Impaired recollection but spared familiarity in patients with extended hippocampal system damage revealed by 3 convergent methods

Seralynne D. Vann[a,1,2], Dimitris Tsivilis[b,1], Christine E. Denby[b,3], Joel R. Quamme[c], Andrew P. Yonelinas[d], John P. Aggleton[a], Daniela Montaldi[b], and Andrew R. Mayes[b]

[a]School of Psychology, Cardiff University, Cardiff CF10 3AT, United Kingdom; [b]School of Psychological Sciences, University of Manchester, Manchester M123 9PL, United Kingdom; [c]Department of Psychology, Grand Valley State University, Allendale, MI 49401; and [d]Department of Psychology, University of California, Davis, CA 95616

Edited by Mortimer Mishkin, National Institute for Mental Health, Bethesda, MD, and approved February 11, 2009 (received for review November 30, 2008)

To understand recognition memory, the detection of stimulus repetition, it first is necessary to resolve the debate between 2 fundamentally different models of recognition. Contemporary single-process models assume that recognition memory relies solely on the neural system required for the recall of prior events. Dual-process models assume that recognition comprises 2 independent forms of memory: one supports recall, and the other detects repeated stimuli by signaling their familiarity, the feeling of previous occurrence without the recall of any associated information. These 2 models were contrasted in patients who had undergone surgical removal of a colloid cyst, a condition associated with memory loss when accompanied by fornix and/or mammillary body atrophy. Comparisons were made between 2 groups of 9 patients that differed only with respect to the extent of mammillary body atrophy. Only the more atrophied group was impaired on tests of recall, but both groups showed normal recognition levels on a task that equates recall and recognition performance in normal participants. To explore the nature of this spared recognition, we estimated recall-based recognition and familiarity-based recognition using 3 distinct methods: self-report, receiver operating characteristics, and structural equation modeling. All 3 methods showed impaired recall-based recognition accompanied by intact familiarity in the most atrophied group, as predicted only by dual-process models. When structural equation modeling was applied to all 62 colloid cyst patients, the recall/familiarity dual-process model best explained the patients' memory pattern. The convergent evidence that mammillary body atrophy impairs recall but spares familiarity-based recognition appears irreconcilable with single-process models.

amnesia | colloid cyst | fornix | mammillary body | recognition

The present study tested one of the most contentious issues in current memory research: whether recognition memory reflects the operation of 2 distinct retrieval processes ("recollection" and "familiarity") or the output of a single common process or memory system. According to single-process models, subjective feelings of "remembering" a target (a form of cued recall) as opposed to "knowing" a target (an isolated sense of familiarity) merely reflect differences in the strength of the recognition signal (1–3). Dual-process models postulate that "recollection-based" and "familiarity-based" recognition rely, in part, on independent functions and distinct brain regions (4–7). Consequently, the 2 classes of model make different predictions about the fate of recognition memory following damage to sites assumed to be vital for recall such as the hippocampus and its interdependent structures (8–11). Only dual-process models predict that a complete sparing of familiarity-based recognition could occur in the presence of recall deficits. The present study tested this critical prediction by examining a large cohort of patients, all of whom had colloid cysts surgically removed from within the third ventricle, a condition associated with recall deficits arising from damage to the fornix and mammillary bodies (MBs) (12–15).

Previous studies have shown that colloid cyst removal often is associated with recall deficits but can appear to spare recognition (12, 16), an outcome more consistent with dual-process models. It first, however, is necessary to ensure that such recognition dissociations are not artifacts that arise from using recall tests that are more difficult than recognition tests. The Doors and People Test (17) controls for this artifact by equating recall and recognition in control participants. With this test it has recently been shown that MB atrophy in colloid cyst cases severely disrupts recall but spares recognition (13). Therefore the same cohort of patients is ideally placed to test the critical prediction of whether spared recognition reflects intact familiarity-based recognition (as assumed only by dual-process models). Techniques for distinguishing familiarity-based from recollection-based processes include (*i*) using subjective measures of "remembering" and "knowing" made during recognition testing (R/K); (*ii*) using confidence judgments to derive receiver operating characteristics (ROCs); and (*iii*) structural equation modeling (SEM) of relationships among recall and recognition scores. No agreed optimal approach exists, so the current study sought convergent evidence from these 3 independent methods.

To date, attempts to dissociate recollection and familiarity in patient studies have been inconclusive. Single-case and small-group studies have shortcomings, because much variation in individual patient performance often may arise because of premorbid differences in memory, variations in effective lesion location, and the inherent variation of cognitive measures (R/K and ROC) based on subjective experiences. Large-group studies provide the best opportunity to counter these shortcomings, but previous studies (11, 18, 19) have combined patients with differing etiologies and/or have not provided volumetric measures of sufficient key brain structures, thus severely limiting any conclusions. In contrast, the present study examined a cohort of patients with both a single etiology (colloid cyst) and quantified estimates of damage from multiple sites, including the extended hippocampal system (hippocampus, fornix, and MBs).

MRI-based volumetric measurements were available for a subset of 26 colloid cyst patients, all of whom had undergone neuropsychological testing. Two subgroups of patients (each $n = 9$) were

Author contributions: S.D.V., D.T., C.E.D., J.R.Q., A.P.Y., J.P.A., D.M., and A.R.M. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]S.D.V. and D.T. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: vannsd@cardiff.ac.uk.

