An exploratory study of the relationship between face recognition memory and the volume of medial temporal lobe structures in healthy young males

Clare E. Mackay[a], Neil Roberts[a,*], Andrew R. Mayes[b], John J. Downes[c], Jonathan K. Foster[d] and David Mann[e]
[a]*Magnetic Resonance and Image Analysis Research Centre, University of Liverpool, UK*
[b]*Department of Clinical Neurology, University of Sheffield, UK*
[c]*Department of Psychology, University of Liverpool, UK*
[d]*Department of Psychology, University of Manchester, UK*
[e]*Department of Neuropathology, University of Manchester, UK*

Received 29 October 1997
Revised 19 February 1998

A rigorous new methodology was applied to the study of structure function relationships in the living human brain. Face recognition memory (FRM) and other cognitive measures were made in 29 healthy young male subjects (mean age = 21.7 years) and related to volumetric measurements of their cerebral hemispheres and of structures in their medial temporal lobes, obtained using the Cavalieri method in combination with high resolution Magnetic Resonance Imaging (MRI). Greatest proportional variability in volumes was found for the lateral ventricles (57%) and least for the cerebral hemispheres (8%). No significant difference was observed in the mean volumes of the hippocampus, parahippocampal gyrus, amygdala, caudate nucleus, temporal pole and temporal lobe on the right and left sides of the brain. The volumes of the right and left parahippocampal gyrus, temporal pole, temporal lobe, and left hippocampus were, prior to application of the Bonferroni correction to take account of 12 multiple comparisons, significantly correlated with the volume of the corresponding hemisphere ($p < 0.05$). The volumes of all structures were highly correlated ($p < 0.0002$ for all comparisons) between the two cerebral hemispheres. There were no positive relationships between structure volumes and FRM score. However, the volume of the right amygdala was, prior to application of the Bonferroni correction to take account of 38 multiple comparisons, found to be significantly smaller in the five most consistent high scorers compared to the five most consistent low scorers ($t = 2.77$, $p = 0.025$). The implications for possible relationships between healthy medial temporal lobe structures and memory are discussed.

*Corresponding author: Neil Roberts, Magnetic Resonance and Image Analysis Research Centre, University of Liverpool, P.O. Box 147, Liverpool, L69 3BX, UK. Tel.: +44 151 794 5629; Fax: +44 151 794 5635; E-mail: neil@liverpool.ac.uk.

Behavioural Neurology 11 (1998) 3–20
ISSN 0953-4180 / $8.00

1. Introduction

The present study is the first to use unbiased stereological techniques with associated error prediction formulae to explore the possibility of a relationship between the volume of brain structures and the efficiency with which they perform their putative functions in young male subjects with healthy brains. The first aim of the study was to explore whether a positive relationship exists between face recognition memory (FRM) and the volume of the hippocampus, parahippocampal gyrus and amygdala. In particular, it can be inferred from the work of Sergent et al. [38], Sabbah et al. [34], Grady et al. [17] and Young et al. [41] that FRM is likely to be better in healthy young subjects in whom the parahippocampal gyrus (and possibly also the hippocampus) on the right and the amygdala in both hemispheres are more efficient, and therefore perhaps larger. We tested this hypothesis using a combination of psychometric testing, high resolution MRI and the Cavalieri volume estimation method. A second aim, which was specifically neuroanatomical, was to determine how much structures such as the hippocampus vary in volume in healthy young people with no history of brain damage, and whether the volumes of structures in the medial temporal lobe correlate with each other and with the volume of the whole cerebral hemisphere.

Unilateral brain damage has frequently been found to produce material specific psychological deficits and has indicated that the right and left cerebral hemi-

spheres are respectively more important for handling verbal and non-verbal information (see [20] for a review). The way in which functions are localised to particular structures, or groups of structures, within each hemisphere is being addressed by MRI. There are two main approaches. The first of these is functional MRI (fMRI), which, like Positron Emission Tomography (PET), has been used to locate the regions of the brain activated when psychological processes (for example, memory, speech or voluntary movement) are carried out, by measuring an indirect correlate of neuronal activity such as blood oxygenation or blood flow. These techniques can be used to describe inherent biological variation in neuronal activation among healthy individuals when performing particular cognitive operations and also the changes that occur as a result of brain damage. The mechanism thought to be responsible for the subtle changes of typically between 2 and 5% in signal intensity between baseline and activation images in fMRI is an increase in local susceptibility resulting from an inflow of oxygenated blood to the region of the brain being used.

The other approach to investigating the localisation of brain function involves structural imaging of brain regions that are believed to mediate specific psychological processes. MRI as applied to healthy subjects reveals considerable variations in the sizes of brain structures. This suggests that MRI can be used to examine whether there is a positive relationship between the size of brain structures and the efficiency with which they perform their hypothesised psychological functions. Such a 'neophrenological' approach rests on three kinds of assumption. First, it assumes that the psychological functions of different brain regions can be identified by studying the effects of brain lesions in patients and by functional brain imaging of healthy subjects who engage the same processes. There is massive general support for the validity of this assumption. Second, it assumes that, in general, the efficiency with which a brain region performs its hypothesised operations is a function of the number of neurons which it comprises and the complexity of their synaptic interconnections. Third, it assumes that MRI-derived measures of the volume of different brain structures are at least partially determined by the number and size (and hence of the complexity of the synaptic connections) of the neurons in each structure so that greater volumes should mean that a structure works more efficiently. This in turn means that individuals with larger brain regions should perform the functions mediated by the regions better. There is some support for this third assumption that can be drawn from several MRI studies of pathological conditions. For example, the volume of the hippocampus in temporal lobe epilepsy has been found to be proportional to the number of neurons that it contains [9, 23, 24]. The major question, therefore, relates to the second assumption that there is a positive relationship between structural size and efficiency. This was, of course, the central assumption of the phrenologists.

Whereas fMRI necessitated the development of sophisticated image analysis techniques to co-register and identify changes in signal intensity between baseline and activation images, testing the central assumption of the phrenologists only depends on the development of appropriate sampling techniques for unbiased and efficient volume estimation. In this paper, we use the Cavalieri method of modern design stereology for this purpose. Structure volume is estimated without bias as the sum of the area of the transects through the structure on consecutive systematic sections multiplied by the distance between sections. Generally, no more than 5–10 sections need to be analysed to achieve a Coefficient of Error (CE) of 5% on the Cavalieri estimate of volume (see [31, 33]).

Several studies have attempted to examine the central assumption that underlies the neophrenological approach considered here by relating the level of functions such as intelligence, musical ability and memory performance in groups of healthy subjects to the volume of the cerebral hemispheres and to the volume of specific structures within them. None of these studies employed stereological methods. Willerman et al. [42] studied a group of undergraduate students and reported that IQ score was proportional to total brain volume in males. Somewhat similar results have been reported by Andreasen et al. [3], who found correlations between intelligence and not only whole brain volume, but also the volume of several specific structures, such as the hippocampus. It is unclear whether the correlations with specific structures would have disappeared if whole brain volume (or hemisphere volume) had been partialled out. Schlaug et al. [36] have investigated the possibility of a correlation between subjects' musical ability and the size of their planums temporale (which includes the auditory association cortex and hence should be concerned with the processing of music and language). The planum temporale was larger on the left in those with perfect pitch suggesting that an ability to interpret music as a language, rather than a propensity for handling non-verbal material, underlies this ability.

The medial temporal lobes have been implicated as important for memory since it was observed that the temporal lobe epileptic, H.M., became amnesic after a bilateral temporal lobectomy [37], but the roles of specific structures within this region remain a source of contention. Early observations using animals suggested that the hippocampus and amygdala were of primary importance since lesions that included both caused more profound amnesia than lesions that included either individually [28]. However, these lesions also included the surrounding perirhinal and parahippocampal cortices [44], damage to which was subsequently shown to cause severe memory impairments [8]. The hippocampus certainly seems to be involved with explicit memory in humans as is implied by the positive correlation between its volume and memory performance, which has been found in several studies of patient groups in whom this structure is likely to be damaged. These groups include schizophrenics [29], patients with Alzheimer's disease [12, 14, 22] and patients with temporal lobe epilepsy [24, 35].

Recent studies also suggest that although amygdala lesions do not contribute to global amnesia, they may impair non-verbal visual recall and recognition, and, in particular, the recognition of faces and facial information processing [1, 41]. Unfamiliar faces are difficult to verbalise and are therefore considered to be non-verbal stimuli. Memories of them are likely to be mediated via right temporal lobe structures. However, Young et al. [41] reported a patient in whom MRI revealed focal atrophy of the amygdala with a more marked reduction on the left side than the right. This patient had impaired FRM as well as poor ability to read emotions from facial expressions. It would seem that the amygdala is bilaterally concerned with visual explicit memory.

In a PET study of the brain activation produced by different face processing tasks, Sergent et al. [38] compared a gender judgement task with a face identity judgement task. The identity task produced activation anterior to that produced by the gender judgement task. This activity was found in structures in both hemispheres such as the fusiform gyrus and temporal pole, but was stronger in the right hemisphere. The greatest level of activation was found in the right parahippocampal gyrus with no activation being detected in the left parahippocampal gyrus. Sabbah et al. [34] have replicated the study of Sergent and her colleagues using fMRI and found essentially the same results except that right hippocampal activation was also produced. Another PET study [17], which explicitly examined the encoding and retrieval of face stimuli, also found that the right hippocampus as well as the right parahippocampal cortex was activated, but only during encoding of face stimuli. These PET and fMRI studies in conjunction with lesion studies, therefore, suggest that FRM may depend on activity in the right parahippocampal gyrus and the amygdala bilaterally. There are, however, inconsistencies between the PET studies and the lesion studies regarding the extent to which the right hippocampus is involved in face recognition. It was intended that the study reported here would cast further light on this issue.

2. Method

2.1. Subjects

The study was run in two stages that were a year apart, each of which had three phases. Each stage included 100 healthy male undergraduates between the ages of 18 and 25 (mean $= 21.7$). In phase one of both stages, preliminary memory tests were given in order to select the subjects with best and worst FRM. In phase two of both stages, which was given one to two months later, further tests were given to the selected extreme scoring subjects for several reasons, one of which was to see how consistent their FRM performance was across time. Finally, in phase three which was given around the time of phase two, the selected extreme scorers on FRM were given an MRI scan. There were some differences in the tests given in the two stages (designed to reduce the ceiling effects found in stage 1) so they will be described and analysed separately. The handedness of the subjects who were selected for MRI scanning was ascertained using the Edinburgh Handedness Inventory (EHI) (short form) which is a 10 item questionnaire giving a laterality quotient percentage from -100 (left handed) to $+100$ (right handed). Twenty-six of the 29 scanned subjects completed the questionnaire. Four scored in the negative range indicating left handedness.

