

Plasticity in the Human Hippocampus

Katherine Woollett

Submitted for PhD in Cognitive Neuroscience

October 2010



University College London

Supervisor: Eleanor A. Maguire

Declaration:

I, Katherine Woollett, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Date:

Abstract

If we are to approach rehabilitation of memory-impaired patients in a systematic and efficacious way, then it is vital to know if the human memory system has the propensity for plasticity in adulthood, the limiting factors on such plasticity, and the timescales of any plastic change. This thesis was motivated by an attempt to develop a body of knowledge in relation to these questions. There is wide agreement that the hippocampus plays a key role in navigation and memory across species. Evidence from animal studies suggests that spatial memory-related hippocampal volume changes and experience-related hippocampal neurogenesis takes place throughout the lifespan. Previous studies in humans indicated that expert navigators, licensed London taxi drivers, have different patterns of hippocampal grey matter volume relative to control participants. In addition, preliminary evidence also suggested there may be functional consequences associated with this grey matter pattern.

Using licensed London taxi drivers as a model for learning and memory, the work undertaken centered on four key issues: (1) In a set of studies, I characterised the neuropsychological profile of licensed London taxi drivers in detail, which included devising a number of new table-top associational memory tests. This enabled me to assess the functional consequences of their expertise and hippocampal grey matter pattern in greater depth than previous studies. (2) In order to explore the effects of taxi drivers' expertise in more naturalistic settings, I also examined how well they could learn the layout of an unfamiliar town compared with a group of non-taxi drivers, and how effectively taxi drivers could integrate a new district into their existing spatial representation of London. (3) I then conducted a study on experts whose knowledge was much less spatial than taxi drivers in order to examine if the effects on hippocampal grey matter and neuropsychology were general or whether they were specific to the spatial domain. (4) Given that previous taxi driver studies were cross-sectional, the question of whether the human hippocampus can exhibit spatial memory-related structural plasticity in adulthood was uncertain. I therefore conducted a longitudinal study which assessed participants both pre and post taxi driver training using structural MRI and neuropsychological measures. This enabled me to investigate, within subjects, whether hippocampal volume changes can be acquired in response to intense spatial stimulation. In addition, I explored

whether ceasing to be a taxi driver (i.e. retiring after many years on the job) resulted in 'reverse' plasticity.

I found evidence for hippocampal plasticity within individuals as a result of their intense acquisition of spatial knowledge over a number of years that was associated with qualifying to be a licensed London taxi driver, and preliminary evidence of reverse plasticity when taxi drivers retire. This suggests that hippocampal structure and memory ability can be modified in response to environmental factors and are not necessarily hard-wired. However, my results also provide some insights into the boundaries within which human memory operates, as I identified both positive and negative cognitive consequences of being an expert navigator, and also established that the MRI and neuropsychology effects of expertise on the hippocampus may be restricted to the spatial domain.

Acknowledgements

The work in this thesis could not have been accomplished without the generous help of a number of people. I wish to thank Prof. Eleanor Maguire, my principal supervisor, for her invaluable support and advice throughout my PhD, and for being so encouraging, enthusiastic and patient, and Prof. Neil Burgess, my secondary supervisor.

Next there are my friends and colleagues to thank from UCL: Nikolaus Weiskopf, Chloe Hutton and John Ashburner for their technical advice; Alex Leff and Jenny Crinion for being so helpful and generous with their time; Alice Neal, Debbie Talmi and Jen Summerfield for making me smile; the support staff at the FIL; Peter Aston, David Bradbury, Janice Glensman, Ric Davis and Chris Freemantle for help with scanning and computers.

Finally, I would like to thank my mother, Beto and Marga Martin for teaching me to never give up, without which I would not have been here today doing this PhD.

Table of Contents

Chapter 1- Does experience change the brain?	12
1.1 Introduction	13
1.2 Structural effects of expertise in the healthy adult human brain	17
1.3 Hippocampal anatomy	19
1.4 The functions of the hippocampus	26
1.5 Structural changes in the hippocampus	31
1.6 Licensed London taxi drivers – a model system	33
1.7 What are the underlying mechanisms of plasticity?	42
1.8 Functional differentiation along the long-axis of the hippocampus	49
1.9 Thesis overview	54
Chapter 2 - Materials and Methods	56
2.1 Introduction	57
2.2 Participants	57
2.3 Biophysics of magnetic resonance imaging	60
2.3.1 Magnetic fields	60
2.4 Generating an MR signal	62
2.5 Tomographic image formation	64
2.6 Scanning parameters	64
2.7 Structural MRI sequences	66
2.8 Automated analysis of structural MRI scans	69
2.8.1 Standard voxel-based morphometry (VBM)	69
2.8.2 Diffeomorphic Anatomical Registration using Exponentiated Lie algebra (DARTEL)	71
2.8.3 High Dimensional Warping	71
2.9 Statistical inference	72
2.10 Neuropsychological tests	73
Chapter 3 - The effects of navigational expertise: an in-depth analysis	75
3.1 Introduction	76
3.2 Experiment 1	78
3.2.1 Materials and methods	78
3.2.2 Procedure	89
3.2.3 Data analysis	89
3.3 Results	90
3.3.1 MRI data	90

3.3.2	Neuropsychological data	91
3.4	Discussion	97
3.5	Experiment 2	101
3.5.1	Material and Methods	102
3.5.2	Procedure and data analyses	107
3.5.3	Results	107
3.5.4	Discussion	108
Chapter 4 -The effects of navigational expertise on wayfinding in new environments		111
4.1	Introduction	112
4.2	Experiment 3	114
4.2.1	Materials and methods	114
4.2.2	Procedure	123
4.2.3	Data analysis	123
4.2.4	Results	124
4.2.5	Discussion	132
Chapter 5 - Non-spatial expertise and hippocampal grey matter volume		139
5.1	Introduction	140
5.2	Experiment 4	142
5.2.1	Material and methods	142
5.2.2	Procedure	144
5.2.3	Data analyses	144
5.2.4	Results	144
5.2.5	Discussion	148
Chapter 6 - Hippocampal plasticity: a longitudinal study		151
6.1	Introduction	153
6.2	Experiment 5	158
6.2.1	Materials and methods	158
6.2.2	Procedure	162
6.2.3	Data analyses	162
6.2.4	Results: Time 1	163
6.2.5	Results: Time 2	167
6.2.6	Results: Time 1 vs. Time 2	171
6.2.7	Discussion	174
6.3	Experiment 6	177
6.3.1	Material and methods	178
6.3.2	Procedure	182

6.3.3	Data analysis	183
6.3.4	Results	183
6.3.5	Discussion	186
Chapter 7 - General discussion		188
7.1	Plasticity in the human hippocampus	190
7.2	Neuropsychological consequences of navigational expertise	196
7.3	Spatial and non-spatial expertise	201
7.4	Implications for the hippocampus	203
7.5	Conclusions	205
References		208

List of Figures

Figure 1. Example of a sagittal view from a structural MRI scan	14
Figure 2. The human hippocampal formation	19
Figure 3. Schematic representation of the hippocampal network	20
Figure 4. Hippocampal place cells	24
Figure 5. Navigation performance	29
Figure 6. Map of London	34
Figure 7. Taxi driver qualification process	35
Figure 8. Hippocampal volume differences between taxi drivers and control subjects	37
Figure 9. The effect of taxi driving experience on grey matter volume	39
Figure 10. Main effects for the neuropsychological tests	40
Figure 11. Newly generated cells can be detected in the adult human hippocampus in patients	48
Figure 12. 3D modelling of molecularly defined subdivisions of DG, CA3, and CA1 delineated by gene expression	52
Figure 13. Protons align parallel or anti-parallel to the magnetic field (B_0)	62
Figure 14. Generation of MR signal	63
Figure 15. T1 weighted scan - Structural MRI image showing different tissue intensities	66
Figure 16. The Rey-Osterrieth complex figure	84
Figure 17. Example of London landmark recognition test stimuli	86
Figure 18. Example of London proximity judgements test stimuli	87
Figure 19. Object–place associations test	89
Figure 20. Anterior hippocampal grey matter volume differences between taxi drivers and control subjects	92
Figure 21. London proximity judgements	94
Figure 22. Object-place test	96
Figure 23. Visual paired associates test	104
Figure 24. Visual paired associates within a scene context test	104
Figure 25. Face-name associations test	105
Figure 26. Object-sound associations test	106
Figure 27. Verbal paired associates	107
Figure 28. Map of New Town	116
Figure 29. Map of London	120
Figure 30. Example views from the three environments	121
Figure 31. Example sketch maps	126
Figure 32. Scenes correctly recognised across environments	129
Figure 33. Road segments recalled across environments	130
Figure 32. Road junction recalled across environments	131
Figure 35. Knowledge in medical doctors	147
Figure 36. Grey matter differences between qualified London taxi drivers	173
Figure 37. The virtual environment of London (UK)	182

Figure 38. MRI findings in retired taxi drivers	185
Figure 39. Rey complex figure	186
Figure 40. Navigation performance	187

List of Tables

Table 1 - Participant characteristics	80
Table 2 - Basic cognitive measures	94
Table 3 - Stress measures	94
Table 4 - Memory measures	96
Table 5 - Participant characteristics	104
Table 6 - Associative memory measures	109
Table 7 - Performance of both groups on the New Town tests	126
Table 8 - Performance of taxi drivers on the London tests	129
Table 9 - Performance of taxi drivers on tests of existing and new London (scored separately), and New Town	132
Table 10 - Participant characteristics	144
Table 11 - Memory measures	147
Table 12 - Stress measures	147
Table 13 - Time 1 Volunteer characteristics	160
Table 14 - Time 2 Volunteer characteristics	161
Table 15 - Time 1 Memory measures	166
Table 16 - Time 1 Memory measures	166
Table 17 - Time 1 mean scores for the stress and anxiety tests	167
Table 18 - Time 1 mean scores for the stress and anxiety tests	167
Table 19 - Training-related measures	168
Table 20 - Interval between testing	169
Table 21 - Time 2 memory measures	170
Table 22 - Time 2 stress and anxiety measures	171
Table 23 - Volunteer characteristics	180
Table 24 - Cognitive tests	181
Table 25 - The routes in the virtual reality navigation task	184

Chapter 1

Does experience change the brain?