[3]Present address: Department of Medical Physics and Clinical Engineering, Royal Liverpool University Hospital, Liverpool L7 8XP, United Kingdom.

This article contains supporting information online at www.pnas.org/cgi/content/full/0812097106/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.0812097106

Table 1. Summary of neuropsychological profiles in the small mammillary body (small MB) and large mammillary body (large MB) groups

Characteristic	Small MB (n = 9)	Large MB (n = 9)
Age (years)	46.6 ± 12.9	37.8 ± 9.2
Gender	male, 5; female, 4	male, 4; female, 5
Time between surgery/first test	8.5 ± 5.1 years	4.8 ± 4.2 years
Neuropsychological tests–standardized scores		
WAIS-III (norm = 100)		
VIQ	101.0 ± 16.3	107.8 ± 12.4
PIQ	102.9 ± 12.6	105.7 ± 9.8
FSIQ	102.3 ± 14.8	107.7 ± 11.3
WMS-III (norm = 100)		
Auditory immediate	87.2 ± 15.7	102.6 ± 9.8
Visual immediate	80.4 ± 13.0	95.4 ± 12.5
Immediate memory	80.9 ± 16.1	99.0 ± 12.0
Auditory delayed	82.7 ± 14.1	105.1 ± 10.2
Visual delayed	81.8 ± 13.9	98.9 ± 11.6
auditory recognition delayed	88.9 ± 11.4	100.6 ± 9.5
General memory	80.7 ± 13.7	102.2 ± 11.9
Working memory	106.6 ± 23.9	105.4 ± 11.0
Doors and People (norm = 10)		
Verbal recall (People)	6.4 ± 3.0	11.1 ± 3.3
Visual recognition (doors)	9.0 ± 2.4	10.8 ± 2.3
Visual recall (shapes)	6.4 ± 3.8	10.3 ± 2.9
Visual recognition (names)	10.2 ± 2.3	11.0 ± 2.3

Data are presented as means ± SD.

created based solely on differences in MB volume (large versus small). Estimates of recollection and familiarity were derived for these 2 subgroups using established R/K and ROC tests of word recognition. The critical prediction from dual-process models (8, 20) was that the subgroup of patients with the smallest MBs would show a disproportionate loss of recollection-based processes but spared estimates of familiarity. Next, SEM of visual recall and visual recognition data were applied to the results from a larger set of colloid cyst patients (n = 62), many of whom performed only IQ and standard memory tests. These analyses allowed comparisons of how well single- and dual-process models explained the nonverbal recognition data and also provided derived estimates of familiarity and recollection for the 2 subgroups of patients.

Results

Small-MB Group (n = 9) Versus Large-MB Group (n = 9). A total of 38 colloid cyst patients received a standardized, structural MRI protocol (13, 21, 22). Of these, 26 patients completed all cognitive tests, including R/K and ROC. Two subgroups, each of 9 patients, were drawn from this smaller cohort of 26 cases. The critical difference was that 1 subgroup (the "small-MB" group) contained the 9 patients with the smallest combined left and right MB volumes, and the other subgroup (the "large-MB" group) comprised the 9 colloid cyst cases with the largest MB volumes. The subgroups therefore represented the top and bottom thirds of the patients, based on MB volume, to maximize likely mnemonic differences between the 2 groups. There was no evidence that the large-MB group suffered any memory loss when comparisons were made with population norms (e.g., the Wechsler Memory Scale–third edition [WMS-III] including Face Recognition, Doors and People Test) (Table 1). Likewise, in the large-MB group there was no difference between full-scale IQ predicted memory performance on the WMS-III and actual performance (e.g., General Memory index, $t < 1$). The next set of analyses revealed how well-matched these 2 subgroups were on factors other than memory performance.

Surgical Status, IQ Performance, and Brain Volumetric Measurements. The small-MB group (n = 9) and the large-MB group (n = 9) did not differ in age ($t(16) = 1.66$, $P = 0.12$), length of time from surgery to first test session ($t(16) = 1.66$, $P = 0.12$), or gender balance (Table 1). Likewise, performance of the small-MB and large-MB groups did not differ on verbal, performance, or full-scale IQ as measured by the Wechsler Adult Intelligence Scale–third edition (WAIS-III) (all $t < 1$), which were all within normal limits (Table 1).

Structural volumes were calculated for 19 sites of interest, all of which are implicated in memory processes or are susceptible to damage from colloid cysts. Comparisons carried out using both raw volumes and intracranial volume (ICV) normalized scores found no volume differences between the 2 subgroups (all $P > 0.05$) with the inevitable exception of MB volume (both raw scores and ICV-normalized $P < 0.001$; supporting information (SI) Table S1). Indeed, the only other comparisons in which $P < 0.1$ were fornix volume [ICV normalized $t(16) = 1.92$, $P = 0.07$; fornix volumes were smaller for the small-MB group] and the entorhinal cortex volume [raw volume $t(16) = 1.98$, $P = 0.065$; entorhinal cortex volumes were larger for the small-MB group]. Finally, it should be noted that the MB volumes of both the small-MB and large-MB groups were smaller than those of a normal control group (13) used for volumetric analyses [mean raw volumes ± SD: controls = 0.068 cm^3 ± 0.010; large-MB group = 0.058 cm^3 ± 0.008; small-MB group = 0.015 cm^3 ± 0.01; for the large-MB group versus controls, raw scores: $t(10.8) = 7.9$, $P < 0.001$; for the large-MB group versus controls, ICV-normalized scores: $t(19) = 19.3$, $P < 0.001$].