2.2. Materials

In phase one of stage one, the face version of the Warrington Recognition Memory Test (RMT) was used. This is a two-choice forced-choice face recognition test. However, the good subjects scored too near to ceiling levels on this 50 item forced choice test so

for phase two a 100 item FRM test was used. This was constructed from the faces of undergraduate students (FRM Test A). This test was also used in phase one of stage two. In phase two of stage two, a second 100 item FRM test (FRM Test B) was used. This test was constructed from the faces of actors and actresses who were judged likely to be unknown to the general public. Both FRM Test A and FRM Test B were constructed from close-up, full face, black and white photographs of males and females. In each test pair, the face of the foil was matched to the target in gender, age and broadly in terms of features. The study faces were mounted on individual cards and the test faces were mounted in pairs on cards with the right/left position of targets being counterbalanced.

Subjects in stage one, phase one were also given the word version of the RMT (50 items) and those in stage two, phase one were given a 150 item two-choice forced choice word recognition test. This test comprised common, concrete target and foil words, and like the face recognition tests, study words were mounted on individual cards and test cards each contained a target and a foil word with target right/left position being counterbalanced. The test was developed because it was clear from the results of stage one that the subjects were scoring at ceiling levels on the word version of the RMT. It was also found that even with a 100 item test, subjects were still performing around ceiling levels, which was the reason why 150 items were used.

Subjects in phase one of both stages were also given a 24 item three-choice forced choice recognition test for random visual patterns. This is referred to as the wallpaper pattern test as the test materials are black and white versions of wallpaper patterns. Test order in all these tests was fixed and different from the study presentation order.

In phase two of both stages, the Benton Face Processing Test [6] was given in order to assess subjects ability to process face material. This test requires subjects to decide whether different face views are of the same or different people. The Gestalt Picture Fragment Test was also given in phase two of both stages in order to assess the ability to interpret visual materials. This test requires subjects to identify familiar objects from incomplete pictures.

2.3. Neuropsychology procedure

In stage one, phase one, 100 male undergraduates were selected on the basis of being in their 20s, having no history of head injury or illness that could have caused brain damage, and willingness to undergo an MRI scan if asked to do so. Each subject was tested in a quiet room on the word and face tests from the RMT and on the wallpaper recognition test. With all tests there was no delay between study and test, and exposure during the study phase to each item was three seconds. Scores on the face version of the RMT were used to select the best and worst subjects on face recognition memory for phase two, which was run about two months later immediately before phase three. Six subjects were selected for this phase on the basis of having scored 34 or less on the face version of the RMT and nine subjects were selected on the basis of scoring 49 or more. These subjects were then given FRM Test A, the Benton Face Processing Test and the Gestalt Picture Fragment Test. The study time was three seconds per item and the delay between the end of the study phase and the test phase was 15 minutes. Five subjects with poor face recognition memory and six with good face recognition memory were scanned in phase three. The other four subjects were unfortunately unavailable for scanning in the time available.

In stage two, phase one, 100 male undergraduates were selected according to the same criteria as in stage one. Subjects were tested individually in a quiet room on the 150 word recognition test with study exposures of one second per word, FRM Test A with study exposures of three seconds per face, and the wallpaper recognition test with study exposures of 1.5 seconds per pattern. There was a delay of about 15 minutes between the end of each study phase and the beginning of testing. Thirty six subjects were selected for phase two, which was begun about two months later, on the basis of having the best and worst scores on the FRM test. More subjects were included in phase two because we wished to use stricter selection criteria for phase three than was possible in stage one. In phase two, the selected subjects were given FRM Test B, the Benton Face Processing Test, and the Gestalt Picture Fragment Test. There was a filled delay of 30 minutes between the end of the study phase and the beginning of the test phase of the face recognition test. Nine low scoring and 11 high scoring subjects in this test were selected for the third phase on the basis of their having similar scores on the Benton and Gestalt tests so that it could reasonably be argued that the high level visual processing abilities of the good and poor face recognizers were matched. Two of the 20 MRI scans from phase three were not usable because they were severely

degraded due to subject movement. This meant that there was a total of 29 usable scans from both stages of the study.

2.4. MR image acquisition and analysis procedure

The brains of the twenty-nine subjects from both stages of the experiment, who had been selected to have similar face (and high level visual stimulus) processing ability and yet differing face recognition memory, were imaged using a 1.5 T SIGNA whole body MR imaging system (General Electric, Milwaukee, USA). One hundred and twenty four coronal T1-weighted images were obtained using a 3D spoiled gradient echo (SPGR) pulse sequence (TR of 34 ms, TE of 9 ms and flip angle of 45°). The two, as opposed to one, NEX acquisition, which took 27 minutes and 52 seconds, gave increased contrast between the grey and white matter, and therefore more ready definition of structure boundaries. The Field of View (FOV) of the images was 20 cm, and each image refers to a contiguous section of tissue of 1.6 mm thickness. The MR images showed no evidence of movement or chemical shift artefacts, and partial voluming effects were minimal.

The acquired images were transferred to ANALYZE (MAYO Foundation, Minnesota, USA) software running on a SPARC 10 workstation (SUN Microsystems, CA, USA). The $256 \times 256 \times 124$ acquired voxels of side 0.78 mm $\times$ 0.78 mm $\times$ 1.6 mm were linearly interpolated to $256 \times 256 \times 254$ cubic voxels of side 0.78 mm. The left and right hippocampus, parahippocampal gyrus, amygdala and temporal pole are optimally visualised and their volumes best measured on image sections oriented perpendicular to the long axis of the hippocampus [4, 23]. These sections were conveniently obtained by reformatting oblique sections through the cubic voxel data within ANALYZE (MAYO Foundation, Minnesota, USA) software. The four main steps of the reformatting procedure are illustrated in Fig. 1. Firstly, the images were reformatted sagitally and the image in which the long axis of the left hippocampus was most clearly visualised selected (Fig. 1a). A line was drawn perpendicular to the long axis of the hippocampus and the entire data set reformatted so that the brain was exhaustively sectioned parallel to this direction. Next, the direction of reformatting was adjusted (Fig. 1b) slightly so that the brain appeared symmetrical on the reformatted sections (Fig. 1c). The resulting 0.78 mmthick contiguous coronal sections lie approximately perpendicular to the mean direction of the right and left temporal lobe structures. This was also a convenient sectioning direction for measurement of right and left cerebral hemisphere volume and the volume of the lateral ventricles, and a suitable sectioning direction for estimating the volume of the caudate nucleus, which served as a control structure not thought to play a critical role in memory processes.

The procedures for identifying the boundaries of structures on the reformatted MR images were established by a neuroanatomist (DM) and through the use of a neuroanatomy atlas [30]. The SPGR sequence provides optimum contrast between tissues on the basis of the value of their T1 relaxation times. The shorter the T1 the higher the signal intensity so that white matter appears brighter than grey matter, and CSF appears black. The structure definitions used are given below and an illustration of the application of the stereological point counting technique to sections through the hippocampus, parahippocampal gyrus, amygdala and caudate nucleus is illustrated in Fig. 2.

hippocampus – anteriorly the boundary was the first slice on which the hippocampus could be differentiated from the amygdala by visualisation of the alveus and/or CSF between the structures. The posterior boundary is reached when the lateral ventricles divide into the frontal and temporal horns.

parahippocampal gyrus – this is the first gyrus inferior to the hippocampus bounded by an imaginary line drawn between the temporal horn and the end of the sulcus. The posterior and anterior limits were identical for those used for the hippocampus.

amygdala – anteriorly measurement stopped at the last section on which its boundary could be clearly identified in respect of the adjacent white matter of the temporal lobe. The amygdala lies anterior and superior to the neighbouring hippocampus, separated by the alveus and typically additionally by a region of CSF superior to the alveus on the most posterior sections containing both structures. The posterior limit was the last slice on which the grey matter superior to the hippocampus can be distinguished.

caudate nucleus – the head of this grey matter structure lying proximal to the left ventricle is clearly visualised along its length within surrounding white matter on the T1-weighted images.

temporal pole – the posterior boundary was the slice marking the anterior limit of the temporal pedicle.

temporal lobe – as for the hippocampus, the posterior boundary of the temporal lobe is reached when the