1.1 Introduction

Whether the structure of the human brain can change in response to experience is a challenging and long-standing question of great interest to neuroscientists, educationalists and popular culture alike. The issue of structural plasticity of the human brain is fundamental to understanding if cognitive abilities are dictated by innate hard-wiring, or if they are amenable to environmental stimulation, so called nature versus nurture. Another manifestation of this debate comes in the guise of concepts such as 'talent' and 'expertise' [1-3]. It has been suggested that expertise is acquired, involves accrual of considerable knowledge, deliberate training and requires the individual to be able to monitor and control cognitive processes in order to perform a set of tasks competently and effectively [2]. Others see expertise as being innate, where training merely enhances innate hard-wired talent by permitting higher levels of expertise to be achieved and at a faster rate [4, 5].

Media hype and an increasing market for 'brain training' and 'brain gym' gadgets that claim to improve cognitive abilities, demonstrates the wide public interest in whether experience changes the brain [6]. This is unsurprising, as the answer to this question will have universal relevance by informing what each of us can hope to achieve, and by providing much-needed insights into the basis of individual differences. In addition, there could be important implications for patients with cognitive impairments. If we are to approach rehabilitation in a systematic and efficacious way, then it is vital to know if a particular cognitive system has the propensity for plasticity in adulthood, the limiting factors on such plasticity, and the timescales of any plastic change.

In general, however, there is a paucity of systematic neuroscience research into structural brain plasticity in humans and its relationship to cognition. This is

understandable given the near impossibility of bringing traditional microscopic techniques to bear on this issue in vivo in humans. Magnetic resonance imaging (MRI) [7], however, now enables in vivo examination of human brain structure (see Figure 1). Over the last two decades improvements in image acquisition, image quality, and spatial resolution permit detailed analysis of grey matter, white matter, and even specific fibre tracts [8]. In parallel with these advances, methods for quantitative data analysis of structural MRI brain scans have also progressed. On such approach is the region-of-interest (ROI) method [9] where an area of interest is manually delineated, and the total volume (in mm³) is obtained by counting the number of pixels in each ROI and multiplying them by the appropriate voxel size (determined by the resolution of the MRI scan). The main issue with this approach is that it is extremely time-consuming (and can also introduce user bias). Consequently, automatic segmentation of cortical and sub-cortical regional volumes in programmes such as Freesurfer [10] is becoming more widespread owing to the improvement in accuracy of tissue classification in recent years and speed of processing this approach offers. The most-widely used automated whole-brain analysis method is voxel-based morphometry (VBM) [11-13], implemented in the software package Statistic Parametric Mapping (SPM) [14]. This involves a mass-univariate voxel-by-voxel analysis of the whole brain, grey and white matter, and requires no *a priori* regions of interest. VBM operates within the framework of the general linear model, and thus allows groups of participants to be compared in terms of their brain structure, longitudinal studies within subjects, and the examination of the relationship between grey or white matter volumes and cognitive measures.

Structural MRI's fine resolution of soft tissue and superior anatomic resolution compared to other imaging techniques (such as computerised tomography) means it is now the mainstay in the clinical setting for identifying and evaluating brain lesions.

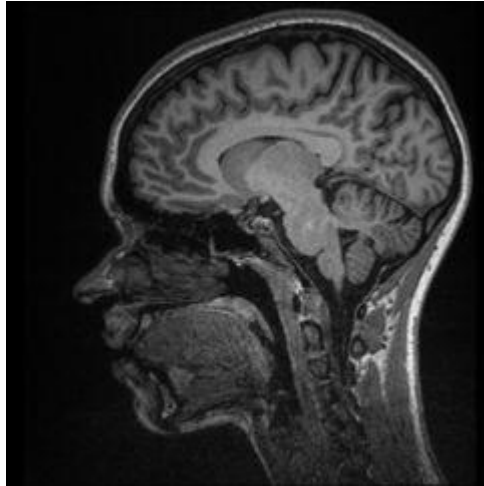


Figure 1. Example of a sagittal view from a structural MRI scan.

Moreover, in the realm of neuroscience research, the location and size of lesions/atrophy, and their relationship with cognition has been widely assessed. An elegant example of this is a study that examined memory in patients who had damage to the fornix following surgery for a colloid cyst [15]. A strong correlation was found between fornix and mammillary body volumes. Moreover, the results showed that volume of the mammillary bodies correlated with performance on tests of memory recall but not so with tests of recognition.

As well as examining the effects of clearly apparent brain lesions on cognition, MRI has revealed more subtle structural abnormalities in a range of disorders. For example, in schizophrenia, structural deficits in a wide variety of brain regions, including the hippocampus, have been identified, with variation between subsyndromes in terms of structural changes [16]. Volume decreases in a number of regions, also including the hippocampus, have been associated with depression [17]; decreased volume in the right posterior hippocampus has been observed in blind individuals [18, 19] but see [20]; increases in grey matter in medial and dorsolateral frontal areas, and in the parietal lobes, alongside decrease in frontal, parietal, temporal and occipital white matter have been found in autism [21]; and reduction of hippocampal head volume in posttraumatic stress disorder [22]. This knowledge has helped advance the

understanding of the biological factors implicated in a realm of neurological and psychiatric disorders.

The fact that MRI can be used with children, and longitudinally, has also made it possible to begin to construct quantitative maps of normal child brain development *in vivo*. Developmental trajectories in relation to grey matter development [23], IQ and cortical development [24], gender and cortical development [25], have been established and knowledge is expanding rapidly [26, 27]. In addition, MRI has proved invaluable to detect abnormal brain development, which has been linked to neurodevelopmental disorders such as ADHD, Autism, and childhood-onset schizophrenia [28-30].

While the utility of structural MRI is in no doubt for patients and in development, another major advantage of structural MRI brain scanning is the ability to examine 'normal' healthy adult brains, and to consider how brain structure relates to cognitive variables. Examining healthy adult participants eschews confounds such as medication effects, structural and functional re-organization after brain injury, covert damage in patients, and structural brain changes that occur through development. Studies of healthy adults have found differences in brain structure relating to IQ [31, 32], gender [33, 34], handedness [34, 35], as well as mapping out patterns of grey and white matter changes during normal aging [36-38]. Studies combining MRI techniques and genetic profiling have found genetic variability to modulate performance on a range of tasks including those assessing memory and hippocampal function [39, 40].

1.2 Structural effects of expertise in the healthy adult human brain

While it is interesting to consider the brain correlates of background characteristics such as gender and handedness, the main interest of this thesis is in structural brain plasticity in response to experience or external stimulation, and what this can tell us about the functions of particular brain regions. The most obvious way to examine this is in relation to acquired skills or expertise. Skill or expertise-related structural plasticity has been reported in numerous groups including musicians [41-45], typists [46], golfers [47], mathematicians [48], bilinguals [49], dancers [50], and novelty-seekers [51]. Many of these studies have also observed a positive correlation between the volumetric change in specific brain structures and the time spent learning and practising their specialization, e.g. increase in the grey matter of the left inferior parietal cortex of bilinguals is more pronounced in early bilinguals compared to late bilinguals (see [49]); professional musicians show a greater increase in grey matter volume in motor and auditory areas [42] and a prefrontal region [43]; in professional dancers the volume of the right anterior hippocampal formation correlated negatively and the right posterior hippocampal formation positively with the amount of time spent training for ballet/ice dancing or slacklining at the time of the study [50]. Such correlations suggest an association between skill acquisition and brain structure, however, the studies were all cross-sectional, and within-subject effects could not be examined. Moreover, in many of the participants, skill acquisition began during childhood, thus conflating brain changes during development with learning and practicing of the skills.

One way to more directly examine the effect of external stimulation on brain structure is to conduct longitudinal studies, where learning and outcome can be measured at different time points within individuals. Repeated measures designs are increasingly popular in the clinical literature but despite their usefulness, only a relatively small

number of studies have used this methodology in healthy populations. This is no doubt due to the time-consuming nature of longitudinal experiments. In one such study, young adults learned to juggle over a three month period and were scanned before and after training [52]. Increases in grey matter volume were observed in the intraparietal sulcus, and the human movement territory (hMT). A third scan was acquired three months after they stopped training which showed the grey matter changes had reversed [52]. Similar findings were also demonstrated in medical students during learning for final exams [53] (see also [54]). More recently, Boyke et al.[55] found that elderly people that learnt to juggle showed grey matter volume changes in the hMT and area V5. By the same token, Hyde et al. [56] followed children through musical training over a period of fifteen months and found that improvements in musically relevant motor and auditory skills correlated with structural brain changes. These studies show that the brain is amenable to structural plasticity in response to stimulation. However, the majority of these studies were focused in the motor domain [57], looking at changes over short timescales, and in children and students, where the latter may still be on a developmental trajectory [58]. Furthermore, the structural brain changes were mainly observed in cortical areas.