Confirmation of Recall:Recognition Difference from Doors and People Test. An ANOVA carried out on the 4 subtest scaled scores (2 recall, 2 recognition) of the Doors and People Test revealed a significant subgroup difference ($F_{(1, 16)} = 8.3$, $P = 0.01$; Table 1). The group x subtest interaction did not reach significance ($P = 0.1$), probably because of the relatively small group numbers in the present study (13). Even so, the small-MB group was significantly worse than the large-MB group on both tests of recall (People, $P = 0.001$; Shapes, $P = 0.005$) but not on the 2 tests of recognition (Doors, $P = 0.19$; Names, $F < 1$). These results did not reflect abnormal performance of the large-MB group, because their mean scores for the 4 subtests (Table 1) were all just above the normalized

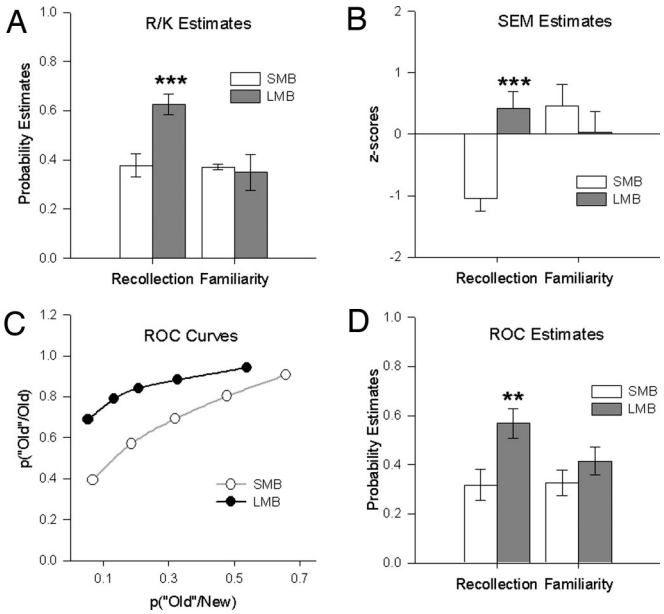

Fig. 1. Derived probability estimates of ''recollection'' and ''familiarity'' for the small-MB (SMB) and large-MB (LMB) groups, drawn from a cohort of postsurgery colloid cyst patients. Data are presented in histograms as means ± standard error of the mean. (*A*) Estimates from the Remember/Know (R/K) procedure. (*B*) Estimates from structural equation modeling (SEM). (*C*) Receiver operating characteristic (ROC) curve. (*D*) Estimates from receiver operating characteristic (ROC) procedure. **, $P < 0.01$; ***, $P < 0.005$.

scaled score of 10 (largest t (8) = 1.31, $P = 0.23$). Meanwhile, the small-MB group differed from the normalized scaled scores on both recall tests (People: t (8) = −3.55, $P = 0.008$; Shapes: t (8) = −2.80, $P = 0.023$) but not on the recognition tests (Doors: t (8) = −1.23, $P = 0.26$; Names: t <1).

Remember/Know Procedure. The R/K procedure, as described by Yonelinas et al. (19), provided estimates of familiarity and recollection for target sets of words. Fig. 1*A* shows the derived probability estimates for recollection and familiarity. Recollection was estimated as the probability of responding "remember" to an old item minus the probability of responding "remember" to a new item. Familiarity was estimated as the probability of a "familiar" response given that the item was not recollected, assuming it is not possible to give a "familiar" response to a recollected item, i.e., F/(1 − R). Familiarity was estimated in the same way for old and new items. This estimate assumes that the probability of familiarity occurring is independent of whether an item is recollected. New-item familiarity then was subtracted from that of the old items to account for potential differences in false-alarm rates.

Inspection of Fig. 1*A* immediately shows that although the 2 groups showed comparable levels of familiarity ("know"), the small-MB group had relatively reduced levels of recollection ("remember"). This pattern was supported by an overall ANOVA with the factors group, encoding condition (deep/shallow), and memory type (recollection/familiarity). There was an overall group difference ($F_{(1, 16)} = 6.2$, $P = 0.024$) as well as a group x memory-type interaction ($F_{(1, 16)} = 5.8$, $P = 0.028$) because the small-MB group had significantly lower estimates of recollection ($P = 0.002$), but there was no difference in estimates of familiarity ($F < 1$). There also was an overall effect of encoding condition ($F_{(1, 16)} = 20.3$, $P < 0.001$), with overall performance being better on the deep encoding condition. However, the lack of a significant 3-way interaction reflects the same pattern of responses across both deep and shallow conditions, i.e., for familiarity estimates the small-MB and large-MB groups were comparable, but for recollection the small-MB group scores remained lower (see Fig. S1). Consistent with this pattern of results, ICV-normalized MB volumes ($n = 26$) significantly correlated with the estimates of recollection ($r = 0.621$, $P = 0.001$) but not familiarity (r = −0.070, $P = 0.73$); these correlations are upheld even when controlling for patient age and time between surgery and test (see Table S2).