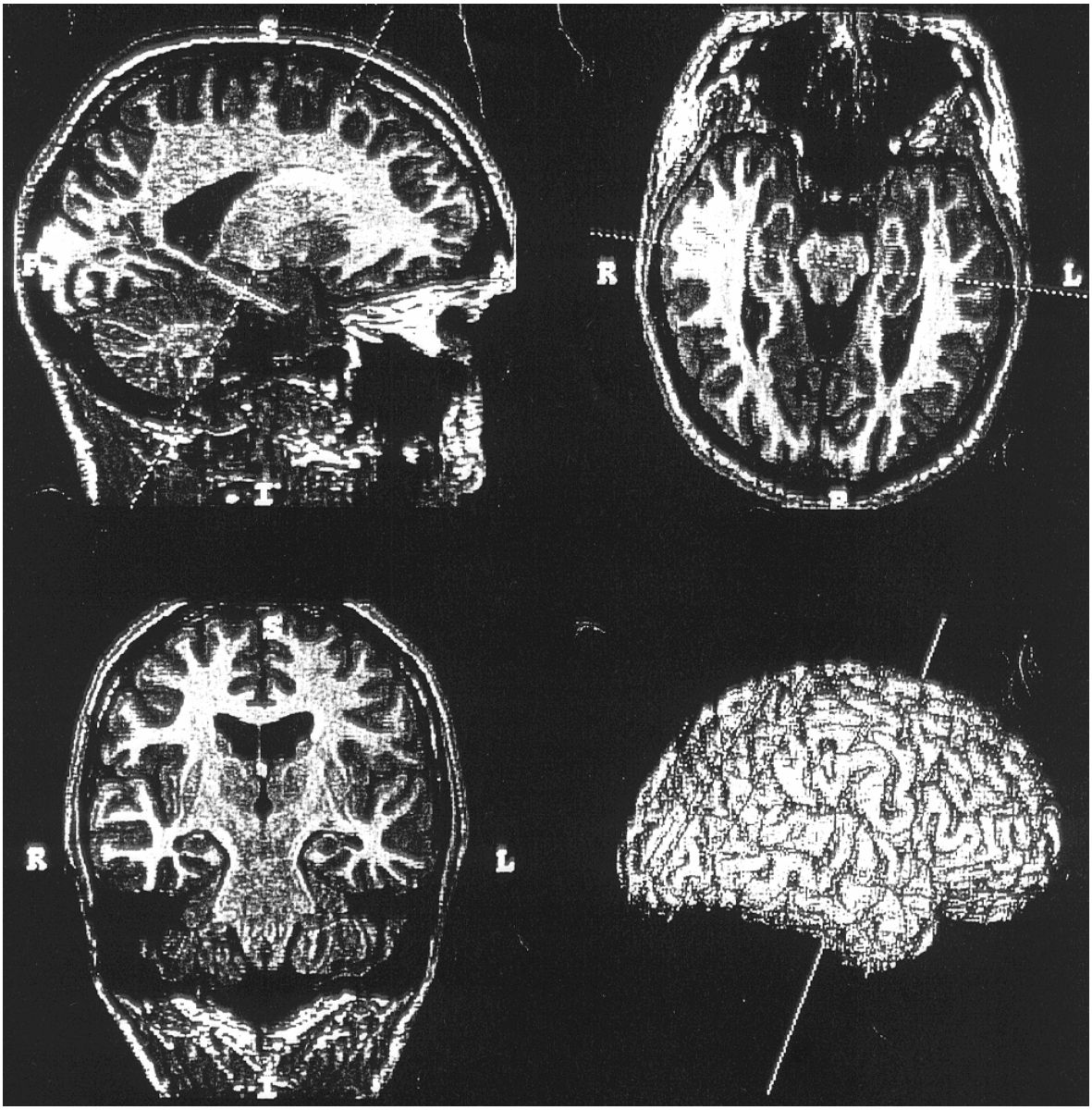

Fig. 1. For optimum visualisation of the structures of the medial temporal lobe the acquired images were reformatted perpendicular to the long axis of the hippocampus. The procedure involved (top-left) drawing a line along the long axis of the hippocampus from a sagittal slice on which this is clearly visible and (top-right) making adjustments so as to render the data symmetrical using an axial slice. A resulting coronal slice after reorientation is shown in (bottom-left), and (bottom-right) shows a 3-dimensional rendering of the brain with the plane of reorientation shown.

lateral ventricles divide into the frontal and temporal horns.

lateral ventricles – the cerebral ventricular system is well delineated as regions of very low signal intensity on T1-weighted images.

cerebral hemispheres – not including the cerebellum, were separated from the brain stem at the superior limit of the pons.

An unbiased estimate of the volume of a structure of arbitrary shape and size may be obtained efficiently and with known precision using the Cavalieri method of modern design stereology. The method requires that the structure is sectioned from end to end with a series of parallel planes a constant distance apart. Provided that the position of the first section is random within the sectioning interval an unbiased estimate of volume is obtained by multiplying the total area of the

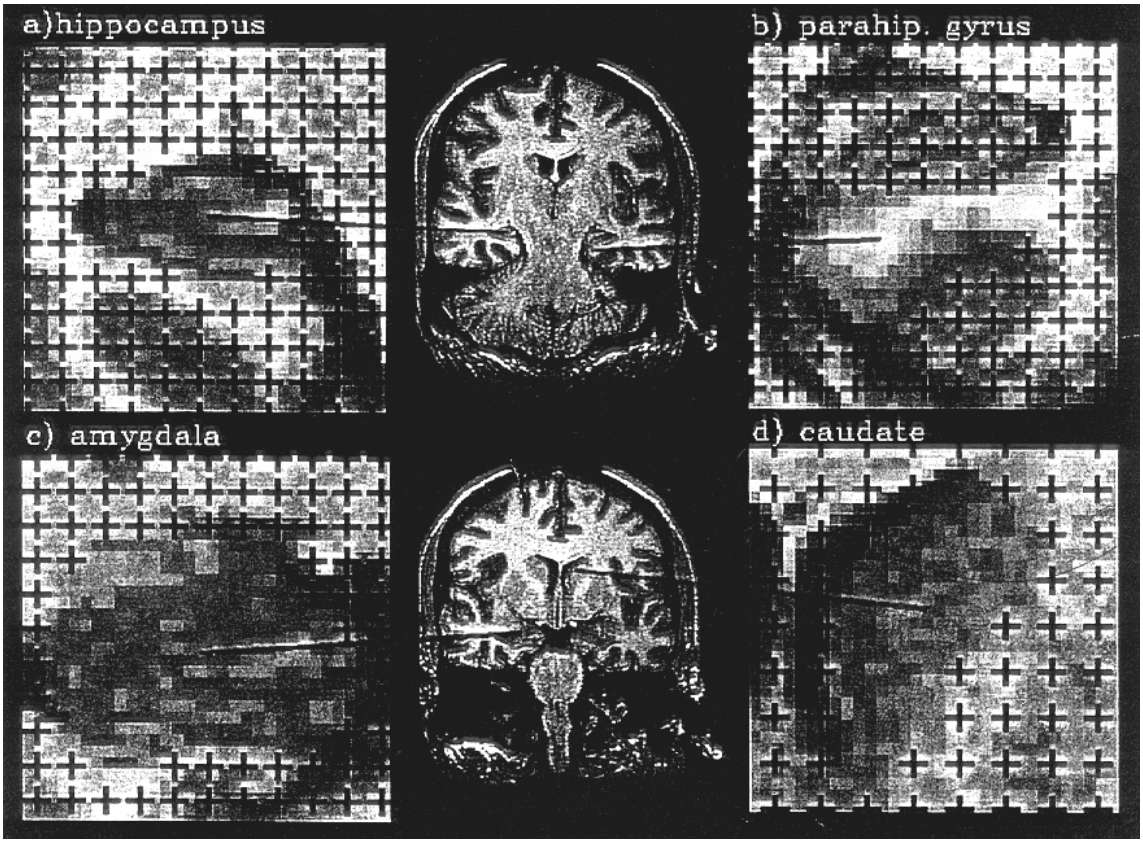

Fig. 2. MR images reformatted perpendicular to the long axis of the hippocampus (see Fig. 1) overlain by test system to illustrate estimation of the section areas of the hippocampus, parahippocampal gyrus, caudate, and amygdala by point counting. Points overlying the structure of interest have been removed.

transects through the structure on all the sections by the sectioning interval. If the section areas cannot be properly segmented and measured automatically, then the semi-automatic approach based on manual tracing of the transect boundaries on the digital images should never be adopted. Interactive point counting techniques represent a more reliable and efficient approach. In the present study point counting was carried out via stereology menus within ANALYZE (MAYO Foundation, Minnesota, USA) software.

Structure volume may be estimated either from a series of images as originally acquired or else, when a high resolution 3D data set is available, from images reformatted along an axis defined retrospectively. The later approach, which was used in the present study, offers the convenience of enabling one to define the sectioning direction on a structure by structure basis and may be especially convenient when an anatomical boundary cannot be clearly defined for the entire structure on the MR images. This applies to the temporal pole, for example.

The precision of a volume estimate obtained using the Cavalieri method may be measured by its Coefficient of Error (CE) or 'relative standard error'. Systematic sampling, whereby there is a constant interval between sections, is a more efficient sampling strategy than random sampling by a factor equal to the square root of the number of sections; for example, only ten systematic sections are required to give the same precision as 100 random sections. However, prediction of the CE for systematic sections is not straightforward. The transect areas on successive sections are not independent quantities and so conventional statistical formulae cannot be applied. A formula for predicting the CE, which involves assessment of the correlational structure of the data, was first developed by Gundersen and Jensen [18], based on the theory of Matheron [26]. More recently, Cruz-Orive [11] has developed a formula to take full account of the situation where the section areas have been estimated by point counting rather than measured exactly and it is this approach that has been used in the present study. Inter-rate re-

liability was addressed by two observers. Intra-class correlation coefficients for the hippocampus, parahippocampal gyrus, amygdala, caudate nucleus, temporal lobe, lateral ventricle and whole cerebral hemisphere are 0.97, 0.98, 0.99, 0.84, 0.96, 1.0 and 0.94 when each observer obtained measurements using the same image sections and grid positions, and 0.91, 0.91, 0.91, 0.84, 0.91, 0.97 and 0.87 when each observer obtained measurements from a different systematic random series of sections and with a new random throw of the test system.

Besides knowing the volume of a structure and its predicted uncertainty for an individual it is also of interest to consider the variation in structure volumes between individuals. Gundersen and Osterby [19] have shown that the square of the Coefficient of Variation (CV) on the estimate of mean volume for a group of subjects is equal to the square of the inherent biological CV in the volume of the structure among the group members plus the mean of the squares of the predicted CE's on the individual volume estimates. Consequently, since the former quantity is measured and the CE's on the individual estimates could be predicted, the inherent biological variation in the volumes of the right and left hemisphere brain structures could be determined for the 29 subjects analysed in the present study.

3. Results

We will report the results in two steps. First, we will consider the MRI data that relate to the degree of individual variability in structure volumes and whether the volumes of specific structures, such as the hippocampus, correlate with each other and with the volume of the whole hemisphere. The data from all 29 subjects are considered together for this analysis. Second, we will report the relationship between the neuropsychological measures and the volume measures.

The volumetric data for the 29 subjects are shown in Table 1. For each subject it took about one hour to obtain unbiased estimates of the volumes of all the listed structures with predicted CE's of between 3 and 6%. We wished to compare the biological variability of the estimated structure volumes in order to assess to what extent this differed across structures. In order to do this it was necessary to do two things. First, one must express the variability in volume as a proportion of the mean. Second, one must separate the biological variability from the total variability by eliminating the component due to sampling error on the individual volume estimates. Both these aims are achieved using the approach described by Gundersen and Osterby [19]. The results obtained are shown in Table 2. These data indicate that there was considerable variability in the volumes of the structures that were measured in order to test our neuropsychological hypotheses. The ventricular volumes were the most variable, and the hemisphere and temporal lobe volumes were the least variable. This has clinical relevance since it supports the view that there is great variability in the ventricular volumes of normal subjects, which justifies the reluctance of clinicians to interpret the observation of large ventricular volumes in a patient as a definitive marker for pathology. All other structures varied to a similar degree and only slightly more than the hemispheres. There was also a weak indication that the hippocampal volumes varied more on the right than the left, whereas all other structures varied slightly more on the left. This may relate to the fact that we chose this population on the basis of varied face recognition memory (which may be more right hemisphere related) but limited variability in verbal memory and processing ability (which may be more left hemisphere related). The data indicate that there is adequate variability in the volumes of structures of interest in order for our study to provide a fair test of the neuropsychological hypotheses about face recognition memory outlined in the Introduction.