Consequently, studies of the long-term effects of the acquisition of expertise during adulthood that involve higher cognitive processes in individuals of average IQ, are still lacking. There is also a lack of consistent evaluation of the functional consequences of skill acquisition. These may include knock-on cognitive benefits but also possible detrimental effects – in terms of brain structure and cognition (see more on this in Section 1.5). Finally, we remain ignorant about what happens to the brain in adulthood once an individual stops using a skill that had been deployed over many years.

Compared with humans, structural plasticity has been studied more extensively in other mammals [59, 60] and in the song-system of birds [61-63]. However, the majority of

animal research has focused on the hippocampus, perhaps because it is widely accepted that neurogenesis (the production of new neurons – see more on this in Section 1.7), a candidate mechanism for plasticity, continues throughout adulthood in this structure. As the main focus of this thesis is the plasticity of the human hippocampus, I will briefly describe the anatomy of the hippocampus and pertinent background details before discussing hippocampal plasticity.

1.3 Hippocampal anatomy

The hippocampus is phylogenetically one of the oldest structures in the mammalian brain. The anatomy of the hippocampus has been most thoroughly studied in non-humans, however, there is a good degree of consistency across species concerning its neuronal architecture and connectivity (Figure 2). The hippocampus is divided into three segments along its long-axis: hippocampal head, hippocampal body and hippocampal tail. The structure of the hippocampus is the same in its different segments which are divided into sub-regions known as CA1, CA2 and CA3, in which the principal layer is made up of pyramidal cells. An adjoining area, the dentate gyrus (DG), comprises granule cells. Throughout this thesis I will use the anatomical terminology of Amaral [64] where 'hippocampus' includes CA1-CA3 and the DG (Figure 2). Other neighbouring regions include the subiculum, presubiculum, parasubiculum, and entorhinal cortex (ERC). These regions are not the focus on this thesis but are mentioned for completeness.

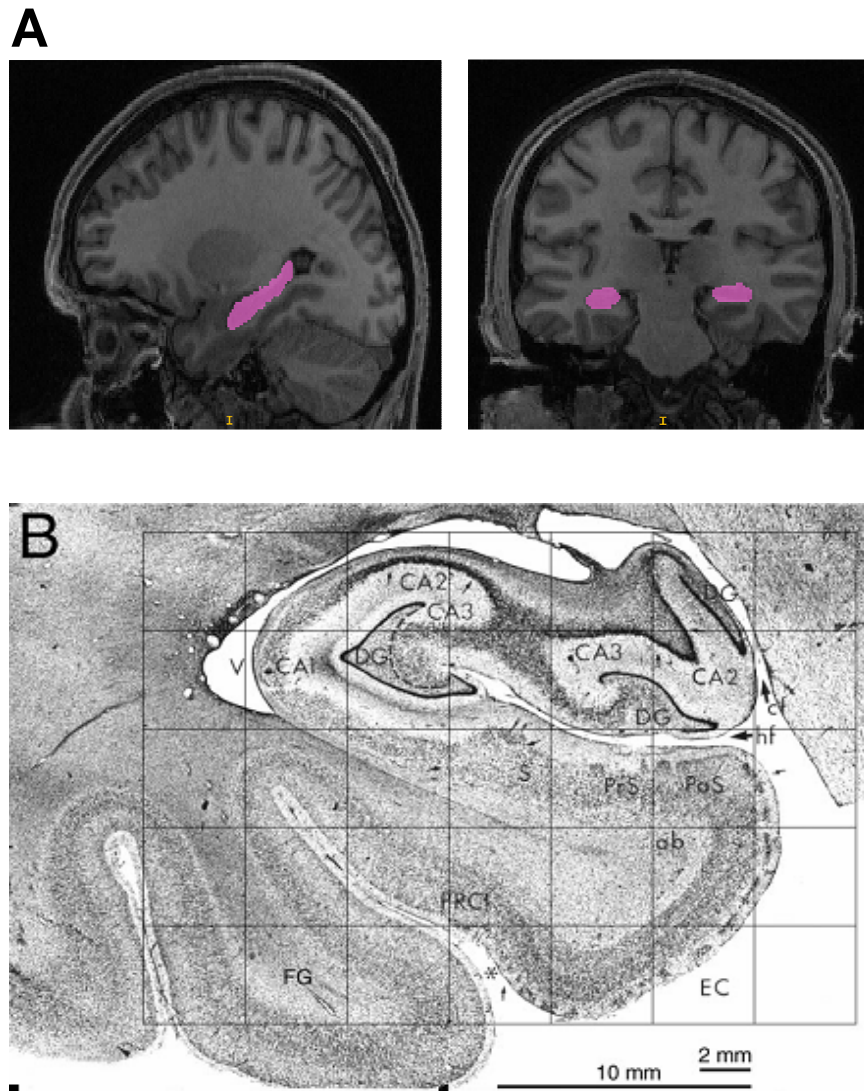


Figure 2. The human hippocampus formation

A) The position of the human hippocampus shown on a structural MRI scan. Sagittal view of the right hippocampus (left upper panel); Coronal view of the left and right hippocampus (right upper panel). B) Nissl-stained coronal sections through the human hippocampal formation. (From [64]).

The hippocampus embodies three connected afferent pathways known as the trisynaptic loop. This circuit, unlike connections in other brain regions that are bidirectional, is mainly unidirectional. For example, the ERC signals information to the granule cells of DG, which connect with pyramidal cells of CA3, then to field CA1, and finally to the subiculum. Yet CA3 cells do not project back to the granule cells of the DG, rather information is fed back to the entorhinal cortex through the subiculum, as

well as to the contralateral hippocampus via the associational commissural pathway (Figure 3).

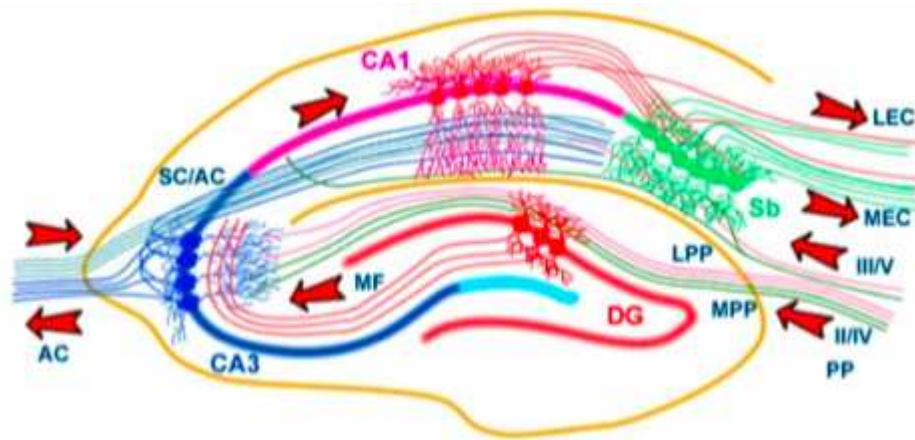


Figure 3. Schematic representation of the hippocampal network.

Input from the ERC forms connections with the DG and CA3 pyramidal neurons via the perforant path. CA3 neurons also receive input from the DG via the mossy fibres. These in turn send axons to CA1 pyramidal cells via the Schaffer collateral pathway, as well as to CA1 cells in the contralateral hippocampus via the associational commissural pathway. CA1 neurons also receive inputs direct from the perforant path and send axons to the subiculum (Adapted from MRC Centre for Synaptic Plasticity web page <http://www.bristol.ac.uk/synaptic/pathways/>).

Cortical input to the hippocampus

The hippocampus receives inputs from parietal, temporal and occipital cortices via the ERC. For example, the parietal cortex connects to the ERC through the parahippocampal cortex. The temporal cortex reaches the ERC through the perirhinal cortex. Thus, the hippocampus receives information that has already been processed from the perirhinal cortex, the parahippocampal cortex and entorhinal cortex. The ERC is made up of six distinct cellular layers and divided into two regions Lateral entorhinal area (LEA) and the medial entorhinal area (MEA).

The axons of the perforant path arise mainly in layers II and III of the ERC. Inputs from projections of layer II and layer III of the ERC are combined in the CA1 and subiculum, these then project back to the ERC in a particular fashion, septal CA1 to the LEA and more temporal areas to MEA. In contrast projections to the CA1 remain segregated.

The LEA projects to the distal portion of CA1 and the proximal part of the subiculum, whereas the MEA projects to the proximal CA1 and subiculum. Axons from layers II and IV of the ERC project to the DG and CA3 [73] (Figure 3). The feedback efferent projections from the hippocampus via the ERC to the perirhinal and parahippocampal cortices are largely reciprocal to the input projections [363].

The ERC therefore has the crucial role of being the main direct input to the hippocampus from cortical areas, with the subiculum acting as the main output.

Fornix system

The hippocampus also receives inputs via the fornix fibres arising from the medial septal nuclei, diagonal band of Broca, and anterior thalamic nucleus which convey important neuromodulatory signals (e.g. acetylcholine) (see: [357]). Moreover, disruption of the hippocampal outputs via the fornix to the anterior thalamic nuclei and mamillary bodies is also known to induce significant loss of episodic memory capacities, in particular recall, in humans [15, 358]. Consequently, hippocampal outputs via the fornix represent an additional anatomical pathway to that directed through the ERC.