Receiver Operating Characteristics. The ROC procedure was identical to that previously used by Yonelinas et al. (23) and assessed confidence in recognition memory for studied words. ROC curves were plotted from the confidence ratings for the recognition responses made in a verbal yes/no recognition test. The average ROCs for the small-MB and large-MB groups are presented in Fig. 1*C*. The proportion of new items accepted as old (false alarms) and the proportion of old items accepted as old (hits) are plotted on the x- and y-axes, respectively. The left-most point on each function represents the proportion of items receiving the most confident old responses (i.e., a 6 response), and each consecutive point includes items receiving the next most confident old response (e.g., the second point includes items receiving 5 or 6 responses). From the graph (Fig. 1*C*), it can be seen that the 2 groups have differently shaped curves. The large-MB group has a curve that is skewed to the left. The curve for the small-MB group is much more symmetrical, indicating a greater reliance on familiarity than on recollection. This difference is supported by the higher value at which the large-MB group's curve intersects the *y*-axis.

The average ROC for each participant was quantified by fitting a nonlinear equation to the observed ROCs using a sum of squares search algorithm (23). The equation P("old" old) = P("old" new) + R + (1 − R) Φ(d'/2 − c_i)–Φ(−d'/2 − c_i) assumes that recognition reflects the contribution of recollection (R) and an independent familiarity process (d' reflects the distance between 2 equal-variance Gaussian strength distributions; c_i reflects the response criterion at point i; and Φ reflects the cumulative response function). To facilitate comparison to recollection, each d' value was converted to the probability of a hit given the average false-alarm rate. Familiarity accuracy then was measured by subtracting the average false-alarm rate from the calculated hit rate. Parameter estimates for recollection and familiarity were derived for each subject and are presented in Fig. 1*D*. Again, it can be seen that the 2 groups have equivalent levels of familiarity but differ in their use of recollective processes. An ANOVA carried out on these estimates revealed an overall group difference ($F_{(1, 16)} = 7.67$, $P = 0.014$). Although there was no group x memory-type interaction ($F_{(1, 16)} = 2.16$, $P = 0.16$), the small-MB group had significantly lower estimates only of recollection ($F_{(1, 32)} = 9.26$, $P = 0.005$); there was no difference in the estimates of familiarity for the large-MB and small-MB group ($F_{(1, 32)} = 1.17$, $P = 0.288$). There was no difference in the d' scores across the 2 groups [t (16) = 1.13, $P = 0.28$; means ± SE; small-MB group = 0.89 ± 0.44, large MB group = 1.14 ± 0.49]. As for R/K, correlations based on ICV-normalized MB volumes (n = 26) yielded a significant positive relationship for the ROC estimates of recollection (r = 0.472, $P = 0.015$) but not familiarity (r = 0.347, $P = 0.082$). Again, these correlation were upheld when controlled for age and time between surgery and test (Table S2).

Structural Equation Modeling. In all colloid cyst cases ($n = 62$), SEM was applied to visual recall and visual recognition data from the entire cohort of colloid cyst patients who had been assessed using the WAIS-III or Wechsler Abbreviated Scale of Intelligence (WASI) and WMS-III tests (13). Verbal memory performance was not examined because these tests do not contain a task that solely measures verbal recognition. The mean age of the patients was 48 years (SD, 12), and their ages ranged from 22 to 74 years. All patients performed within normal limits on tests of intelligence (mean IQ, 104.3; SD 13.1; range, 77–134). Estimates of famil-

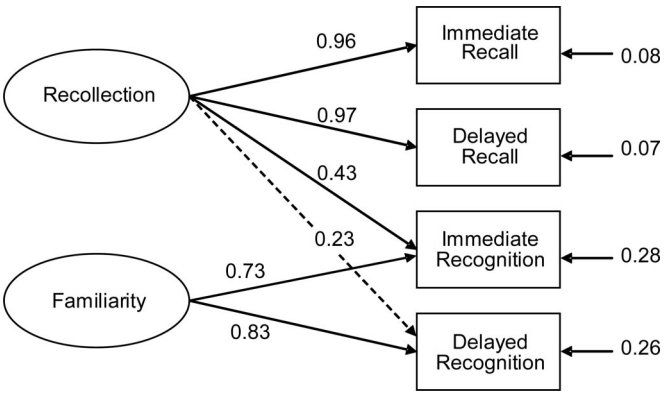

Fig. 2. Best-fitting path model relating recall to recognition in cohort of 62 colloid cyst patients using scores from immediate and delayed face recognition and family pictures tests taken from the WMS-III. The model assumes that recognition relies on recollection and familiarity, whereas recall relies solely on recollection. The ovals represent latent variables, and the rectangles represent measured variables. Solid lines represent significant regression coefficients. Dotted lines represent nonsignificant coefficients.

iarity and recollection were made by fitting a 2-process model to the standardized recognition and recall test data, using SEM (19, 24, 25).