There was no significant difference in the mean volumes of the hippocampus, parahippocampal gyrus, amygdala, caudate nucleus, temporal pole and temporal lobe on the right and left sides of the brain. However, the observation that the left hemisphere was on average 2.7% greater than the right was significant ($t = 2.46$, $p = 0.017$). Fig. 3 shows the relationship between the volumes of all the individual structures and the cerebral hemispheres computed separately for the right and left sides of the brain. Pearson's product moment correlation coefficient was calculated to test the significance of each correlation. The volumes of both the right and left parahippocampal gyrus, temporal pole and temporal lobe correlated significantly with the volume of the corresponding hemisphere ($p < 0.05$). The volume of the hippocampus also tended to correlate with hemisphere volume. This was significant for the left hippocampus ($p = 0.019$) whereas there was only a tendency for the right hippocampus to correlate with hemisphere volume ($p = 0.110$). The volumes of the amygdala, the caudate nucleus and the lateral ventricle were not

Table 1
Estimated volumes (ml) of the hippocampus, parahippocampal gyrus, amygdala, temporal pole, temporal lobe, lateral ventricles, and caudate nucleus in the left and right cerebral hemispheres together with hemisphere volume for the 29 male subjects[a]

Subject	Hippocampus		Parahippocampal gyrus		Amygdala		Temporal pole	
	L	R	L	R	L	R	L	R
01	2.8 (3.1)	3.2 (3.5)	3.7 (6.8)	3.6 (6.0)	2.6 (4.3)	2.9 (3.7)	20.2 (6.4)	16.9 (8.8)
02	2.9 (3.6)	3.1 (3.4)	2.5 (8.0)	2.4 (7.6)	2.1 (5.2)	2.0 (5.4)	18.9 (7.1)	21.5 (6.9)
03	2.6 (3.3)	2.4 (3.3)	2.9 (8.5)	2.9 (7.6)	2.5 (4.0)	2.6 (3.6)	24.9 (5.5)	20.6 (5.8)
04	2.9 (3.0)	3.0 (2.9)	3.0 (7.0)	3.5 (7.0)	2.5 (3.3)	2.4 (3.2)	20.2 (6.2)	21.1 (6.2)
05	1.9 (4.9)	2.2 (5.1)	4.2 (7.0)	3.6 (7.9)	2.5 (4.0)	2.7 (3.2)	22.5 (6.7)	26.7 (6.3)
06	2.7 (4.4)	3.2 (3.2)	3.4 (6.4)	3.7 (6.5)	2.4 (4.0)	2.5 (4.0)	21.1 (5.8)	18.7 (7.4)
07	3.1 (3.5)	3.5 (2.6)	3.1 (6.2)	3.1 (5.9)	2.3 (3.5)	2.6 (3.4)	20.9 (6.0)	22.7 (6.4)
08	2.2 (4.6)	2.0 (5.4)	2.7 (7.8)	2.9 (8.9)	2.3 (4.0)	2.4 (4.3)	16.8 (6.9)	11.2 (9.4)
09	2.1 (4.2)	2.8 (3.8)	2.5 (7.5)	2.6 (7.5)	3.2 (3.3)	3.5 (3.2)	18.7 (6.3)	20.0 (5.4)
10	2.2 (4.0)	2.5 (3.3)	3.9 (4.4)	4.0 (5.4)	2.8 (2.9)	2.6 (2.5)	22.1 (6.5)	23.4 (5.9)
11	3.0 (4.2)	3.0 (2.8)	2.4 (8.5)	2.9 (6.8)	3.0 (2.7)	2.8 (3.4)	16.8 (7.6)	17.0 (7.5)
12	2.1 (5.2)	2.5 (2.9)	3.3 (6.6)	2.9 (7.6)	2.7 (3.8)	2.6 (4.6)	16.2 (7.1)	16.8 (6.2)
13	2.9 (3.9)	3.3 (3.1)	2.8 (7.1)	3.0 (8.0)	2.7 (2.7)	2.3 (3.2)	22.5 (6.2)	18.1 (7.0)
14	3.0 (3.2)	2.7 (3.6)	2.8 (5.9)	3.6 (5.2)	2.7 (4.0)	2.8 (4.6)	23.6 (5.9)	22.1 (6.1)
15	2.4 (5.0)	3.1 (3.3)	3.3 (6.8)	3.4 (6.2)	2.3 (3.0)	2.3 (3.5)	22.1 (5.4)	20.9 (6.6)
16	2.4 (4.4)	2.3 (5.2)	3.6 (8.3)	3.8 (5.9)	2.7 (5.0)	2.8 (4.0)	18.1 (6.6)	16.8 (8.0)
17	3.0 (2.6)	3.2 (2.8)	3.5 (7.0)	3.4 (7.5)	2.1 (3.6)	2.3 (4.4)	18.5 (6.7)	17.1 (6.5)
18	2.5 (3.1)	2.8 (3.7)	3.0 (7.8)	2.9 (9.4)	2.2 (5.1)	2.3 (4.6)	21.5 (6.5)	23.2 (5.6)
19	2.6 (3.9)	2.8 (3.7)	3.9 (7.4)	3.3 (6.0)	2.9 (3.6)	3.0 (3.5)	19.8 (6.5)	21.0 (6.0)
20	3.1 (3.1)	3.1 (3.5)	4.0 (7.2)	3.6 (6.3)	2.2 (4.0)	2.6 (2.9)	19.8 (6.5)	20.2 (6.4)
21	2.4 (3.9)	2.7 (2.6)	2.4 (7.4)	1.9 (9.8)	2.7 (3.6)	2.4 (3.5)	17.1 (6.9)	18.1 (7.1)
22	2.6 (3.5)	2.7 (4.4)	3.0 (6.9)	3.1 (7.8)	2.0 (3.3)	2.3 (4.4)	16.6 (7.8)	13.2 (6.7)
23	2.4 (3.6)	2.4 (3.7)	3.0 (7.4)	3.4 (6.5)	3.1 (2.5)	3.1 (3.6)	22.1 (5.7)	20.0 (6.1)
24	2.8 (4.7)	2.7 (3.8)	2.5 (7.3)	2.6 (7.3)	2.3 (3.9)	2.7 (3.3)	17.3 (6.3)	21.6 (6.1)
25	2.2 (4.8)	2.5 (3.8)	3.7 (6.3)	4.2 (6.8)	2.4 (4.1)	2.6 (3.7)	21.3 (6.2)	19.0 (7.1)
26	2.5 (3.7)	2.6 (3.5)	3.9 (6.3)	3.7 (6.9)	2.5 (3.7)	2.7 (3.7)	24.4 (4.5)	24.6 (5.4)
27	2.8 (3.8)	2.9 (2.9)	3.5 (6.3)	2.8 (7.8)	3.0 (2.7)	2.8 (4.2)	17.9 (6.8)	16.6 (7.0)
28	2.5 (3.3)	2.3 (3.6)	3.9 (5.8)	3.5 (4.9)	2.5 (3.6)	2.6 (3.8)	20.2 (6.0)	23.2 (6.6)
29	3.3 (3.2)	3.7 (3.4)	3.5 (9.7)	3.3 (7.2)	3.1 (2.8)	2.9 (3.1)	23.0 (5.9)	21.7 (5.7)
max	3.32	3.74	4.23	4.15	3.22	3.54	24.95	26.70
min	1.94	2.02	2.42	1.88	2.00	1.98	16.19	11.23
mean	2.62	2.72	3.24	3.23	2.55	2.62	20.17	20.00
SD	0.35	0.41	0.53	0.50	0.33	0.30	2.49	3.09

[a]The percentage predicted CE's are shown in parentheses, and the maximum, minimum and mean volumes of each structure are also presented. *(continued)*

significantly correlated with the volume of the hemisphere. The volumes of the individual structures on the right and left sides of the brain were highly correlated ($p < 0.0002$ for all comparisons). The correlations between the volumes of the structures of the medial temporal lobes were also investigated. These data are summarised in Table 3. The volume of the right hippocampus is significantly correlated with the volume of the right temporal lobe ($p = 0.018$) which is significantly correlated with the volume of the right temporal pole ($p = 0.014$). There is no evidence that any of the above correlations were influenced by subjects' handedness because their handedness quotient did not correlate with any of the structure volumes (all comparisons: $p > 0.1$).

The neuropsychological data obtained from stages 1 and 2 of this study are presented in Table 4. Since the tests that were administered in the two stages of the study differed slightly, the results will be considered separately for stages 1 and 2.

3.1. Stage 1

The means and standard deviations of the results of the RMT 50 faces and 50 words recognition memory tests and the 24 item pattern recognition test, given in phase 1 of stage 1, are presented in Table 4. The values presented in the table refer to the eleven subjects who went on to complete the three phases of stage 1. There was a ceiling affect in the 50 words test, and a less obvious one in the 50 faces test. Pearson's product moment correlation was used to test for evidence of