Internal circuitry of the hippocampus

The Dentate Gyrus

The DG projects via the mossy fibres to the CA3 region. The DG does not project to any other areas of the hippocampus. Each neuron within CA3 receives input from a small number of granule cells (approximately 70). Nonetheless, the mossy fibre projections are deemed to be powerful owing to the distinctive characteristics of its synapses i.e. the presence of large dendritic spines in multiple active zones located

near the soma of CA3 neurons. A single (or few) mossy fibre (s) is able to trigger the firing of a CA3 neuron [361].

CA3 Subfield

CA3 neurons receive input from three sources: 1) the DG; 2) the perforant path input from layer II of the ERC; 3) Recurrent connections from other CA3 neurons. CA3 neurons in this recurrent projection can connect with other CA3 neurons throughout the hippocampus. This system can be considered as one in which associational connections are more numerous.

The Schaffer collaterals constitute the major source of input leading from CA3 to CA1. Given that there are fewer CA3 neurons (3×10^5) as compared to CA1 neurons (4×10^5), this connection is divergent with information spreading out, following the convergence previously effected by the DG (10^6 neurons) to CA3 projections. Additionally, a single CA3 neuron may contact CA1 cells throughout most of the length of the hippocampus.

CA1 and subiculum

The CA1 region projects to the subiculum and to the deep layers of the ERC (layers V,VI). As such, inputs from the ERC to the hippocampus originate primarily in the superficial layers, and outputs from the hippocampus are directed to the deep layers. The superficial and deep layers of the ERC are connected by intrinsic associational connections, allowing the hippocampal circuit to form a complete loop. It is possible that this closed loop may support the reverberation of information over short delays within the hippocampus. Efferent pathways from the ERC largely reciprocate the pattern of inputs, with projections from the ERC to the perirhinal and postrhinal cortices, and back to neocortical association areas.

Place cells

In 1971 O'Keefe and Dostrovsky made recordings from pyramidal cells in the hippocampus of freely moving rats. They discovered that these neurons showed location-specific firing. That is, a neuron fired when the rat was in a certain part of the environment, irrespective of the animal's viewpoint, giving rise to the name 'place cell'. The location where each place cell fires is called place field. Place fields emerge rapidly upon exposure to a novel environment, once formed they remain stable over long periods of time [65, 66]. Place fields vary not only with the shape and size of the environment in which testing takes place but also with the anatomical location of the recordings in the hippocampus. Place cells rarely have more than one field in a single environment and the sizes of the place field vary along the long axis from dorsal to ventral. Dorsal hippocampal place fields are smaller, and can generally be seen along the edges and centre of open platforms, or evenly distributed in closed cylinders, rectangular boxes [67] and mazes [68] (Figure 4). It has been suggested that the difference in size of the place fields may serve to distinguish between two different environments, being mediated by the larger fields of the ventral hippocampus, or two different locations within in any one environment, mediated by the dorsal hippocampus [69, 70]. In addition, place fields of neighbouring cells are not closer to each other than would be expected by chance for cells located at any two points in the hippocampus [71] but see [72].

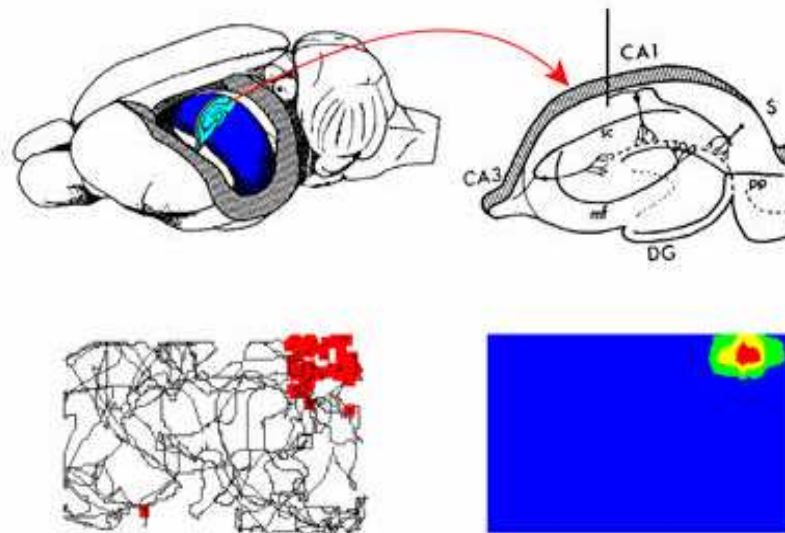


Figure 4. Hippocampal place cells.

Rodent brain exposed (upper left panel) with schematic diagram of electrode within CA1 region (upper right panel). Traversals by rat in rectangular environment spike locations (red) are superimposed on the animal's trajectory in the recording enclosure (black) (lower left panel) with rate map of a place cell firing in top right corner of enclosure (lower right panel). Colour coding: red indicates a large response; yellow, moderate response; green, low response. Adapted from O'Keefe Lab web page (www.anat.ucl.ac.uk/research/okeefe/okeefe_research.html).

The hippocampus has two main modes of electrical activity, theta, which is a regular oscillation, and large irregular activity (LIA). Based on the observation that there is a period of a few seconds between the onset of a sharp wave and LIA associated behaviour (e.g. immobility), LIA has been proposed to be the 'at rest' mode of the hippocampus. Alternative functions such as strengthening synaptic modifications that took place immediately preceding LIA periods may play a role in memory consolidation have also been proposed [73]. The theta mode occurs during states of active behaviour (e.g. locomotion) and during REM sleep. In hippocampal place cells a phenomenon referred to as theta phase precession occurs, in which the cell fires late in the theta cycle when an animal (rat) enters the place field of the cell, and fires at earlier phases as the rat moves through the place field [362]. Several functions have been suggested for hippocampal theta, such as binding together activity in different cortical areas, and acting as a clock signal against which single spikes in pyramidal cells can be timed. The relationship between the firing of single spikes and theta recorded from

the hippocampus of freely moving rats provides a robust example of phase-coding of a cognitive variable (location) in the mammalian brain [74].

1.4 The functions of the hippocampus

The human hippocampus has been acknowledged to play a critical role in memory since Scoville and Milner [75] reported the seminal case of patient HM. Following bilateral removal of much of his temporal lobes for the relief of intractable epilepsy, HM developed a severe and permanent post-operative amnesia while his general intellectual abilities remained relatively intact. Whilst his case highlights the vital importance of the temporal lobes in supporting memory, H.M.'s lesions were extensive [75-78] and therefore did not allow conclusions to be drawn regarding the precise mnemonic roles of areas within the temporal lobes. In spite of this, his pattern of spared and impaired function allowed the proposal of multiple memory systems in the brain [79, 80]. Long-term memory was proposed to comprise the declarative memory system mediated by the medial temporal lobes, which encompasses memories for facts and events, and the non-declarative system which includes conditioning, priming, and procedural memory, held to be independent of the medial temporal lobes. In the 50 years since HM, other patients have been reported with more selective hippocampal damage [81-84], and a number of other theories of memory and the hippocampus have been put forward [85, 86].

Theories of hippocampal function share some commonalities, but there are also important differences. Disagreements concern the nature of the memory representation, e.g. spatial or non-spatial [87], the duration of hippocampal involvement in representing memories, e.g. short-term or long-term [79], and aspects of mnemonic processing such as recall/recognition and familiarity [88-90], pattern separation/pattern

completion [73, 91-93]. Below I outline in brief the main theories of hippocampal function and consider some of the evidence supporting each (for other theories see [72, 94-96]).

The discovery of 'place cells' [87, 97] led to the development of the cognitive map theory of hippocampal function [97]. This posits that a fundamental function of the hippocampus is the construction and maintenance of spatial maps of the environment. This representation contains information about places in the environment, their spatial relations, and knowledge that specific objects are located in specific places. It allows the animal to locate itself in a familiar environment without relying on a specific input, such as an object, and link together different parts of the environment that may not have been experienced at the same time, allowing the animal to take novel routes. The system that enables this type of flexible navigation was coined the "locale system". The theory suggests that the hippocampus stores an allocentric representation of the environment that allows for flexible wayfinding. This cognitive map of the environment is thought to be distributed possibly along the longitudinal and transverse axes of at least the dorsal hippocampus [97-99]. The cognitive map theory suggests that other non-hippocampal regions form the "taxon system" that relies on egocentric (i.e. body centred) spatial representations to support less flexible navigational strategies, such as those that would be employed when heading towards a salient landmark.

Since the discovery of place cells, it has been shown that lesions of the hippocampus impair a rat's ability to navigate, for example in the now classic task the Morris water maze, where a rat has to remember the location of a hidden platform in a milky swimming pool [100]. The cognitive map theory argues that the fundamentally spatial role of the hippocampus in animals extends to humans, and may have evolved to be the basis of more general event or episodic memories, which incorporates both time and language. This is suggested to be mediated by a lateralisation of function, with the

right hippocampus implementing spatial and episodic tasks, and the left hippocampus involved in language and verbal memory [97]. Thus, the cognitive map theory advocates that the role of the human hippocampus in memory is to represent the spatial and temporal context in which events occur [101].

