group and the general population [MQ mean and range, respectively: patients, 74, 45–111; controls, 105, 74–139: $t_{(53)} = 6.2$, $P < 0.001$]. The controls and the patients did not differ in age [independent-samples t test, $t_{(53)} = 0.7$, $P = 0.5$], or in IQ [IQ mean and range, respectively: patients, 100, 80–124; controls, 106, 88–130: $t_{(53)} = 1.7$, $P = 0.095$]. The patients' mean academic attainments were all in the average range (mean and range: mathematical reasoning, 97, 62–120; numerical operations, 95, 65–123; word reading, 101, 82–124; spelling, 94, 69–117; reading comprehension, 103, 67–122). Fig. 1B illustrates the scores of the patient group relative to the control group. The IQ, MQ, and Literacy and Numeracy standard scores are based, respectively, on the Wechsler Intelligence Scales, either the Wechsler Memory Scales or the Children's Memory Scales, and the Wechsler Individual Achievement Test.

All subtest scores, except names, were significantly below zero (all values of $t > -3.0$, $P < 0.002$), indicating that patients were impaired relative to controls (who represent zero, as their scores were used to convert raw scores into Z scores). Using the Friedman's test, we found that the patients' four subtest scores differed significantly from each other [X^2 (3, $n = 29$) = 16.5, $P = 0.001$]. Post hoc Wilcoxon signed-ranks tests indicated that the Shapes subtest yielded a significantly greater deficit than all of the other subtests (all values of $P < 0.002$), whereas the degree of deficit on the other subtests did not differ from each other. Collapsing across subtests, we found a greater deficit in recall compared with recognition ($Z = -2.41$, $P = 0.016$), and a greater deficit in memory for visual compared with memory for verbal material ($Z = -3.4$, $P = 0.001$). Comparison across recall and recognition, the scores from the two subtests that involve verbal material only (i.e., People vs. Names), yielded a trend toward a difference in the correlations: $Z = 1.8$, $P = 0.07$.

Given the large range in HV loss among the 29 patients, our primary aim was to determine whether the subtests differed in their degree of correlation between extent of hippocampal atrophy and size of deficit. The results (Fig. 2) indicated that performance on the People and the Shapes subtests (the two recall measures) correlated significantly with HVs [$r = 0.41$, $P = 0.03$ (note that the correlation with the People subtest does not survive the Bonferroni correction); and $r = 0.64$, $P < 0.001$, respectively], whereas correlations with Doors and Names subtests (the two recognition measures) were not significant ($r = 0.24$, $P = 0.2$; $r = 0.06$, $P = 0.8$, respectively). Collapsing across test material confirmed that scores on the recall subtests correlated with HVs ($r_s = 0.60$, $P = 0.001$), whereas scores on the recognition subtests did not ($r = 0.11$, $P = 0.5$); furthermore, these two correlations were significantly different from each other (Fisher's exact test: $Z = 2.5$, $P = 0.01$). The results were the same for HVs in each hemisphere, i.e., both the left and right HVs correlated with scores on the recall subtests (both $r_s = 0.5$, $P < 0.001$) but not with those on the recognition subtests (both $r_s = 0.2$, $P > 0.3$).

We also examined the interaction of left versus right HV as a function of verbal vs. visual test material, and found that both HVs correlated with visual (both $r > 0.6$, $P < 0.001$), but not with verbal material (both $r < 0.4$, $P > 0.06$).

Fig. 2. The patients' scores on People and Shapes, the two recall subtests, correlated significantly with HVs ($r = 0.41$, $P = 0.03$; $r = 0.64$, $P < 0.001$, respectively), whereas their scores on the Doors and Names subtests, the two recognition measures, did not ($r = 0.24$, $P = 0.2$; $r = 0.06$, $P = 0.8$, respectively).

Finally, we examined forgetting on the recall subtests, calculated by subtracting immediate memory scores from the delayed memory scores. Correlations between mean HV and forgetting scores revealed a significant forgetting for visual material ($r_s = -0.56$, $P = 0.002$), and for verbal material ($r_s = -0.38$, $P = 0.039$), but the latter did not survive correction for multiple comparisons. With regard to the hemispheric effects of hippocampal damage, again only visual material survived the Bonferroni correction (visual forgetting: both left and right hippocampi: $r > 0.6$, $P < 0.001$; verbal forgetting: both $r < 0.4$ $P > 0.02$), suggesting that the two hippocampi are equally involved in memory for visual material.