Table 1 *(continued)*

Subject	Temporal lobe		Caudate nucleus		Lateral ventricles		Whole hemisphere	
	L	R	L	R	L	R	L	R
01	85.8 (3.4)	80.9 (3.2)	5.9 (8.2)	5.5 (7.1)	6.8 (3.9)	11.3 (2.6)	597.9 (3.6)	556.1 (3.9)
02	87.6 (4.1)	66.6 (3.4)	3.5 (7.2)	3.7 (7.3)	25.5 (1.4)	17.6 (1.9)	632.2 (3.4)	571.3 (3.4)
03	68.3 (4.3)	72.3 (4.1)	5.0 (6.9)	5.1 (9.0)	6.6 (3.5)	5.9 (3.9)	613.1 (3.5)	575.1 (3.3)
04	84.6 (3.5)	83.5 (3.5)	4.7 (5.4)	4.6 (6.1)	6.7 (3.6)	4.8 (4.7)	594.1 (3.3)	552.2 (3.3)
05	88.9 (3.0)	80.0 (2.9)	5.1 (6.6)	5.0 (5.2)	6.3 (4.2)	7.6 (3.8)	719.8 (3.0)	700.8 (3.1)
06	88.6 (2.8)	85.8 (3.3)	4.1 (7.1)	4.1 (8.6)	5.3 (4.5)	5.8 (3.8)	624.6 (3.6)	582.7 (3.3)
07	89.2 (3.5)	85.5 (3.4)	4.4 (6.2)	4.6 (6.0)	8.2 (3.1)	5.5 (4.7)	655.1 (3.4)	601.8 (3.4)
08	69.4 (4.3)	68.3 (4.3)	4.5 (9.9)	4.7 (7.8)	13.9 (2.6)	9.8 (3.5)	552.2 (3.5)	487.5 (4.1)
09	95.5 (4.2)	82.0 (3.5)	5.0 (7.1)	5.2 (7.4)	4.0 (5.3)	4.1 (6.6)	457.0 (4.4)	441.8 (4.2)
10	76.0 (3.5)	80.6 (3.5)	5.7 (5.6)	6.0 (7.5)	7.8 (3.5)	7.3 (4.6)	605.6 (3.8)	586.5 (3.4)
11	86.1 (3.7)	84.6 (3.5)	5.4 (6.9)	5.2 (7.9)	8.3 (3.3)	7.6 (3.4)	529.4 (3.9)	514.5 (3.5)
12	68.9 (3.6)	71.4 (4.5)	3.8 (9.6)	4.0 (8.2)	5.8 (4.5)	11.3 (2.8)	529.4 (3.6)	529.4 (3.5)
13	77.5 (3.8)	74.6 (4.0)	4.6 (9.3)	4.5 (6.5)	6.9 (3.9)	6.7 (3.7)	601.8 (3.5)	586.5 (3.4)
14	87.8 (3.8)	80.3 (3.6)	5.1 (7.2)	4.8 (6.0)	5.7 (4.1)	7.5 (3.7)	647.5 (3.3)	590.3 (3.2)
15	82.6 (3.2)	82.9 (3.5)	4.1 (6.1)	4.5 (7.9)	5.1 (5.2)	6.3 (6.7)	556.1 (4.1)	533.2 (3.9)
16	77.7 (3.3)	86.6 (3.7)	4.3 (7.9)	4.4 (7.2)	2.4 (6.6)	3.5 (7.8)	548.4 (3.5)	533.2 (3.3)
17	74.3 (3.6)	74.9 (3.4)	4.4 (7.0)	4.7 (7.9)	8.4 (3.5)	6.6 (3.4)	567.5 (3.4)	548.4 (3.4)
18	89.8 (3.1)	89.2 (3.4)	4.6 (5.5)	4.5 (5.8)	20.4 (1.6)	16.6 (2.6)	624.6 (3.7)	578.9 (3.4)
19	77.2 (3.3)	76.6 (3.7)	4.6 (6.7)	4.4 (7.6)	16.1 (2.1)	12.6 (2.6)	624.6 (3.3)	540.8 (3.4)
20	82.6 (3.5)	80.9 (3.7)	3.8 (7.7)	3.4 (6.7)	6.9 (3.5)	4.2 (4.5)	666.5 (3.4)	609.4 (3.1)
21	72.0 (3.7)	68.3 (3.6)	3.1 (8.8)	3.2 (9.4)	4.5 (5.2)	3.7 (6.8)	518.0 (3.8)	466.5 (3.7)
22	71.4 (3.4)	74.3 (3.4)	3.3 (4.9)	4.1 (6.5)	9.3 (3.2)	9.6 (2.9)	533.2 (4.2)	483.7 (3.8)
23	86.6 (3.5)	84.3 (3.4)	4.7 (7.0)	4.9 (6.5)	10.5 (3.2)	7.9 (3.3)	616.9 (3.7)	548.4 (3.4)
24	77.5 (3.4)	78.0 (3.6)	4.3 (7.3)	4.6 (7.6)	4.2 (4.7)	4.4 (5.7)	578.9 (3.7)	575.1 (3.8)
25	77.7 (3.3)	76.9 (3.7)	4.1 (8.0)	4.4 (7.0)	4.1 (4.9)	3.9 (5.3)	544.6 (4.1)	537.0 (3.7)
26	78.6 (3.3)	71.1 (3.7)	4.6 (6.8)	5.0 (7.3)	3.9 (5.6)	3.8 (4.8)	628.4 (3.4)	582.7 (3.3)
27	85.2 (3.2)	79.2 (3.4)	4.8 (8.6)	5.0 (5.7)	7.5 (3.2)	9.4 (2.8)	575.1 (3.8)	578.9 (3.4)
28	82.6 (3.6)	83.8 (3.2)	5.0 (5.0)	4.5 (6.9)	6.1 (3.5)	4.7 (5.1)	613.2 (4.1)	575.1 (3.7)
29	91.5 (3.6)	83.2 (3.1)	5.8 (5.4)	5.2 (6.5)	9.1 (3.3)	6.7 (3.4)	571.3 (3.7)	590.3 (3.7)
max	89.2	95.5	5.91	6.03	25.50	17.58	719.8	700.8
min	66.6	68.3	3.11	3.23	2.40	3.47	457.0	441.8
mean	78.9	81.4	4.57	4.64	8.15	7.47	590.6	557.2
SD	6.06	7.42	0.70	0.58	5.05	3.64	53.6	49.7

Table 2
Analysis of variance of the volumes of the brain structures for the 29 subjects in Table 1[a]

Subject	H'ocampus		P'hip. gyrus		Amygdala		Temp. pole		Caudate		Ventricles		H'sphere	
	L	R	L	R	L	R	L	R	L	R	L	R	L	R
$\%CV_T$	17.9	20.7	16.4	15.6	22.6	21.7	15.5	12.3	15.3	12.6	63.9	50.7	9.1	8.9
IQR 25%	2.96	3.19	2.83	2.94	1.75	1.66	18.09	18.09	4.14	4.39	5.30	4.69	552.2	533.2
75%	3.74	3.89	3.63	3.56	2.42	2.37	21.71	22.09	5.00	5.00	8.40	9.39	624.6	582.7
%CE	10.8	11.6	7.2	7.1	7.4	7.6	6.4	6.7	7.2	7.2	4.0	4.5	3.6	3.5
$\%CV_B$	14.3	17.1	14.8	13.9	21.4	20.3	14.1	10.3	13.5	10.3	63.8	50.5	8.4	8.2
IQR 25%	3.03	3.12	2.91	2.92	1.79	1.77	18.24	18.60	4.15	4.32	4.61	4.90	556.9	526.1
75%	3.69	3.94	3.57	3.54	2.39	2.33	22.10	21.40	4.99	4.96	11.69	10.04	624.3	588.3

[a]The total Coefficient of Variation (CV_T), expressed as a percentage, is the standard deviation amongst the 29 volume estimates divided by the mean volume. This contains a contribution from the uncertainty that is due to each volume having been estimated by a procedure which involves both sectioning and point counting (CE_T) together with a contribution due to the inherent biological variation among individuals (CV_B). A description of the relation between these quantities is given in the text. The 25% and 75% values for the interquartile volumes (ml) for both the raw data IQR(CV_T) and the biological variation IQR(CV_B) are also shown.

any relationship between performance on the different tests. This was computed for all 100 subjects who completed phase 1. The only significant correlation was between the results of the 50 faces and 24 patterns test ($r = 0.5$, $p = 0.0001$).

The results of FRM Test A and the 100 item word recognition test, the Benton Face Processing Test and the Gestalt Picture Fragment Test administered in

Table 3
Pearson's product moment correlation for a linear regression analysis between the volumes of medial temporal lobe structures in the left and right cerebral hemisphere for the 29 subjects in Table 1[a]

		Hippocampus			Amygdala			PHG			Temporal pole		
		L	R	C	L	R	C	L	R	C	L	R	C
	L	−0.079	-	-	-	-	-	-	-	-	-	-	-
Amygdala	R	-	−0.110	-	-	-	-	-	-	-	-	-	-
	C	-	-	−0.069	-	-	-	-	-	-	-	-	-
	L	0.313	-	-	0.319	-	-	-	-	-	-	-	-
PHG	R	-	0.262	-	-	0.330	-	-	-	-	-	-	-
	C	-	-	0.301	-	-	0.330	-	-	-	-	-	-
	L	0.283	-	-	0.271	-	-	0.355	-	-	-	-	-
T. pole	R	-	0.111	-	-	0.343	-	-	0.219	-	-	-	-
	C	-	-	0.169	-	-	0.31	-	-	0.280	-	-	-
	L	0.176	-	-	0.228	-	-	0.221	-	-	0.297	-	-
T. lobe	R	-	0.435	-	-	0.351	-	-	0.263	-	-	0.454	-
	C	-	-	0.356	-	-	0.304	-	-	0.194	-	-	0.378

[a]The p-values for the correlation are given in parenthesis. For 27 degrees of freedom a value of 0.311 is required for the correlation to reach significance at the 95% confidence level and 0.431 to reach significance at the 99% confidence level. When the results from the left (L) and right (R) are combined (C) the respective values for 56 degrees of freedom are 0.218 and 0.304.

phase 2 of stage 1 are also presented in Table 4. The only significant correlation was between scores on the Gestalt and Benton tests ($r = 0.65$, $p = 0.03$). The subjects selected for MR imaging (phase 3) had a wide variety of scores in FRM Test A, which were not related to verbal recognition ability or to non-verbal processing ability (as indicated by scores on the Gestalt and Benton tests).

For the eleven subjects who completed testing, the relationship between the FRM test scores, obtained in phase 2, and the volumes of the brain structures measured using the MR images obtained in phase 3, was assessed. It had been intended to use analysis of variance since the original selection procedures were designed to create two distinct groups on the basis of FRM performance. However, there was some overlap between the groups due to inconsistent performances between the two phases of testing, and also due to certain of the subjects being unavailable for MR imaging. Therefore, correlational statistics were employed instead. The volumetric data were treated both as absolute values and also as a proportion of the corresponding hemisphere volume. Furthermore, the ratios of the volumes of the structures in the right and left cerebral hemispheres were investigated to see whether there was any interaction between structure asymmetry and performance in FRM Test A. The results are shown in Table 5. There was a significant negative correlation between the volume of the right amygdala and FRM ($r = 0.697$, $p = 0.017$), such that the higher the performance on FRM, the smaller the right amygdala. There was a non-significant tendency for the right temporal lobe to show the same pattern ($r = -0.515$, $p = 0.105$) and also the left amygdala ($r = -0.526$, $p = 0.096$).

3.2. Stage 2

The means and standard deviations of the results of FRM Test A and the 150 word recognition memory test and a 24 item pattern recognition test, given in phase 1 of stage 2, are presented in Table 4. The values presented in the table refer to the eighteen subjects who went on to complete the three phases of stage 2. Calculation of Pearson's product moment correlation coefficient (computed for all 100 subjects who completed phase 1) revealed no significant relationships between performance on any of the memory tests.

The means and standard deviations of the results of FRM Test B, the Benton face processing test and the Gestalt fragmented figures test administered in phase 2 of stage 2 are also presented in Table 4. Correlations were also computed for the performance on the different psychometric tests of these eighteen subjects. No significant relationships were found between the level of performance on any of these tests. The subjects selected had a wide variety of FRM scores, which were not related to word recognition memory ability or to non-verbal processing ability.