Other theories disagree with the premise that the hippocampus is fundamentally specialised for space, the alternative suggestion being that it merely appears so because spatial tasks are the easiest sort to administer to animals. According to Cohen and Eichenbaum [86] the hippocampus processes all manner of associations and sequences of events, and spatial navigation is merely one example of the way in which the hippocampus is specialised for “relational processing” [102]. They argue that the hippocampus allows for the formation of flexible representations of the environment that are later consolidated in the neo-cortex, where they can be accessed without input from the hippocampus. The hippocampus is suggested to support encoding, consolidation and only short/intermediate memory processes including retrieval for any type of relational information [103]. Patients with deficits in associative memory across domains appear to offer support for the relational theory [104]. Functional MRI studies have found hippocampal activations for associative memory that do not seem to involve the spatial domain [105] whilst others have only found activation in the hippocampus during spatial relational tasks [106]. More recently researchers have carefully evaluated the type of material employed when testing patients. Mayes et al ., [104] showed that patients have spared associative recognition only when the pairs are made of the same type of items (e.g. face-face) but remain impaired when the pairs are made of different types of items (e.g. word-face). Whilst others argue that recognition is spared when the elements of an association can be 'unitized' (i.e. processed as a single integrated unit), this 'unitized' encoding relies on cortical learning mechanisms more so than what typically occurs, and thereby enhances the familiarity strength of intact pairs [284].

One of the most influential theories of hippocampal function has been the consolidation theory [107]. The belief here is that the hippocampus plays a short-term role in memory acquisition, and once consolidated, it is transferred to the neocortex and can be retrieved without the involvement of the hippocampus. This is held to be the case for all types of information, including episodic, spatial and sensory. Patients with bilateral hippocampal damage who have limited retrograde amnesia and preserved memory for life events that occurred many years ago offer support to this view [108, 109], as do two cases of patients who were able to navigate in environments learned long ago [110, 111]. However, there are also hippocampal damaged patients whose amnesia does not show a temporal gradient, with remote and recent memories being similarly compromised [109, 112]. In terms of spatial memory, assessment of a London taxi driver (TT) of 40 years experience who sustained bilateral hippocampal damage revealed significant navigation impairment along London's detailed road network [113]. This suggests the hippocampus is necessary for spatial memories in perpetuity (Figure 5).

In contrast to the consolidation theory, the multiple trace theory [85] suggests the hippocampus is necessary for memory retrieval throughout the lifetime, and points to patients with extensive retrograde memory loss as supportive evidence. They argue that memory undergoes consolidation/re-consolidation in which multiple traces of a past event are formed and differentially located along the longitudinal axis of the hippocampus. These traces in turn form different associations with other information that reflect their distinct sources. Therefore remote memories are represented by a larger number of traces making them more accessible than recent ones and this may be why some amnesic patients show a temporal gradient [114].

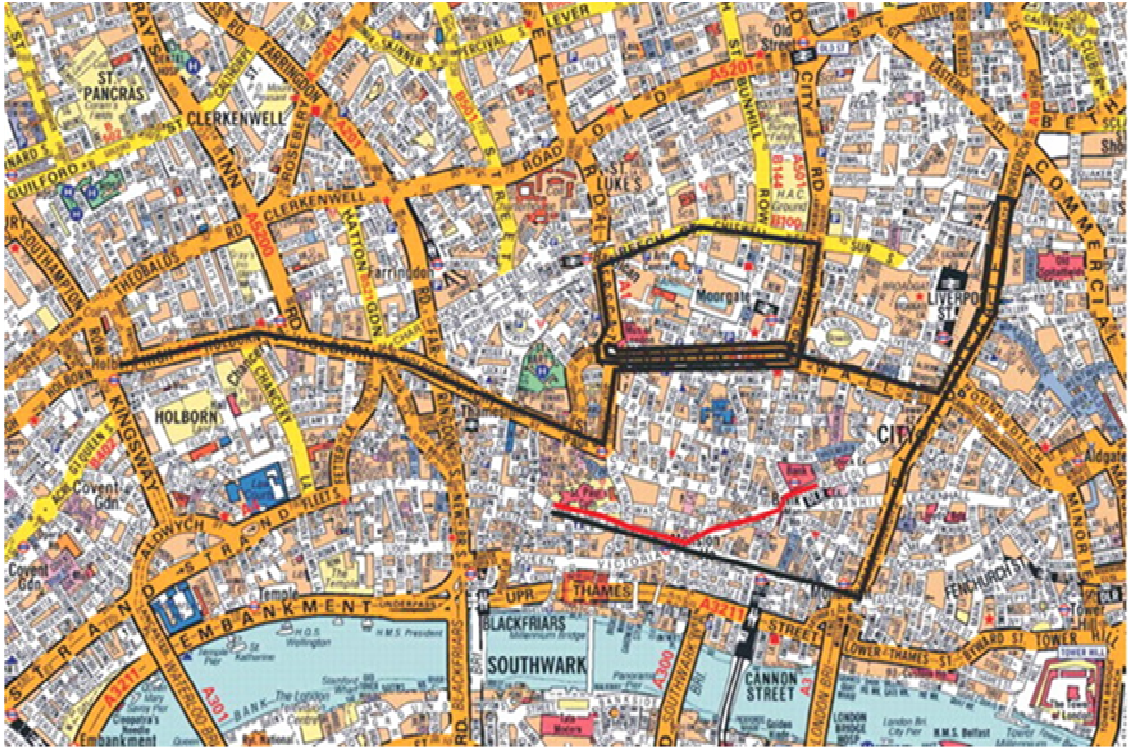


Figure 5. Navigation performance.

An example of navigation performance for a Route in central London (St Paul's Cathedral to the Bank of England). All 10 control subjects performed in an identical manner, shown in red on this map. TT was grossly impaired; his route is shown in black. Map reproduced by permission of Geographers' A-Z Map Co. Ltd © Crown Copyright 2005. All rights reserved. Licence number 100017302. From Maguire et al., 2006 [113].

Gilboa et al. [115], found activation for remote memories to be distributed along the long-axis of the hippocampus whereas activation for recent memories was clustered in the anterior region of the hippocampus. In accordance with the multiple-trace theory the authors suggest this may be why lesions to the hippocampus disproportionately affect recent compared to remote memory. Whilst appearing to account for a good deal of the extant data, the computational efficiency of multiple traces is questionable, and recent fMRI analyses using multivariate pattern classification calls into question the existence of multiple distinct traces for repeatedly retrieved memories [72, 116].

All of these theoretical accounts share a common emphasis of the key role the hippocampus plays in memory, although are in disagreement as to whether the primary role of the hippocampus is spatial (cognitive map theory), long-term (multiple trace

theory), or associative and time limited (relational, consolidation theories). Given this, the question remains as to whether the hippocampus has a bias towards spatial representations, and whether its involvement is of a long duration. In this thesis I do not seek to adjudicate directly between hippocampal theories. However, examining the nature of hippocampal structure-function relationships, and potential stimulants for hippocampal structural plasticity provide another means of contributing to the debates that are engendered by these different views (and I will consider this further in Chapter 7).

1.5 Structural changes in the hippocampus

Volumetric analyses of the hippocampal region in birds and other small mammals have shown that the size of the hippocampus correlates with whether or not the animal engages in tasks that place a demand on spatial memory. For example, animals that store food and have to remember the location of their caches have a larger hippocampus than their non-food storing counterparts (see [117] for a review). Hippocampal volume differences are present even between animals of the same species, depending on the degree to which an animal shows storing behaviour [118]. Even more striking is that some species show volume differences in the hippocampus specifically during seasons where the demand for spatial memory is greatest [119, 120] i.e. the memory load during food-storing periods coincided with an increase in hippocampal neurons. Migratory birds exhibit larger hippocampus volumes compared to more sedentary counterparts [121], as do homing pigeons that acquire navigation expertise compared to birds restricted to indoor flying [122]. There is evidence of similar patterns in mammals such as voles [123], kangaroo rats [124, 125], and gerbils [126]. Taken together these findings suggest that as spatial memory load increases so does the volume of the hippocampus, be it within or between subjects.

Structural imaging of humans with neuropsychiatric disorders has shown variations in size of the hippocampus. Smaller hippocampal volumes have been reported in epilepsy, and the reduction is concordant with the side of the epileptogenic focus [127, 128]. In Alzheimer's disease the loss of hippocampal volume is a hallmark of the disorder [129, 130]. Hippocampal volume reduction is also present in schizophrenia [131], and the size of the reduction has also been related to symptom severity [132]. Volume reductions have also been observed in autism [30] even in high functioning individuals with normal IQ [133]. Studies of both first-episode and multiple episode depression also show reductions bilaterally in the hippocampus [134, 135].

In healthy populations longitudinal assessment of aging and the hippocampus has shown shrinkage in this area that increases with the passage of time [136], and volume decreases correlate with delayed recall memory performance [137]. Gender differences have been reported, with larger hippocampi reported in males [138], while others have failed to find differences [139]. During navigation tasks the choice of strategy by participants has been related to grey matter volume in the hippocampus, with participants who used a spatial memory strategy (learning relationships between environmental landmarks) having greater hippocampal grey matter volume [140]. Overall, however, a large meta-analysis of studies involving memory tasks and their relationship with hippocampal volume in healthy participants found extreme variability across studies, making firm conclusions about the relationship between memory and hippocampal volume in healthy participants difficult [280].