The 2-factor SEM assumes that 1 latent variable (recollection) contributes to both recall and recognition, whereas a second latent variable (familiarity) contributes only to recognition (Fig. 2). The model provided a statistically acceptable account of the data, χ^2 (2) = 2.28, P = 0.31, root mean square error of approximation (RMSEA) = 0.048, consistent Akaike information criterion (CAIC) = 43.30. Of the 6 factor loadings estimated, all were significantly greater than zero except that between delayed recognition on recollection, z = 1.80, P = 0.07. However, because this parameter represents an a priori expectation, it was kept in the model for all model comparisons and for estimating factor scores. The low value of this parameter suggests a diminishing role of recollection in face recognition with increased delay. A single-factor model also was fit in which there was no familiarity factor (i.e., there was only 1 factor underlying all 4 tests). The fit of the single-factor model was statistically rejectable, χ^2 (4) = 30.28, $P <$ 0.001, RMSEA = 0.328, CAIC = 61.04. Moreover, the 2-factor model represented a statistically significant improvement over the 1-factor model, $\Delta\chi^2$ (2) = 28.00, $P < 0.001$. An alternative 2-factor model inconsistent with the dual-process theory of recognition was fit also. In this model, the second factor contributed to recall instead of recognition. The fit of this alternative model also was statistically rejectable, χ^2 (2) = 7.00, P = 0.03, RMSEA = 0.202, CAIC = 48.01. The alternative 2-factor model and the dual-process model are not nested models and thus cannot be compared statistically. However, practical fit measures indicate the dual-process model provided the superior fit, with lower RMSEA (26) and CAIC (27) values than the alternative 2-factor model. Thus, the SEM results indicated 2 factors were necessary to explain the covariances among the 4 tests and that a dual-process account of recognition specifically was necessary.

Small-MB Group Versus Large-MB Group (SEM). From the best-fitting solution of the first 2-factor model (Fig. 2), latent variable scores for recollection and familiarity factors were estimated for each subject in the small-MB and large-MB groups. Recollection and familiarity scores were standardized to a mean of zero and SD of 1.0 across all 62 cases; thus they are interpretable as z-scores. The mean estimates are depicted in Fig. 1*B*. An ANOVA carried out on these estimates revealed no overall difference between the 2 groups ($F_{(1, 16)}$ = 2.63, P = 0.12), but there was a significant group x memory-type interaction ($F_{(1, 16)}$ = 11.31, P = 0.004). This interaction reflected a significant difference between the small-MB and large-MB groups in terms of their recollection estimates ($F_{(1, 32)}$ = 11.87, P = 0.002) but not in their familiarity estimates ($F_{(1, 32)}$ = 1.04, P = 0.32). In addition, although the recollection estimates were significantly lower than the familiarity estimates in the small-MB group ($F_{(1, 16)}$ = 12.18, P = 0.002), there was no difference between the recollection and familiarity estimates in the large-MB group ($F < 1$). Consistent with the group analyses, correlations based on ICV-normalized MB volumes (n = 36) yielded a significant positive relationship for the SEM-derived estimates of recollection (r = 0.552, $P < 0.001$) but not familiarity ($r = -0.143$, P = 0.405). These correlations remained when controlled for age and time between surgery and test (Table S2).

Discussion

There are several reports that the amnesia associated with colloid cyst damage appears preferentially to disrupt recall rather than recognition (12, 13, 16). The present study therefore sought to understand the cause of any apparent recall: recognition dissociation in this patient group. In doing so, the present study also responded to the stated need to use group rather than single-case studies to address the more fundamental question of whether single-process or dual-process models can best explain recognition memory (1). Unlike previous group studies (11, 18, 28), the patients were not selected on the basis of known memory deficits; rather, they represent the continuum of pathological changes associated with a single etiology. This approach was made possible by having detailed neuropsychological assessments and quantified volumetric measures of multiple brain sites, including the MBs, for a cohort of 26 cases. Convergent findings from 3 different approaches all arrived at the same conclusion: pathology associated with MB atrophy produces a selective loss of recollection-based recognition while sparing familiarity-based recognition. These findings strengthen the view that the extended hippocampal system is selectively critical for processes leading to effective recall.

Previously, through the use of the Doors and People Test, Tsivilis et al. (13) showed a recall/recognition dissociation related to MB atrophy that was not an artifact of task difficulty. This same dissociation (poor recall, spared recognition) was confirmed in the small-MB group used in the present study. Earlier single-case studies reported this same dissociation on the Doors and People Test in patients with selective hippocampal damage (29, 30). This pattern, however, was not found in a group of 6 diverse amnesic patients with varying patterns of neuropathology in the hippocampus, frontal lobes, and thalamus (28) or in a group of 7 patients with hippocampal damage following cardiac arrest, carbon monoxide poisoning, drug overdose, or unknown causes (11). Because of such inconsistencies the present study adopted the more rigorous approach of examining a cohort with a single etiology and acquiring detailed volumetric structural information.

Independent estimates of recollection-based and familiarity-based verbal recognition were derived for 2 subgroups, each consisting of 9 colloid cyst patients (the large-MB and small-MB groups) using R/K decisions and ROC. Membership of the 2 subgroups was determined entirely on volumetric analyses compiled after cognitive testing; hence, all testing was blind to eventual group status. The rationale for subdividing the colloid cyst cases on the basis of MB volume arose from specific predictions (8, 13) that damage to this site should preferentially disrupt recollection-based recognition. The subsequent convergent findings of spared measures of familiarity but diminished recollection-based recognition in the small-MB group therefore were consistent both with these predictions and with the patients' previous performance on the Doors and People Test (spared recognition, impaired recall).

Prior indirect evidence of spared familiarity after extended hippocampal damage comes from the remarkable case of a man with selective bilateral MB damage caused by a snooker cue (31).