MR data were obtained for eighteen subjects in phase 3 of stage 2. The relationship between the phase 2 FRM results and the phase 3 anatomical findings was

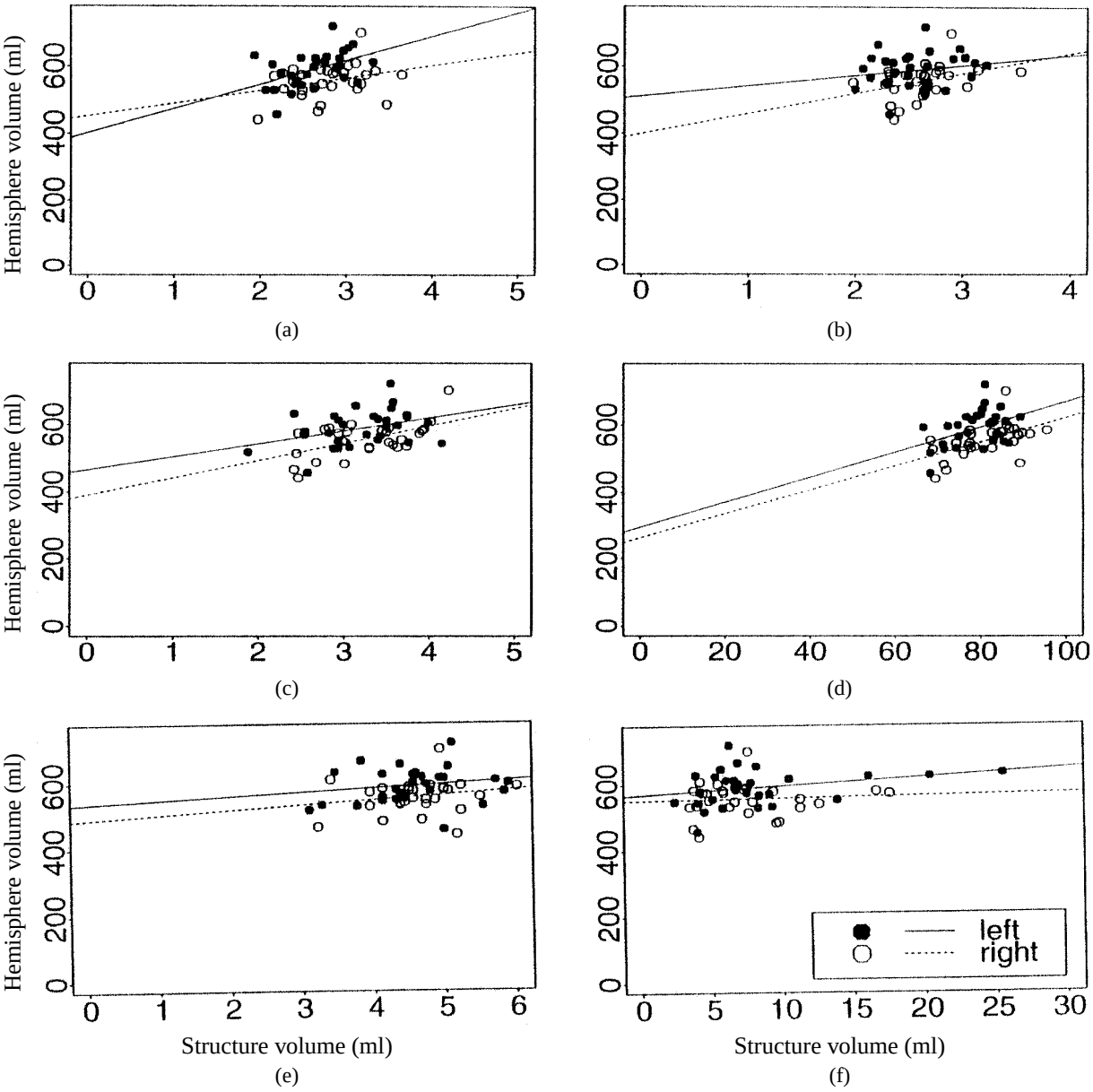

Fig. 3. Relationship between the volumes of the right and left (a) hippocampus, (b) amygdala, (c) parahippocampal gyrus, (d) temporal lobe, (e) caudate nucleus and (f) lateral ventricle and the volume of the respective cerebral hemisphere.

examined for these eighteen subjects using the same statistical methods as for stage 1. The results are presented in Table 6. No significant correlations were found between performance in FRM Test B and either the absolute volumes or the proportional volumes of the measured brain structures. There was, however, a significant correlation between the ratio of the left to the right temporal pole volume and FRM performance. This indicates that the smaller a subjects' right temporal pole was relative to their left temporal pole, the better their FRM. No other correlations between left/right ratios of structural volumes and FRM test performance were significant.

One of the problems with subject selection, as already mentioned, was the inconsistency of some of the subjects' scores between phase 1 and 2 of testing. We, therefore, tested the possibility of a relationship between FRM and structural volume in the subjects with the most consistent scores across the two phases. From the entire sample of 29 subjects, taken from both

Table 4
Results of the 50 faces, 50 words, 24 patterns, 100 faces, Benton and Gestalt psychometric tests given to the eleven subjects in stage 1 and eighteen subjects in stage 2[a]

Subjects (stage 1)	50 faces	50 words	24 patterns	FRM Test A	Benton	Gestalt	100 words	EHI
01_H	50	50	23	99	49	19	100	100
02	50	50	22	89	43	18	97	65
03	49	48	20	84	50	18	99	−57
04_H	49	49	19	98	50	18	99	87
05	49	50	23	91	52	19	84	100
06_H	49	49	23	99	52	19	99	82
07_L	29	40	14	59	50	16	96	−20
08	30	50	14	82	41	7	97	67
09	32	49	16	89	41	14	99	100
10	33	47	16	74	49	12	95	53
11_L	34	48	17	72	50	19	93	85
mean	41.3	48.2	18.8	85.1	47.9	16.3	96.2	
SD	9.4	2.9	3.6	12.7	4.2	3.8	4.6	

Subjects (stage 2)	FRM Test A	150 words	24 patterns	FRM Test B	Benton	Gestalt		EHI
12_H	95	128	18	92	50	9		NA
13_H	92	130	11	94	49	6		100
14	97	135	22	90	47	12		−60
15	93	122	18	89	47	13		100
16	86	125	18	94	50	12		100
17	91	144	21	98	50	14		−90
18	92	143	20	88	49	11		87
19	84		15	91	50	15		NA
20	82		14	86	50	16		67
21	87		21	82	50	16		87
22	89	115	14	78	49	19		100
23_L	72	128	14	67	49	14		43
24_L	71	136	16	63	45	14		60
25	75	136	20	81	42	13		94
26	69	135	23	82	50	17		NA
27	58		15	77	52	15		41
28_L	63		15	70	45	8		85
29	64	121	20	92	*	*		100
mean	81.1	130.6	17.5	84.1	48.5	13.2		
SD	12.3	8.6	3.4	9.9	2.5	3.3		

[a]The subjects in stage 2 additionally received a 150 words test. Subjects with consistently high FRM scores are denoted by $_H$ and those with consistently low scores by $_L$.

stages of the study, the five most consistently high and the five most consistently low scorers on the FRM test were selected. Degree of consistency was measured as the difference in face recognition test scores between phase 1 and phase 2 (to make the comparison between the two stages equivalent, scores on the face RMT test from phase 1 of stage 1 were doubled). The five most consistent high and the five most consistent low scorers were then selected. As Table 4 shows, five subjects were selected from each stage of the experiment. T-tests were used to determine whether structural volumes differed between the consistently high and low scoring groups. The only significant group differences were found with the amygdala. The right amygdala was significantly smaller in the high performance group compared with the low performance group ($t = 2.77$, $p = 0.025$).

4. Discussion

This study applied the Cavalieri method of modern design stereology to high resolution MR images in order to determine the volumes of the cerebral hemispheres and a range of brain structures. The MR imaging was conducted in 29 healthy young males with no history of brain damage, who had undergone a non-verbal memory battery. This study was hypoth-

Table 5
Pearson's product moment correlation coefficient for linear regression analysis of FRM scores against the absolute volume of the hippocampus, amygdala, parahippocampal gyrus, caudate nucleus, temporal pole, lateral ventricles and cerebral hemisphere for the eleven subjects in stage 1[a]

	Absolute volumes		Volumes as a proportion of h'sphere		Ratio of left to right structure volumes
	left	right	left	right	
Hippocampus	0.07	−0.12	−0.14	0.07	−0.34
P'hipp. gyrus	0.19	0.19	0.24	0.24	−0.08
Amygdala	−0.70[b]	−0.53	−0.63[b]	−0.49	0.13
Caudate	−0.06	−0.24	−0.05	−0.15	−0.31
Temp. Pole	0.08	−0.08	0.10	−0.07	−0.15
Temp. Lobe	−0.52	−0.26	−0.37	−0.26	0.20
Ventricles	−0.07	0.14	−0.08	0.13	0.50
Hemisphere	0.01	−0.00			−0.06

[a]Results are also presented for the volume as a proportion of same hemisphere volume and for the ratio of left:right hemisphere volume for each of the structures.

[b]Significance at the 95% confidence level with 9 degrees of freedom requires a correlation coefficient of 0.62.

Table 6
Pearson's product moment correlation coefficient for linear regression analysis of FRM scores against the absolute volume of the hippocampus, amygdala, parahippocampal gyrus, caudate nucleus, temporal pole, lateral ventricles and cerebral hemisphere for the eighteen subjects in stage 2[a]

	Absolute volumes		Volumes as a proportion of h'sphere		Ratio of left to right structure volumes
	left	right	left	right	
Hippocampus	−0.03	0.04	0.09	−0.06	0.12
P'hipp. gyrus	0.20	0.20	0.25	0.19	0.00
Amygdala	0.09	−0.09	0.08	−0.03	−0.27
Caudate	0.05	−0.03	0.11	−0.05	−0.14
Temp. Pole	0.14	−0.28	0.22	−0.33	−0.52[b]
Temp. Lobe	0.06	−0.10	0.04	−0.05	−0.22
Ventricles	0.12	0.20	0.13	0.19	0.22
Hemisphere	−0.06	0.02			0.13

[a]Results are also presented for the volume as a proportion of same hemisphere volume and for the ratio of left:right hemisphere volume for each of the structures.