The volume of the hippocampus has also been related to the presence of specific gene polymorphisms associated with memory. Substitution of the valine (VAL) by methionine (MET) at the 66th codon pair in the gene for the precursor peptide for brain-derived-neurotrophic-factor (BDNF) has been associated with smaller hippocampus volumes [141, 142], but evidence for global effects on brain volume of this

polymorphism also exist [143]. Moreover, the MET allele has been associated with poorer episodic memory, differential hippocampal activation (fMRI), and lower hippocampal n-acetyl aspartate (NAA), assayed with MRI spectroscopy [144, 145]. In a twin study, genetic and environmental influences on hippocampus size were compared [146]. From this experiment it was concluded that most of the variance (more than 60%) in hippocampal size was actually attributable to non-shared environmental components.

Hence there is evidence that the structure of the human hippocampus is affected by a range of pathologies, some involving compromised memory. The evidence for a direct relationship between hippocampal volume and memory performance in healthy participants is less clear, and there appears to be both genetic and environmental influences on hippocampal structure. Evidence from animal studies point to the plasticity of the hippocampus associated with spatial memory. Whether this is the case in humans, is unknown. The main purpose of this thesis is to understand how the structure of the hippocampus adapts to cognitive demands in healthy humans. In animals hippocampal volumetric changes associated with spatial memory were evident in naturalistic settings. Therefore the assessment of memory and navigation in a naturalistic setting would seem to be of direct relevance to our understanding of human hippocampal plasticity. In this thesis a unique and naturalistic learning situation was employed to examine in healthy human adults the relationship between hippocampal volume and memory.

1.6 Licensed London taxi drivers – a model system

Training to become a taxi driver in London (UK) culminates in a licence and green badge from the Public Carriage Office, a government body that regulates and licences taxis in London. Candidates who wish to become taxi drivers may only apply to do so if

they are above the age of 21 years and medically fit and healthy. There are no educational pre-requisites to become a taxi driver. London has a highly complex and irregular layout. In order to learn its intricate detail, candidates undergo extensive training over a period of 2-4 years colloquially known as acquiring 'The Knowledge'. This involves the impressive feat of learning the layout of over 25,000 streets in the city (see Figure 6), and the locations of landmarks and places of interest within a six mile radius of Charing Cross station (official central point to define the geographical scope of London), leading to a stringent set of examinations, which one needs to pass, in order to obtain an operating licence. Some trainee taxi drivers attend a 'Knowledge school' to help them with their training. Initially, they learn 320 routes set out in the 'Guide to Learning the Knowledge of London' published by the Public Carriage Office (known also as 'The Blue Book'). Trainees learn these routes using different kinds of maps, by travelling the routes over and over (usually on small motorbikes), and by 'calling over' which involves verbally repeating the routes. The latter is generally done with a fellow trainee able to point out any mistakes. By the end of their training, trainees will have learnt more than 600 routes in the city of London, a large number of routes into the suburban areas of London, and thousands of places of interest. Once trainees have reached this level of detailed knowledge they will sit a number of oral examinations known as 'appearances'. During these appearances trainees are examined individually by a Public Carriage Office examiner. At each appearance the trainee is asked to indicate the shortest possible route from A to B, with the examiner asking a random set of routes that can be located anywhere within London (see Figure 7). On average trainees will sit between 12-20 appearances.

Thus, licensed London taxi drivers are a group of individuals who have trained from scratch to become expert navigators, they all know the same environment, and have all been tested in a similar fashion by an independent body, so that the acquisition of expertise is verifiable.

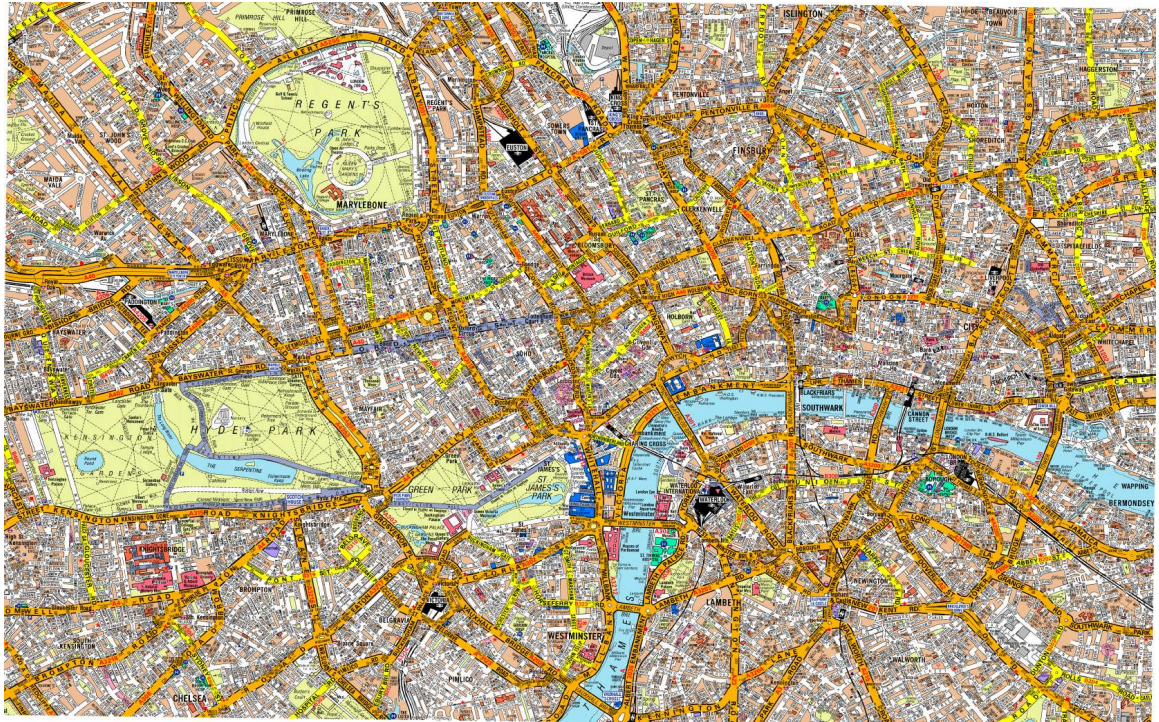


Figure 6. Map of London.

Map of London (note – this does not include all the street taxi drivers are required to learn). Map reproduced by permission of Geographers' A-Z Map Co. Ltd. © Crown Copyright 2005. All rights reserved. Licence number 100017302.

That there are ~25,000 licensed London taxi drivers registered currently with the Public Carriage Office suggests that many people are able to acquire this expertise, although it is worth noting that Public Carriage Office figures suggest that up to 60% of trainees fail to qualify.

Maguire et al. [147] examined whether there were parallels between London taxi drivers and findings of volume differences in the hippocampus of spatially active animals. They compared the structural MRI brain scans of a group of licensed London taxi drivers with those of non-taxi driver control subjects. They used two different techniques to assess hippocampal volume. The first was a traditional region of interest approach, where the hippocampi were manually traced on the MRI scan, and their volumes calculated. The other technique used was voxel based morphometry (VBM) [11].

The qualification process is divided into 3 stages:**Stage 1 – Self assessment**

This takes the form of a written examination. Candidates mark their own assessments by comparing their answers with the correct ones displayed on a screen. At this stage candidates may apply to take the formal written examination. The written examination must be undertaken within two years of the self-assessment.

Stage 2 – Written examination

This examination requires candidates to write out five specified routes within a six mile radius of central London. This is followed by five maps showing the start or end points of five different routes where road names have been removed. Candidates must indicate which part of London it is, this is followed by locating a list of roads and points of interest on the map. The total mark for both sections is 100. The pass mark is 60%.

Stage 3 – Oral examinations

The oral examinations are sub-divided into three stages (3 to 6). Candidates must gain a total of 12 points at each stage to progress to the next one. The grades obtained during the examinations are translated into points as follows:

- Grade AA – 12 points (no mistakes)
- Grade A – 6 points
- Grade B – 4 points
- Grade C – 3 points
- Grade D – 0 points

If at any stage a candidate scores 0 points on four successive examinations (appearances), the previous stage has to be re-started. The time between oral examinations starts off as 56 days, lowering to 28 days, followed by 21 days in the final stage.

Finally, candidates undergo one suburban oral examination covering 25 routes that radiate from the edge of the six mile central London radius into the outskirts of Greater London.

Figure 7. Taxi driver qualification process.

Both techniques gave rise to similar results, namely that there was increased grey matter volume bilaterally in the posterior hippocampi of taxi drivers compared with the control subjects. By contrast, the taxi drivers had decreased grey matter volume bilaterally in the anterior hippocampus (Figure 8). Moreover, in the taxi drivers, the time spent taxi driving was positively correlated with volume of the right posterior hippocampus, and negatively correlated with anterior hippocampal volume. The above findings led Maguire et al. [147] to suggest that there may be a capacity for local plastic change in the structure of the healthy adult human brain in response to environmental

demands, and that in particular the right posterior hippocampus may be the storehouse of the spatial representation of an environment, able to expand to accommodate its elaboration. It is interesting that hippocampal volume correlates with years taxi driving, suggesting that this pattern of grey matter volume is not just associated with the initial learning and qualification phase. Indeed taxi drivers confirm when asked that post-qualification they continue to learn, and their spatial representation of London continues to cohere over time.