This man showed a striking sparing of recognition compared with recall. More specific support comes from the study of a single patient who had mammillothalamic tract damage who appeared to suffer a selective loss of recollection-based recognition when assessed using R/K and ROC tasks (20). The mammillothalamic tract carries projections from the MBs to the anterior thalamus, making these findings highly relevant, a view reinforced by studies in rats showing comparably severe memory impairments following lesions to the MBs or mammillothalamic tract (32). In a clear advance over these single-case studies (20, 31), the present study involved 2 patient subgroups matched for surgical procedure as well as age, IQ, and time since surgery. It also was possible to compare the volumetric status of brain sites outside the MBs. Of the other 18 brain sites measured, including the hippocampus, perirhinal cortex, and septum, the mean volumes in the small-MB group never were significantly smaller than those of the large-MB group. For these reasons the present results seem to provide some of the clearest R/K and ROC evidence to date that recollection and familiarity can be functionally dissociated in patients with MB/fornix atrophy.

An alternative strategy was to apply SEM to data from the entire cohort of colloid cyst cases. This approach again revealed the superior power of 2-process models of recognition to describe the patterns of memory performance. The current SEM results corroborate both the R/K and ROC data, but with a very different operationalization of recollection and familiarity. In the dual-process SEM, recollection is a latent variable that explains the variance shared in recall and recognition tests, and familiarity is a latent variable that explains additional variance shared in recognition tests that is not shared by recall. The SEM analysis indicated that this 2-factor structure provided a good account of the covariance among visual recall and recognition tests. The 2-factor model also was superior to a single-factor model in which there was only 1 underlying memory process. The SEM analysis also showed that an alternative 2-factor model in which the second factor contributed to recall instead of recognition did not provide a statistically acceptable fit. This result is important because it indicates that, despite the simplicity of the model and the small number of variables, it is not the case that just any 2-factor solution is sufficient to explain the data.

The model in the present study was applied to immediate and delayed versions of face recognition and family pictures recall because these tasks were relatively matched in content and provided appropriate data for the largest cohort of subjects (verbal tests were examined by the R/K and ROC methods). Because the recollection and familiarity factors were estimated from only these 2 kinds of tasks, the factor scores include an unknown amount of task-specific variance that probably would not be present across other, different tests of recall and recognition. However, it is precisely because recollection and familiarity factors are estimated from this narrow range of tests that the convergence across methods is notable: greater contamination of factors by task-specific variance should, if anything, have made the SEM results less likely to converge with R/K and ROC results. Furthermore, the present results are consistent with other SEM studies using different recall and recognition tests in different populations. SEMs with the same dual-process assumptions have provided successful accounts of word recall and recognition by patients who have suffered cerebral hypoxia (24, 33) and by elderly participants (25). Moreover, in the hypoxia studies, the 2-factor dual-process model provided a similarly superior account of the data over a single-factor model and an alternative 2-factor recall model. Thus, in addition to providing a cross-method and cross-materials corroboration of the R/K and ROC results of the present study, the current SEM findings provide a cross-sample and cross-materials corroboration of previous SEM studies.

A potential issue is whether the disproportionate loss of recollection-based recognition is merely the consequence of a mild recognition memory impairment, i.e., a loss of memory strength that removes those features of recognition (recollection) that reflect a stronger trace (e.g., ref. 1). One obvious problem with this single-process explanation is that the members of the small-MB group were unimpaired on the recognition components of the Doors and People Test and/or on the ROC task, as measured by d', despite very persistent recall deficits. A single-process memory account must, by its very nature, suppose some overall reduction in recognition performance if the task is sufficiently demanding, but this reduction was not evident. In contrast, dual-process models can predict this pattern of preserved recognition on the assumption that the performance of some recognition tasks depends minimally on recall. In addition, single-process models cannot accommodate a full sparing of familiarity when recall-based recognition is impaired, but all 3 tests showed that the 2 patient subgroups had comparable levels of familiarity. Although additional comparisons were not made with a normal participant group to provide baseline levels of familiarity (because there would have been inevitable group differences relating to illness, surgical procedures, and hydrocephalus), the consistent normal performance of the large-MB subgroup (Table 1) on standard recognition tests (as well as on standard recall tests) must suppose intact familiarity. Consequently, the assumption of intact familiarity also holds for the small-MB subgroup.

To date, the neuropsychological studies that appear to support dual-process models typically report impaired recall relative to recognition (e.g., 4, 19, 29, 34); however, if familiarity and recollection are truly dissociable, then the opposite pattern should be found also. Indeed, a recent study of an epileptic patient who underwent anterior temporal lobe resection that damaged the left rhinal area but left the hippocampus intact has reported this opposite pattern of impaired verbal familiarity with preserved recollection (9). This finding, combined with our finding that recollection-based recognition memory is disproportionately impaired across 3 independent measures in a group of patients selected solely in terms of pathology to the MB, provides some of the most convincing evidence to date in support of dual-process models of recognition memory.

Methods

Participants. A total of 62 participants were drawn from 14 neurological centers across England, Scotland, and Wales. All participants had a colloid cyst surgically removed from within the third ventricle at least 1 year before the investigation. A variety of surgical approaches (transfrontal, transcallosal, and endoscopic aspiration and excision) had been used to remove the cyst (21, 22). Patients were excluded if they had additional neurological disorders or were under 18 years of age. Although all patients (30 males, 32 females) were assessed using the WAIS-III or WASI and WMS-III, only a subset of 38 patients agreed to MRI scanning. The interval between the colloid cyst surgery and the subsequent MRI scan ranged from 12 to 240 months (mean, 79.9 months; SD 66.7).