[b]Significance at the 95% confidence level with 16 degrees of freedom requires a correlation coefficient of 0.40.

esis driven. We sought to test whether FRM is better in healthy young subjects in whom the parahippocampal gyrus (and possibly also the hippocampus) on the right and the amygdala in both hemispheres are larger. A range of tests and correlation analyses have, however, also been performed using the wide variety of neuroanatomical and neuropsychological data obtained. In particular, we have (i) investigated whether there is a significant difference between the volume of the hippocampus, parahippocampal gyrus, amygdala, caudate nucleus, temporal pole, temporal lobe or cerebral hemisphere on the right and left sides of the brain, (ii) investigated whether the volume of the right and left hippocampus, parahippocampal gyrus, amygdala, caudate nucleus, temporal pole and temporal lobe are correlated with the volume of the corresponding whole cerebral hemisphere, (iii) investigated whether the individual (i.e., right or left) or total (i.e., combined) volumes of the hippocampus, parahippocampal gyrus, amygdala, temporal pole and temporal lobe are correlated with each other, (iv) investigated whether significant correlations exist between the results of the different elements of the neuropsychology test batteries given in stages 1 and 2, (v) investigated whether the FRM scores of subjects in stage 1 and stage 2 correlated with the volumes of the hippocampus, parahippocampal gyrus, amygdala, caudate nucleus, temporal pole, temporal lobe, lateral ventricles or whole cerebral hemisphere (reported in absolute terms, as a proportion of hemisphere volume, when relevant, or as the ratio of the volume of the relevant structure on the left and right sides of the brain) and (vi) after combining the data available from stage 1 and stage 2 of the study investigated whether there were any group differences between the absolute, proportional or relative right and left volumes of the structures mentioned in (v) above for the five subjects with the most consistently high FRM scores and the five subjects with the most consistently low FRM scores across the two phases of testing that were an inherent part of each stage. Strictly speaking, it is necessary to make allowance for the fact that in (ii) to (vi) above multiple tests have been performed with the same data and significant effects may have arisen from cumulative chances of error. To keep a result significant at the 95% confidence level despite multiple comparisons the relevant Bonferroni correction requires that the chance of error should be reduced to the value obtained by dividing 5% by the number of tests performed. This correction is commonly regarded as unnecessarily stringent, but it is prudent to bear in mind that 12 different tests were performed in (ii) above, 30 different tests were performed for each of the different structure volumes in (iii) above (i.e., Table 3), 38 different tests were performed for the individual FRM data sets in (v) (i.e., Tables 5 and 6) and for the combined FRM data set in (vi) above (see text). It can be argued that for the results to be truly significant at the 95% confidence level the respective p values in (ii), (iii), (v) and (vi) above should be adjusted

from 0.05 to 0.05/12, 0.05/30, 0.05/38, and 0.05/38, respectively.

Several findings have emerged from our study. First, substantial inter-individual variation in the volumes of different brain structures was demonstrated. The estimated biological Coefficient of Variation in the volumes of the cerebral hemispheres (and temporal lobe), various temporal lobe structures (and the caudate nucleus), and the lateral ventricles was approximately $< 10\%$, 10 to 20% and $> 50\%$ respectively. As can be seen from Table 1, the variability in all of the structures was large despite the narrow population in this sample. This finding is consistent with previous reports of large variations in the volume of the hippocampus among normal young subjects (see [21] for review). It seems likely that this degree of variability will have functional implications. The data concerning the biological variability of structural volumes also have important implications for the design of any study concerned with testing whether the volumes of particular brain structures are significantly larger in one group of subjects than in another. The magnitude of the inherent biological variation determines how many subjects need to be studied in order to confirm whether differences between group means are significant. For example, if the observed difference in the mean volume of a structure between two groups of subjects is 10%, then 7, 22 or 50 subjects would be needed to confirm that there was a significant between group volume difference at the 95% confidence level depending on whether the coefficient of biological variability was 10%, 20% or 30%. This calculation assumes that the samples are independent, that the distribution of volumes is unimodal and symmetric, and that there is no imprecision in the individual volume estimates.

Second, the left cerebral hemisphere of our male subjects was found to be larger than the right hemisphere in 26 of the 29 subjects. Across all subjects the right hemisphere was 2.7% larger than the left and this difference was significant ($p < 0.02$). To our knowledge this has not been reported before either from studies of post-mortem brains or from in vivo MRI analyses. Comparative in vivo data are sparse, owing to difficulties in reliably and efficiently estimating cerebral hemisphere volumes before the advent of modern design stereological methods. Such data as are available tend to be for pathological cases and older populations. Tramo et al. [40] have reported, however, that the surface area of the left cerebral hemisphere is more variable than the right in normal young subjects. This suggests that factors affecting the development of one hemisphere may not equally affect the other. Filipek et al. [13] reported that, in normal young subjects, all bilateral structures were symmetric or nearly symmetric in volume, with the exception of a slightly larger right neocortex and amygdala, and a larger left ventricle. Unfortunately, however, these researchers did not use stereological methods so their conclusions may be questionable. The asymmetry observed in the present study merits further investigation to determine whether it generalises to a wider sample of young male subjects, to young females, and to other subject populations.

Third, the volumes of all structures are highly correlated between the right and left sides of the brain ($p < 0.0002$ for all comparisons). However, only the right and left parahippocampal gyrus, temporal pole and temporal lobe and left hippocampus are significantly correlated with the volume of the corresponding hemisphere, and these correlations lose significance after Bonferroni correction for 12 comparisons. Previous normative MR studies which have examined the temporal region have correlated structure volumes with variables such as age or sex (for example, [7, 13]), but have not focused upon the issue of whether or not to normalise to cerebral hemisphere volume; probably because the latter is difficult to measure using conventional planimetric techniques. Correlations between structure and hemisphere volume might be anticipated from the fact that the structures are projection sites for a complex sequence of to-be-remembered information that has already been processed by the association cortices. The findings raise important issues regarding the normalisation of volumetric data obtained for brain structures in different subjects. If the size of a particular brain structure relates to the volume of the brain, then it may be important to normalise volumes with respect to hemisphere or brain volume before computing neurobehavioural correlations. If there is no systematic regional-global correlation, however, then it may be inappropriate to perform this normalisation procedure. This issue becomes more complex when one considers older subjects in whom appreciable cerebral atrophy may have occurred. In these cases, the only objective means for normalising data may be through the determination of global cranial capacity. Unfortunately, such an approach could render problematic the separate normalisation of right and left sided structures.

Fourth, despite the variability in the volumes of the temporal lobe structures across the different subjects, a positive relationship was not found in either stage

of the study between face recognition memory and the volumes of the hippocampus and parahippocampal gyrus on the right or the volumes of the amygdala bilaterally. Failure to find such a positive relationship between face recognition memory and the measured volumes of these key structures may have arisen for several reasons. Thus, the structures may not be centrally concerned with face recognition memory, face recognition memory may also be determined by the variable contribution of other structures, and structural volume may sometimes be a relatively minor determinant of functional efficiency.

We hypothesised that the volume of the parahippocampal gyrus (and possibly also the hippocampus) on the right and amygdala bilaterally would positively correlate with face recognition memory. It is unclear to what extent hippocampal lesions disrupt recognition memory as some studies of temporal lobe epileptics fail to find a recognition deficit in the face of a recall memory deficit (for example, see [27]). Golomb et al. [15, 16] used MRI and Computed Tomography (CT), and more recently just MRI, to investigate the relationship between hippocampal volume and memory in healthy elderly subjects. They found a positive relationship between subjects' delayed recall memory and their hippocampal volumes. Similar delay specific hippocampal correlations have been obtained by Foster, Kohler and colleagues in patients with Alzheimer's disease and matched control subjects [14, 22]. Since the work described in this paper was performed increasing evidence has emerged that the hippocampus is minimally involved in recognition memory, including face recognition memory. Thus, a meta-analysis of the amnesia literature on the Warrington Recognition Test and recall performance by Aggleton and Shaw [2] led to the conclusion that whereas hippocampal lesions impair episodic recall to a similar extent to that found in other more globally impaired amnesics, they have a minimal effect on recognition memory of words and faces. Baxendale [5] has also shown convincingly that selective hippocampal sclerosis does not impair either word or face recognition memory. The hippocampus may, therefore, be little involved with the processing necessary for good recognition. Instead, it may be profitable to investigate whether the hippocampal volumes of healthy young male subjects correlates positively with recall memory, particularly after a delay, as suggested by the work of Golomb et al. [16]. Furthermore, it is possible that the parahippocampal gyrus is more concerned with the recognition of spatial relationships than it is with the recognition of recently presented visual objects such as faces, which may be the function of the more anterior perirhinal cortex. This has been suggested by Squire et al. [39] on the grounds that the inferotemporal cortex, which processes visual pattern information, projects more strongly to this region whereas the parietal cortex, which processes spatial information, projects to the parahippocampal cortex, but not to the perirhinal cortex. The volume of the perirhinal cortex was not measured in the current study. Nor to our knowledge has it been measured in any other published work. It is also unclear whether the amygdala is involved with face recognition memory primarily or merely as a secondary consequence of its role in the recognition of facially communicated emotions such as fear (see [41]).

The current study failed to find even a hint of a positive relationship between the volumes of any of the measured structures in the medial temporal lobes and face recognition memory. We are currently developing a procedure for measuring the volume of the perirhinal cortex – if the arguments of Squire et al. [39] are correct, there may be a positive relationship between the volume of this structure and FRM test performance. Even so, it may be that a positive relationship does not exist between the volume of structures in the medial temporal lobes and memory efficiency in subjects without brain damage even though these structures play a vital role in mediating memory. This could be because functional efficiency depends mainly on the relative, rather than absolute, sizes of structures making up the critical neural network (as might apply, for example, to structures making up a larger hippocampal circuit, or a larger amygdalar circuit), which we have not assessed. The finding which emerged from analysis of right vs left temporal pole volume indicates that the degree of hemispheric asymmetry in the relative size of homologous bilateral structures may also be important in determining memory performance.

The amygdala and the temporal pole were involved in the only suggestive relationships between structural volumes and face recognition memory to emerge from the current study. Stage 2 of the study found that FRM test performance tended to be better in young males with right temporal pole relatively smaller than the left (there was a weaker trend of the same kind in stage 1). When the data for the consistently good and poor scorers on the FRM test were considered, there was no difference in the size of the left amygdala, but the right amygdala was smaller in the high scorers. These findings, none of which remained significant after Bonferroni correction for multiple comparisons, provide

weak evidence that good face recognition memory may be associated with having an amygdala and temporal pole that is relatively smaller on the right. The possibility that improved performance correlates with a reduction in structure volume (or at least a relative reduction) may be explained by an alternative to the neophrenological hypothesis in which the specificity of a brain structure for a particular task is achieved by customised reduction rather than enlargement of its components. In this sense, a functionally advantageous asymmetry in the volume of the amygdala is achieved by the pruning of one side rather than the enlargement of the other at a particular neurodevelopmental stage. This possibility, which appears plausible in view of the fact that during development the cerebral cortex generates more neurons, axons, synapses and receptors than it finally keeps (see [10] for a review), will need to be examined in a larger study which may confirm or disconfirm the above suggestion that relatively smaller right than left amygdala and possibly temporal pole (which includes perirhinal cortex) volumes in healthy young males are associated with better face recognition memory. This surprise finding is consistent with the amygdala and temporal pole forming part of a system that mediates functions important in face recognition memory.

In future studies, in addition to obtaining volume measures we intend to use MR tissue characterisation procedures such as T2 relaxation time and Magnetisation Transfer Ratio mapping so as to be able to better assess the integrity of structures of the medial temporal lobes. It is also important to improve the criteria used for defining the boundaries of structures of interest, and to develop strategies for measuring the volume of structures such as the perirhinal cortex. In addition, future work will need to control more carefully intersubject variability to the psychological processes of interest. By combining these approaches with alternative memory measures, it should be possible to make further inroads into delineating the anatomical bases of memory.