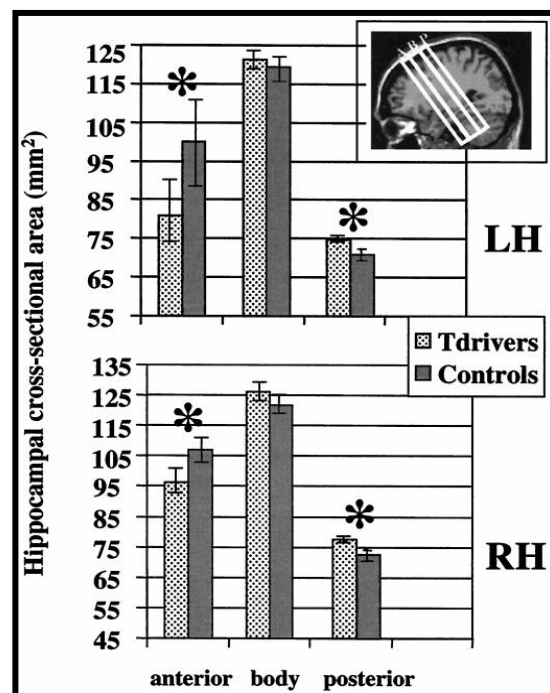


Figure 8. Hippocampal volume differences between taxi drivers and control subjects.

The orientation of the slices measured in the volumetric analysis with respect to the hippocampus is shown (top right inset). A= anterior; B= body; P= posterior. (Upper) The mean of the cross-sectional area measurements (uncorrected for ICV) for the three regions of the left hippocampus (LH) and right hippocampus (RH). Taxi drivers had a significantly greater volume relative to controls in the posterior hippocampus, and controls showed greater hippocampal volume in the anterior. There was no difference between the two groups in the hippocampus body. *, $p < 0.05$. (From Maguire et al. [147]).

Maguire et al. [148] also investigated the relationship between hippocampal volume and navigation ability in the general public using a virtual reality navigation task. Although there was variability in the performance of the participants, the study failed to find an association between hippocampal grey matter volume and navigational ability.

Likewise, Maguire et al. [149] explored whether there was a relationship between hippocampal volume and memory abilities in individuals with exceptional memory capabilities, namely participants in the World Memory Championships, held in London every summer, who are known for memory skills that are wide-ranging including the ability to memorize numerous decks of playing cards and long lists of random digits. Exceptional memory in these individuals was not found to be associated with higher IQ or any structural brain differences relative to control participants [149].

However, fMRI scanning revealed that they activated brain areas associated with navigation, including the hippocampus, more than control participants. The expert memorizers in general utilised a mnemonic strategy known as the Method of Loci [150]. With extensive practice this method is effective at enabling one to remember the order of stimuli by imagining familiar routes and placing the to-be-remembered items along the routes, thus promoting deeper encoding of the stimuli with rich associations. Unlike taxi drivers, however, the restricted set of routes they employed was not sufficient to effect structural brain changes.

Maguire et al. [151] went on to compare licensed London taxi drivers with another group of subjects who also spend all day navigating along the streets of London, in the traffic fumes, dealing with stress, and interacting with customers, namely London bus drivers. The two groups were matched for years experience driving in London, as well age and IQ. The main difference between the two groups was that the bus drivers learned and used a constrained set of routes compared with the large spatial representation of London learned by the taxi drivers. The findings from Maguire et al. [147] were replicated, showing taxi drivers had greater grey matter volume in posterior hippocampi and less grey matter volume in anterior hippocampus compared to bus drivers (Figure 9). Years navigating in London correlated with hippocampal grey

matter volume only in taxi drivers and not bus drivers, positively for the posterior hippocampus and negatively for the anterior hippocampus (Figure 9).

As well as comparing the taxi drivers with a more closely matched control group, the Maguire et al. [151] study extended previous findings by including a preliminary examination of the neuropsychological profile of taxi drivers. Unsurprisingly, relative to bus drivers, taxi drivers were significantly more knowledgeable about London landmarks and their spatial relationships (Figure 10). However, they were significantly worse at the delayed recall of a complex figure (Figure 10). This was in the context of learning and recognition memory for individual items being comparable with bus drivers, as were working memory, retrograde memory, perceptual and executive functions. This led to the suggestion that the anterior hippocampus, reduced in volume in taxi drivers, may be less efficient at forming associations, and that there may also be a limited capacity for further consolidation and storage in the posterior hippocampus, given its involvement in supporting the complex spatial representation of London. This suggests that there might be a cognitive trade-off in acquiring 'The Knowledge', and a limiting factor on the resources available for hippocampal plasticity. More broadly, this finding points to the need to consider not only the cognitive benefits but also the costs of acquiring a skill, as this could be just as informative in elucidating the functions of a brain region such as the hippocampus.

Not that much is known about the broader cognitive advantages or costs associated with skills and their concomitant neuroanatomy. A handful of cases have evaluated benefits that extend beyond those directly associated with the skill at hand.

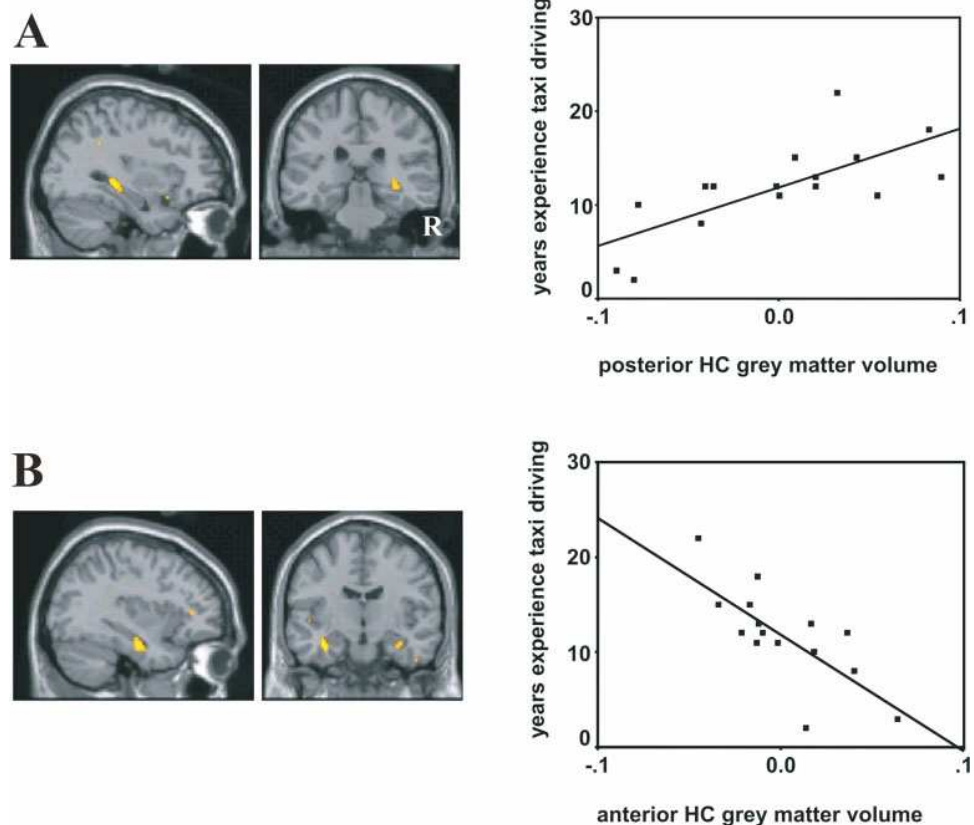


Figure 9. The effect of taxi driving experience on grey matter volume.

A) Sagittal and coronal sections (left and middle panels) showing right posterior hippocampal grey matter volume varies positively with time taxi driving. Right panel shows a plot of the correlation, with increasing grey matter volume the more years of taxi driving. B) Sagittal and coronal sections (left and middle panels) showing anterior hippocampal grey matter volume varies negatively with time taxi driving. Right panel shows a plot of the correlation (at the peak voxel in the right anterior hippocampus) with decreasing grey matter volume the more years of taxi driving (From Maguire et al. [147]).

Musicians have been shown to have greater brainstem responses compared to controls associated with both auditory and audiovisual processing of speech and music stimuli [152], enhanced perception of vocally expressed emotion [153], and music students have shown improved acoustic temporal novelty associated with stronger activation in the hippocampus [154], while bilinguals have shown greater efficiency of inhibitory mechanisms the higher the level of expertise [155].

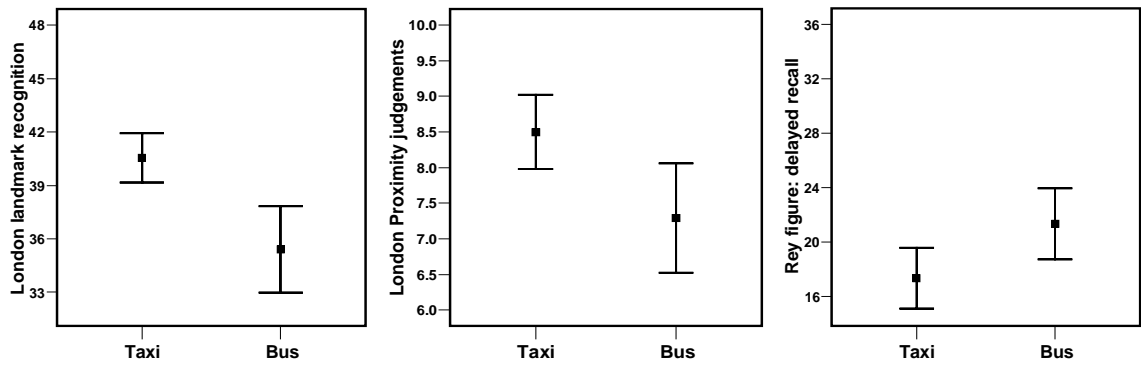


Figure 10. Main effects for the neuropsychological tests.