Discussion

A relationship between extent of selective hippocampal damage and degree of memory impairment has been reported in rats trained on a spatial learning task (13), as well as in a heterogeneous groups of adult patients with variable site and extent of damage to the brain (14–16); to our knowledge, however, there are no studies in the literature examining the effects of hippocampal damage in relation to different, but equally difficult, mnemonic processes. Part of the explanation for this is that any such demonstration requires a sizeable sample combined with a sizeable range in extent of hippocampal atrophy. Equally important is the need shown here to test the effect of varying extents of damage on recall memory, specifically. Recognition memory, by contrast, had already been found to be largely spared, regardless of the extent of hippocampal damage (17). Although Manns and Squire (18) did report a deficit in recognition memory in patients with selective hippocampal damage, their normal control group performed above average on the recognition subtest, possibly explaining the patient group's recognition deficit in that particular case.

Previous reports using the D&P test had already indicated a selective effect of hippocampal damage on recall (7–9, 19–22), but those studies did not examine whether this selective mnemonic effect was correlated with extent of hippocampal damage. Because our patient cohort was a large one and had wide variation in extent of hippocampal atrophy, we were able to examine this structure/function relationship and found a sharp distinction between recall and recognition performance, with only the former showing a correlation with extent of HV loss. This result is supported by a recent study from our group, showing that HV was correlated with context, but not item memory, whereas parahippocampal volume was correlated with item, but not context memory (ref. 23, supplemental information). Our current study extends these findings by providing additional support for the recall–recognition distinction in the role of the hippocampus in a larger sample than reported before, using tasks specifically equated for difficulty to test for differences in these two mnemonic processes.

Interestingly, there was also an effect of task materials, suggesting that the visual domain may be more sensitive to hippocampal damage than the verbal domain, a conclusion in line with recent evidence in humans that the hippocampus may play a role in tasks involving difficult visual discriminations (24–26). Indeed, a deficit in discrimination rather than in recognition may be part of the explanation for the patients' difficulty with the Doors subtest, which requires processing highly detailed visual information.

Finally, our results regarding the forgetting scores show that, besides the low baseline recall performance, patients' memory also exhibits additional decay over time. Unfortunately, the D&P test does not allow for testing forgetting in recognition; however, previous research has shown that forgetting is only detectable in recall or recollection tasks where there is a form of context reinstatement (27).

Our results have important implications for understanding the impact of the extent of early hippocampal injury on different sensory and memory processes. They indicate not only that recall and recognition are dissociable mnemonic processes that depend on different structures within the medial temporal lobe but also that the extent of hippocampal atrophy correlates with deficits in recall memory selectively.