Acknowledgements

We wish to acknowledge the contributions of Sarah Hindmarch, Ian Ensum, Chris Spellerberg and Kieron Harvey in obtaining the data presented here. Valerie Peeck is thanked for her assistance with some of the statistical analyses. We are also grateful to Professor Robb of the MAYO Foundation, Minnesota, USA for provision of a collaborative agreement for the development of the ANALYZE software. This research was supported by Grant No. 68922713N from the Medical Research council of the United Kingdom awarded to Mayes, Chadwick and Roberts.

References

[1] J.P. Aggleton, ed., *The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction,* Wiley-Liss, New York, 1992.

[2] J.P. Aggleton and C. Shaw, Amnesia and recognition memory: A re-analysis of psychometric data, *Neuropsychologia* **34** (1996), 51–62.

[3] N.C. Andreasen, M. Flaum, V.H. Swayze, T. Cizadlo, M. Flaum, D. O'Leary, J. Ehrhardt and W.T.C. Yuh, Intelligence and brain structure in normal individuals, *American Journal of Psychiatry* **150** (1993), 130–134.

[4] G. Bartzokis, J. Mintz, P. Marx, D. Osborn, F. Chiang, C. Phelan and S. Marder, Reliability of in vivo volume measures of hippocampus and other brain structures using MRI, *Magnetic Resonance Imaging* **11** (1993), 933–1006.

[5] S.A. Baxendale, The role of the hippocampus in recognition memory, *Neuropsychologia* **35** (1997), 591–598.

[6] A.L. Benton, K. Hamsher, N.R. Varney and O. Spreen, *Facial Recognition: Stimulus and Multiple Choice Pictures*, Oxford University Press, New York, 1978.

[7] S. Bhatia, S.Y. Bookheimer, W.D. Gaillard and W.H. Theodore, Measurement of whole temporal-lobe and hippocampus for MR volumetry – normative data, *Neurology* **43** (1993), 2006–2010.

[8] M.W. Brown, F.A. Wilson and I.P. Riches, Neuronal evidence that the inferomedial temporal cortex is more important than the hippocampus in certain processes underlying recognition memory, *Brain Research* **409** (1987), 158–162.

[9] G.D. Cascino, C.A. Jack, J.E. Parisi, F.W. Sharbrough, K.A. Hirschorn, F.B. Meyer, W.A. Marsh and P.C. O'Brien, Magnetic Resonance imaging-based volume studies in temporal lobe epilepsy: pathological correlations, *Annals of Neurology* **30** (1991), 31–36.

[10] W.M. Cowan, J.W. Fawcett, D.D.M. O'Leary and B.B. Stanfield, Regressive events in neurogenesis, *Science* **225** (1984), 1258–1265.

[11] L.M. Cruz-Orive, Systematic sampling in stereology. Bulletin of the International Statistical Institute 55, 1993, 451–468.

[12] B. Deweer, S. Lehericy, B. Pillon, M. Baulac, J. Chiras, C. Marsault, Y. Agid and B. Dubois, Memory disorders in probable Alzheimer's disease – the role of hippocampal atrophy as shown with MRI, *Journal of Neurology, Neurosurgery and Psychiatry* **58** (1995), 590–597.

[13] P.A. Filipek, C. Richelme, D.N. Kennedy and V.S. Caviness, Young-adult human brain – an MRI-based morphometric analysis, *Cerebral Cortex* **4** (1994), 344–360.

[14] J.K. Foster, S.E. Black, B.H. Buck and M.J. Bronskill, Executive and associated functions in normal and pathological aging: a neuroimaging perspective, in: *Methodology of Frontal and Executive Functions*, P. Rabbitt, ed., Lawrence Erlbaum, Hove, in press.

[15] J. Golomb, M.J. de Leon, A. Kluger, A.E. George, M.A. Tarshish and S.H. Feris, Hippocampal atrophy in normal aging: an association with recent memory impairment, *Archives of Neurology* **50** (1993), 967–973.

[16] J. Golomb, A. Kluger, M.J. de Leon, S.H. Feris, A. Convit, M.S. Mittleman, J. Cohen, H. Rusinek, S. De Santi and A.E. George, Hippocampal formation size in normal human aging: a correlate of delayed secondary memory performance, *Learning and Memory* **1** (1994), 45–54.

[17] C.L. Grady, A.R. McIntosh, B. Horwitz, J.M. Maisog, L.G. Ungerleider, M.J. Mentis, P. Pietrini, M.B. Schapiro and J.V. Haxby, Age-related reductions in human recognition memory due to impaired encoding, *Science* **269** (1995), 218–221.

[18] H.J.G. Gundersen and E.B. Jensen, The efficiency of systematic sampling in stereology and its prediction, *Journal of Microscience* **147** (1987), 229–263.

[19] H.J.G. Gundersen and R. Osterby, Optimising sampling efficiency of stereological studies in biology: or 'Do more less well!', *Journal of Microscopy* **121** (1981), 65–73.

[20] J.B. Hellige, *Hemispheric Asymmetry: What's Right and What's Left*, Harvard University Press, Cambridge, 1993.

[21] C.R. Jack, W.H. Theodore, M. Cook and G. McCarthy, MRI-based hippocampal volumetrics – data-acquisition, normal ranges, and optimal protocol, *Magnetic Resonance Imaging* **13** (1995), 1057–1064.

[22] S. Kohler, S.E. Black, D'ý. Kidron, C. Szekely, B.H. Buck, J.K. Foster, J.P. Szalai, G. Winocur, M. Moscovitch and M. Bronskill, Hippocampal atrophy and memory impairments in Alzheimer's disease: correlations between MRI and neuropsychological measures. Poster presented at the Sixth Annual Rotman Research Institute Conference, March 1996, Toronto, Canada.

[23] R.I. Kuzniecky and G.D. Jackson, *Magnetic Resonance in Epilepsy*, Raven Press, New York, 1995.

[24] T. Lencz, G. McCarthy, R.A. Bronen, T.M. Scott, J.A. Inserni, K.J. Sass, R.A. Novelly, J.H. Kim and D.D. Spencer, Quantitative magnetic resonance imaging in temporal lobe epilepsy: relationship to neuropathology and neuropsychological function, *Annals of Neurology* **31** (1992), 629–637.

[25] K. Light, N. Roberts, G.H. Whitehouse and R.H.T. Edwards, Unbiased and efficient estimation of bladder volume with MR imaging, *Journal of Magnetic Resonance Imaging* **5** (1995), 33–41.

[26] G. Matheron, *The Theory of Regionalised Variables and its Applications. Les Cahiers du Centre de Morphologie Mathematique de Fontainebleu, No 5*, Ecole National Superieure des Mines de Paris, Fontainebleu, 1971.

[27] L.A. Miller, D.G. Munoz and M. Finmore, Hippocampal sclerosis and human memory, *Archives of Neurology* **50** (1993), 391–394.

[28] M. Mishkin, Memory in monkeys severely impaired by combined but not separate removal of amygdala and hippocampus, *Nature* **273** (1978), 279–299.

[29] P.G. Nestor, M.E. Shenton, R.W. McCarley, R. Haimson, S.R. Smith, B. O'Donnell, M. Kimble, R. Kikinis and A.J. Ferec, Neuropsychological correlates of MRI temporal lobe abnormalities in schizophrenia, *American Journal of Psychiatry* **150** (1993), 1849–1855.

[30] R. Nieuwenhuys, J. Voogel and C. Van Huijzen, *The Human Central Nervous System: A Synopsis and Atlas, 3rd Edition (Revised)*, Springer-Verlag, Berlin, 1988.

[31] N. Roberts, L.M. Cruz-Orive, N.M.K. Reid, D.A. Brodie, M. Bourne and R.H.T. Edwards, Unbiased estimation of human body composition by the Cavalieri method using magnetic resonance imaging, *Journal of Microscopy* **171** (1993), 239–253.

[32] N. Roberts, S. Barbosa, L.D. Blumhardt, R.A. Kawoski and R.H.T. Edwards, Stereological estimation of the total volume of MR visible brain lesions in patients with multiple sclerosis, *MAGMA* **2** (1994a), 375–378.

[33] N. Roberts, A.S. Garden, L.M. Cruz-Orive, G.H. Whitehouse and R.H.T. Edwards, Estimation of fetal volume by magnetic resonance imaging and stereology, *British Journal of Radiology* **67** (1994b), 1067–1077.

[34] P. Sabbah, S. de Schonen, G. Salamon, F. Briant and O. Pascalis, A fMRI study of face recognition, *Human Brain Mapping Supplement* **1** (1995), 46.

[35] K.J. Sass, A. Sass, M. Westerveld, T. Lencz, R.A. Novelly, J.H. Kim and D.D. Spencer, Specificity in the correlation of verbal memory and hippocampal neuron loss: Dissociation of memory, language and verbal intellectual ability, *Journal of Clinical and Experimental Neuropsychology* **14** (1992), 662–672.

[36] G. Schlaug, L. Jancke, Y. Huang and H. Steinmetz, In vivo evidence of structural brain asymmetry in musicians, *Science* **267** (1995), 699–701.

[37] W.B. Scoville and B. Milner, Loss of recent memory after bilateral hippocampal lesions, *Journal of Neurology, Neurosurgery and Psychiatry* **20** (1957), 11–22.

[38] J. Sergent, S. Ohta and B. MacDonald, Functional anatomy of face and object processing, *Brain* **115** (1992), 15–36.

[39] L.R. Squire, B. Knowlton and G. Musen, The structure and organization of memory, *Annual Review of Psychology* **44** (1993), 453–495.

[40] M.J. Tramo, W.C. Loftus, C.E, Thomas R.L. Green, L.A. Mott and M.S. Gazzaniga, Surface area of human cerebral cortex and its gross morphological subdivisions: in vivo measurements in monozygotic twins suggest differential hemisphere effects of genetic factors, *Journal of Cognitive Neuroscience* **7** (1995), 292–301.

[41] A.W. Young, J.P. Aggleton, D.J. Hellawell, M. Johnson, P. Broks and J.R. Hanley, Face processing impairments after amygdalotomy, *Brain* **118** (1995), 15–24.

[42] L. Willerman, R. Schultz, J.N. Rutledge and E.D. Bigler, In vivo brain size and intelligence, *Intelligence* **15** (1991), 229–246.

[43] S.M. Zola-Morgan, L.R. Squire and D.G. Amaral, Lesions of the amygdala that spare adjacent cortical regions do not impair memory or exacerbate the impairment following lesions of the hippocampal formation, *Journal of Neuroscience* **9** (1989), 1922–1936.

[44] S.M. Zola-Morgan and L.R. Squire, The primate hippocampal formation: Evidence for a time limited role in memory storage, *Science* **250** (1990), 288–290.