Of the 38 patients who had MRI scans, 26 patients had completed all memory tasks, including the R/K and ROC tests, without any difficulty in understanding the task demands. This decrease in patient numbers reflected dropout by the participants rather than selective targeting by the investigators. The participants' data subsequently were placed into 2 groups based on the combined left and right ICV-normalized MB volumes (21), with 9 patients in each group (with the largest or smallest volumes, respectively). These groups, therefore, represented the top and bottom thirds of the 26 patients (based on MB volume). These 2 sets of 9 subjects comprised respective parts of the small-MB ($n = 11$) and large-MB ($n = 11$) groups described by Tsivilis et al. (13). Approval for this study was provided by a United Kingdom MultiCentre Research Ethics Committee (MREC). All participants gave their informed consent before inclusion in the study.

MRI Assessments of Neuropathology. The patients and 20 age-matched normal controls were scanned at the same center with the same protocols (21, 22). A list of the 19 regions for which volumes were assessed is given in Table S1. Anatomical and MRI delineations for most of these regions have been published elsewhere (21, 22).

Standard Neuropsychological Tests. Patients were tested on the WAIS-III (35), WMS-III (36), WASI (37), and the Doors and People Test of Recall and Recognition (17).

Cognitive Tests for SEM Analysis. Two measures of recognition memory and 2 measures of recall memory for visual materials were used for SEM analysis. All were taken from the WMS-III and comprised the immediate and delayed face recognition test and the immediate and delayed family pictures test. These tasks were chosen to provide measures of recollection and familiarity from visual materials (verbal materials are already represented in the R/K and ROC procedures) and because they are relatively closely matched in information content.

Remember/Know Procedure. The R/K procedure was identical to that previously described by Yonelinas et al. (19). Briefly, participants heard 100 words and had to make a deep or shallow decision about each word. They subsequently were given a recognition memory test in which they were read the 100 target words intermixed with 50 novel foils. Participants were asked to decide if the word was explicitly remembered (R), merely felt familiar (K), or was new. Further details are given in *SI Methods*. During a single session, participants first heard 25 words and made a shallow decision about each word (how many syllables the word contained). In the second stage, 50 new words were read out, and the subject made a deep decision about each word (giving the word a ''pleasant'' or ''unpleasant'' rating on a 6-point scale). These words were followed by another 25 words that required a shallow decision. Words were read out at a subject-paced rate of about 1 every 10 seconds. A recognition test was given immediately afterward in which subjects were read the 100 target words intermixed with 50 novel foils. Participants were asked to decide if the word was explicitly remembered (R), merely felt familiar (K), or was new. For the first 20 words and for several items spread throughout the test list, participants were required to explain why they made a particular response. None of the participants seemed to have any difficulties understanding the instructions.

ROC Analysis. Participants rated the confidence of their recognition responses. The procedure and test stimuli (words) were identical to those used by Yonelinas et al. (23), except that patients only received 1 session, and all words were subject to deep encoding. This task contained 160 target words and 80 foils. During the encoding phase participants decided whether the words were concrete or abstract. Words were presented at a subject-paced rate of about 1 word every 10 seconds. Immediately following the end of the encoding stage, the participants received a recognition test in which they were asked to rate the their confidence in their recognition responses on a 6-point scale.

Statistical Analysis. Group comparisons used parametric tests (e.g., ANOVA). When significant interactions were found, the simple effects were analyzed as recommended by Winer using the pooled error term (38); when there was a significant main effect but no interaction, the simple effects were examined to identify the specific tests in which performance differed significantly between groups (39). The probability level of < 0.05 was taken as being statistically significant. Structural equation modeling was performed in LISREL 8.3; details are given in *SI Methods*.

ACKNOWLEDGMENTS. The authors thank the participants and their families for their generous contributions to this project and the neurosurgeons for facilitating access to their patients. The authors also acknowledge the contribution of D. McMackin in the initial stages of this research project. This research was funded by Grant G0001371 from the U.K. Medical Research Council. S.D.V. is funded by a U.K. BBSRC David Phillips Fellowship.