1. Düzel E, Vargha-Khadem F, Heinze HJ, Mishkin M (2001) Brain activity evidence for recognition without recollection after early hippocampal damage. *Proc Natl Acad Sci USA* 98(14):8101–8106.
2. Eichenbaum H, Yonelinas AP, Ranganath C (2007) The medial temporal lobe and recognition memory. *Annu Rev Neurosci* 30:123–152.
3. Mishkin M, Vargha-Khadem F, Gadian DG (1998) Amnesia and the organization of the hippocampal system. *Hippocampus* 8(3):212–216.
4. Staresina BP, Fell J, Do Lam AT, Axmacher N, Henson RN (2012) Memory signals are temporally dissociated in and across human hippocampus and perirhinal cortex. *Nat Neurosci* 15(8):1167–1173.
5. Vargha-Khadem F, Gadian DG, Mishkin M (2001) Dissociations in cognitive memory: The syndrome of developmental amnesia. *Philos Trans R Soc Lond B Biol Sci* 356(1413): 1435–1440.
6. Squire LR, Wixted JT, Clark RE (2007) Recognition memory and the medial temporal lobe: A new perspective. *Nat Rev Neurosci* 8(11):872–883.
7. Tsivilis D, et al. (2008) A disproportionate role for the fornix and mammillary bodies in recall versus recognition memory. *Nat Neurosci* 11(7):834–842.
8. Vann SD, et al. (2009) Impaired recollection but spared familiarity in patients with extended hippocampal system damage revealed by 3 convergent methods. *Proc Natl Acad Sci USA* 106(13):5442–5447.
9. Adlam AL, Malloy M, Mishkin M, Vargha-Khadem F (2009) Dissociation between recognition and recall in developmental amnesia. *Neuropsychologia* 47(11):2207–2210.
10. Baddeley A, Emslie H, Nimmo-Smith I (1994) *Doors and People: A Test of Visual and Verbal Recall and Recognition* (Thames Valley Test Company, Bury St Edmunds, UK).
11. Cooper JM, et al. (2015) Neonatal hypoxia, hippocampal atrophy, and memory impairment: Evidence of a causal sequence. *Cereb Cortex* 25(6):1469–1476.
12. Ashburner J, Friston KJ (2000) Voxel-based morphometry—the methods. *Neuroimage* 11(6 Pt 1):805–821.
13. Moser MB, Moser EI, Forrest E, Andersen P, Morris RG (1995) Spatial learning with a minislab in the dorsal hippocampus. *Proc Natl Acad Sci USA* 92(21):9697–9701.
14. Allen JS, Tranel D, Bruss J, Damasio H (2006) Correlations between regional brain volumes and memory performance in anoxia. *J Clin Exp Neuropsychol* 28(4):457–476.
15. Kopelman MD, et al. (2001) Structural MRI volumetric analysis in patients with organic amnesia, 2: Correlations with anterograde memory and executive tests in 40 patients. *J Neurol Neurosurg Psychiatry* 71(1):23–28.
16. Kopelman MD, et al. (2007) Recall and recognition memory in amnesia: Patients with hippocampal, medial temporal, temporal lobe or frontal pathology. *Neuropsychologia* 45(6):1232–1246.
17. Zola SM, Squire LR (2001) Relationship between magnitude of damage to the hippocampus and impaired recognition memory in monkeys. *Hippocampus* 11(2):92–98.
18. Manns JR, Squire LR (1999) Impaired recognition memory on the Doors and People Test after damage limited to the hippocampal region. *Hippocampus* 9(5):495–499.
19. Aggleton JP, et al. (2005) Sparing of the familiarity component of recognition memory in a patient with hippocampal pathology. *Neuropsychologia* 43(12):1810–1823.
20. Baddeley A, Vargha-Khadem F, Mishkin M (2001) Preserved recognition in a case of developmental amnesia: Implications for the acquisition of semantic memory? *J Cogn Neurosci* 13(3):357–369.
21. Mayes AR, Holdstock JS, Isaac CL, Hunkin NM, Roberts N (2002) Relative sparing of item recognition memory in a patient with adult-onset damage limited to the hippocampus. *Hippocampus* 12(3):325–340.
22. Metzler-Baddeley C, Jones DK, Belaroussi B, Aggleton JP, O'Sullivan MJ (2011) Frontotemporal connections in episodic memory and aging: A diffusion MRI tractography study. *J Neurosci* 31(37):13236–13245.
23. Horner AJ, et al. (2012) A rapid, hippocampus-dependent, item-memory signal that initiates context memory in humans. *Curr Biol* 22(24):2369–2374.
24. Graham KS, Barense MD, Lee ACH (2010) Going beyond LTM in the MTL: A synthesis of neuropsychological and neuroimaging findings on the role of the medial temporal lobe in memory and perception. *Neuropsychologia* 48(4):831–853.
25. Graham KS, Gaffan D (2005) The role of the medial temporal lobe in memory and perception: Evidence from rats, nonhuman primates and humans. *Q J Exp Psychol B* 58(3-4):193–201.
26. Lee ACH, Barense MD, Graham KS (2005) The contribution of the human medial temporal lobe to perception: Bridging the gap between animal and human studies. *Q J Exp Psychol B* 58(3-4):300–325.
27. Green RE, Kopelman MD (2002) Contribution of recollection and familiarity judgements to rate of forgetting in organic amnesia. *Cortex* 38(2):161–178.

NEUROSCIENCE