1. Squire LR, Wixted JT, Clark RE (2007) Recognition memory and the medial temporal lobe: A new perspective. *Nat Rev Neurosci* 8:872–883.
2. Donaldson W (1996) The role of decision processes in remembering and knowing *Memory and Cognition* 24:523–533.
3. Wixted JT (2007) Dual-process theory and signal-detection theory of recognition memory. *Psychol Rev* 114:152–176.
4. Mayes AR, et al. (2004) Associative recognition in a patient with selective hippocampal lesions and relatively normal item recognition. *Hippocampus* 14:763–784.
5. Montaldi D, Spencer TJ, Roberts N, Mayes AR (2006) The neural system that mediates familiarity memory. Hippocampus 16:504–520.
6. Mandler G (1980) Recognizing—The judgment of previous occurrence. *Psychol Rev* 87:252–271.
7. Rugg MD, Yonelinas AP (2003) Human recognition memory: A cognitive neuroscience perspective. *Trends in Cognitive Science* 7:313–319.
8. Aggleton JP, Brown MW (1999) Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behav Brain Sci* 22:425–444; discussion 444–489.
9. Bowles B, et al. (2007) Impaired familiarity with preserved recollection after anterior temporal-lobe resection that spares the hippocampus. *Proc Natl Acad Sci USA* 104:16382–16387.
10. Yonelinas AP, Otten LJ, Shaw KN, Rugg MD (2005) Separating the brain regions involved in recollection and familiarity in recognition memory. *J Neurosci* 25:3002–3008.
11. Manns JR, Hopkins RO, Reed JM, Kitchener EG, Squire LR (2003) Recognition memory and the human hippocampus. *Neuron* 37:171–180.
12. McMackin D, Cockburn J, Anslow P, Gaffan D (1995) Correlation of fornix damage with memory impairment in six cases of colloid cyst removal. *Acta Neurochir* 135:12–18.
13. Tsivilis D, et al. (2008) A disproportionate role for the fornix and mammillary bodies in recall versus recognition memory. *Nature Neuroscience* 11:834–842.
14. Mayes AR, Montaldi D (1997) The value of neuroradiological approaches in the study of organic amnesia in *Case Studies in the Neuropsychology of Memory* Parkin AJ, ed (Erlbaum Taylor and Francis, Hove, United Kingdom).
15. Hodges JR, Carpenter K (1991) Anterograde amnesia with fornix damage following removal of IIIrd ventricle colloid cyst. *J Neurol Neurosurg Psychiatry* 54:633–638.
16. Aggleton JP, et al. (2000) Differential cognitive effects of colloid cysts in the third ventricle that spare or compromise the fornix. *Brain* 123 (Pt 4):800–815.
17. Baddeley A, Emslie H, Nimmo-Smith I (1994) *The Doors and People Test: A Test of Visual and Verbal Recall and Recognition.* (Thames Valley Test Company, Bury St Edmonds, United Kingdom).
18. Kopelman MD, et al. (2007) Recall and recognition memory in amnesia: Patients with hippocampal, medial temporal, temporal lobe or frontal pathology. *Neuropsychologia* 45:1232–1246.
19. Yonelinas AP, et al. (2002) Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nature Neuroscience* 5:1236–1241.
20. Carlesimo GA, et al. (2007) Bilateral damage to the mammillo-thalamic tract impairs recollection but not familiarity in the recognition process: A single case investigation. *Neuropsychologia* 45:2467–2479.
21. Denby CE, et al. The frequency and extent of mammillary body atrophy associated with surgical removal of a colloid cyst. *Am J Neuroradiol*, in press.
22. Denby CE, et al. (2008) MRI measurement of fornix pathology: Evidence of extensive fornix damage following surgical removal of colloid cysts in the third ventricle *Neurosci Imag* 2:109–126.
23. Yonelinas AP, Kroll NE, Dobbins I, Lazzara M, Knight RT (1998) Recollection and familiarity deficits in amnesia: Convergence of remember-know, process dissociation, and receiver operating characteristic data. *Neuropsychology* 12:323–339.
24. Quamme JR, Yonelinas AP, Widaman KF, Kroll NE, Sauve MJ (2004) Recall and recognition in mild hypoxia: Using covariance structural modeling to test competing theories of explicit memory. *Neuropsychologia* 42:672–691.
25. Yonelinas AP, et al. (2007) Memory in the aging brain: Doubly dissociating the contribution of the hippocampus and entorhinal cortex. *Hippocampus* 17:1134–1140.
26. Browne MW, Cudeck R (1993) Alternative ways of assessing model fit in *Testing Structural Equation Models*. Bollen KA, Long JS, ed. (Sage, Beverly Hills, CA), pp. 136–162.
27. Bozdogan H (1987) Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika* 52:345–370.
28. Manns JR, Squire LR (1999) Impaired recognition memory on the Doors and People Test after damage limited to the hippocampal region. *Hippocampus* 9:495–499.
29. Aggleton JP, et al. (2005) Sparing of the familiarity component of recognition memory in a patient with hippocampal pathology. *Neuropsychologia* 43:1810–1823.
30. 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* 12:325–340.
31. Dusoir H, Kapur N, Byrnes DP, McKinstry S, Hoare RD (1990) The role of diencephalic pathology in human memory disorder. Evidence from a penetrating paranasal brain injury *Brain* 113:1695–1706.
32. Vann SD, Aggleton JP (2003) Evidence of a spatial encoding deficit in rats with lesions of the mammillary bodies or mammillothalamic tract. *J Neurosci* 23:3506–3514.
33. Yonelinas AP, et al. (2004) Mild hypoxia disrupts recollection, not familiarity. *Cognitive, Affective, and Behavioral Neuroscience* 4:393–400; discussion 401–406.
34. Vann SD, et al. (2008) Memory loss resulting from fornix and septal damage: Impaired supra-span recall but preserved recognition over a 24-hour delay. *Neuropsychology* 22:658–668.
35. Wechsler D (1997) *Wechsler Adult Intelligence Scale – Third Edition* (The Psychological Corporation, San Antonio, TX).
36. Wechsler D (1997) *Wechsler Memory Scale – Third Edition* (The Psychological Corporation, San Antonio, TX).
37. Wechsler D (1999) *Wechsler Abbreviated Scale of Intelligence* (The Psychological Corporation, San Antonio, TX).
38. Winer BJ (1971) *Statistical Principles in Experimental Design* (McGraw-Hill, New York).
39. Howell D (1987) *Statistical Methods for Psychology* (Duxberry Press, Belmont, CA).