London taxi drivers were significantly better than London bus drivers at identifying London landmarks from among visually similar distractors (left panel) and making judgments about proximal relations between London landmarks (middle panel). London taxi drivers were significantly poorer at recalling newly-learned visual information (Rey-Osterrieth complex figure) after a delay than London bus drivers (right panel). Bars represent SE.

Expertise is typically viewed in a positive light. However, the potential costs of expertise also warrant consideration. As well as the Maguire et al. [151] finding of possible negative cognitive effects of being a London taxi drivers, there have been several case reports of individuals whose memories become so good that it actually causes them great distress as they are unable to forget anything [156, 157]. In another domain, some expert musicians suffer focal dystonia, a loss of control and degradation of skilled hand movements [41].

The neuropsychological findings from Maguire et al. [151] were preliminary and based on a limited battery of tests. The neuropsychological profile of taxi drivers needs to be much more fully characterised in order to ascertain if the positive and negative cognitive effects of their expertise replicate, and if so, to thoroughly explore their origin. Thus, is associative memory processing compromised across all domains in expert navigators or just in the spatial domain? Other key questions include: does being an expert navigator promote or impair learning the layout of new environments? Does it facilitate or inhibit updating the layout of an already well-known environment? Finally, after continued use of their expertise throughout their professional life, what happens to

their hippocampi and cognitive profile once taxi drivers retire? In this thesis I intend to bridge these gaps in our knowledge.

1.7 What are the underlying mechanisms of plasticity?

Given evidence of hippocampal plasticity in animals, and indirect evidence for skill-induced hippocampal volume changes from human cross-sectional taxi driver studies [146, 150], what mechanisms might underpin these volume changes? Plasticity could be the result of a number of processes, including proliferation of glial cells, synaptogenesis, and neurogenesis. All three processes are active in the adult mammalian brain. Glial cells continue to be produced, be it at a slow pace, throughout adulthood [158] and have been shown to increase in volume with the addition of synapses following learning [159]. Synaptogenesis, the formation of new synapses, has also been observed in neocortical areas of rats raised in enriched environments [160]. Neurogenesis, the birth of new neurons, is widely thought of being restricted to two regions, the subventricular zone (SVZ) of the lateral ventricle and the subgranular zone of the dentate gyrus of the hippocampus [178]. New neurons are generated in SVZ throughout adulthood and migrate through the rostral migratory system to the olfactory bulb in rats and non-human primates [161], and in monkeys it is also suggested that these cells migrate through fiber tracts to the neocortex [162]. However other studies have failed to detect neurogenesis in the neocortex of adult primates [163] and humans [164, 165]. A recent, postmortem analysis has shown migration of neuroblasts to the olfactory bulb in humans [166] yet the neuronal markers and methodology employed in the study have been criticised [167] and the general consensus in the scientific community is that neurogenesis does not occur in the neocortex. This scope of neurogenesis in the brain is an area of contention, nevertheless, there is agreement that neurogenesis occurs in the hippocampus

throughout the lifespan. In this section, I will limit my discussion to the role of dendritic changes and neurogenesis in learning and memory in relation to the hippocampus.

Dendritic changes

Synaptogenesis is a process that mainly takes place during development, beginning during embryogenesis and extending postnatally in the rat to approximately day 90. It can be seen in adults in the form of reactive synaptogenesis [168], and in response to exercise and environmental enrichment [169]. There is limited evidence of learning induced structural plasticity that involves the production of new synapses or reorganization of existing synapses, and what little there is, pertains mostly to the neocortex [159]. In terms of hippocampal synaptogenesis, exercise has been shown to increase spine density on dentate gyrus granule cells [169]. Others have found no effects of exercise but have reported increases in both cell numbers and dendritic branch numbers in animals in enriched environments [170, 171] and the number of synapses [172, 173]. Spine density and the size of dendritic trees continue to increase in aging animals living in enriched environments [174].

Leuner et al. [175] demonstrated that an increase in spine density on hippocampal pyramidal cells in CA1 occurs as a result of learning. Likewise hippocampal dependent learning has been shown to induce a specific form of synaptogenesis that leads to the formation of multiple-synapse boutons (synapse with more than one dendritic spine) and does not involve a net synaptogenesis [176, 177]. After long term potentiation (LTP - a form of synaptic plasticity induced by artificial tetanic (high-frequency) stimulation of an afferent pathway, and often associated with learning) a range of morphological changes take place in the hippocampus, such as synapse curvature, size of the active zone and spine volume, increases in synaptic length and appearance of perforations of the synapses [178]. A recent meta-analysis [179] suggests that the changes related to learning in the hippocampus are small and specific, with most of the

morphological parameters used to typically measure correlates of learning showing modification. For example, CA1 showed changes in apical spines but not overall synapse density. By contrast, area CA3 showed significant increase in synapse density [178]. A positive relationship between spatial learning and development of the dendritic arbour of new adult-born neurons in the hippocampus has also been observed in mice. Moreover, the rate at which this takes place is dependent on the cognitive demand on the hippocampus during learning [179]. Finally, spatial memory retention has been associated with increased numbers of mushroom spines in the dendritic spines of CA1. Mushroom spines are proposed to act a structural storage sites for long-term associative memory [180].

Neurogenesis

Altman et al. [73] were the first to demonstrate that new neurons were produced in the hippocampus during adulthood and posited that these late generated neurons play an important role in learning and memory. Altman's hypothesis has been validated by a number of studies that clearly established that neurogenesis is ongoing during adulthood in a number of different species, including in (post-mortem) humans [181, 182]. New neurons are not generated throughout the hippocampus, rather they are specifically generated in the subgranular layer of the dentate gyrus. During maturation, new neurons migrate short distances to the granule cell layer [183, 184] where, if rescued, they should eventually become functional, their axons reaching the dendrites of pyramidal cells in subfield CA3 [185].

After birth, new cells need to commit to a neuronal phenotype, undergo morphological and physiological maturation, followed by integration into existing neural networks. Traditionally it is believed that information is stored at excitatory synaptic connections. Therefore, young neurons only become useful once they can relay and process information. Whether new born cells survive and integrate into functional circuits in the

brain is determined by a number of regulatory factors. A number of factors have been identified affecting hippocampal neurogenesis, such as environmental enrichment [186], exercise [187], learning [188], and antidepressant drugs [189]. On the other hand, stress can reduce the rate of neurogenesis in primates [162] and rats [190], as can inflammation [191].

The contribution of neurogenesis to cognitive function is a topic at the forefront of current debates; over the last two decades hippocampal neurogenesis has become a candidate as one of the processes needed to sustain learning and memory. A number of studies have begun, through indirect evidence, to shed some light on this issue. Evidence shows that learning enhances neuronal survival for longer than the hippocampus is thought to be necessary for the maintenance of these memories (up to two years in the human hippocampus [181, 192], and there is continuous neuronal addition, but it is discrete and strongly regulated [183, 192, 193] by a number of factors including hormones, environmental enrichment, physical exercise, and learning, amongst others.

Furthermore, the number of newborn cells correlate with effective hippocampal dependent memory [194-196]. This idea is supported by the view that immature neurons have increased excitability and thus are more responsive to exploration of a new environment [359]. A growing body of evidence has indicated that hippocampal dependent learning, including spatial learning, enhances survival of newborn neurons [197-199]. In contrast, a number of learning tasks have been reported as having no effect on the number of new neurons in the hippocampus [200]. This contradictory evidence can in part be attributed to differences in experimental conditions e.g. differences in cell labelling paradigms [201] and number of training trials [201, 202]. Most of the studies that have found no relationship between learning and neurogenesis have involved learning experiences that are rapidly acquired in a single exposure, as

opposed to training on difficult tasks. Notably, more recent studies have demonstrated that it is only certain types of learning tasks have an effect on neurogenesis - trace conditioning and spatial maze learning tasks (both hippocampal dependent) [202]. In addition, the learning experience must involve cognitive effort, and learning has to take place [202]. Moreover, the animals that learn best have more new neurons after training than those that do not learn, or do not learn efficiently [203-205]. Likewise, mice that regularly exercised, leading to an increase in neurogenesis, were better at pattern separation than controls, suggesting a role for new neurons in encoding information [203]. Furthermore, the same study also noted a decrease in neurogenesis in aged animals despite exercising. Different recruitment levels of new neurons have also been observed between adult and young chickadees, with the young birds recruiting more new neurons. The authors suggested that this may be a result of greater novelty in the environment for young versus older adults, and recruiting more new neurons allows young animals to learn more [204].

A computational model of hippocampal function that has incorporated neurogenesis suggests that the constant turnover of neurons in the dentate gyrus may ensure that each new event is encoded uniquely, avoiding interference with old or newly stored memories [205]. Moreover this model suggests a role for neurogenesis at different time scales. At short time scales, it may have an integrative role in mediating contextual associative learning, whilst at longer time scales it may be important for representing distinct contexts, thereby reducing interference of overlapping memories [206]. This is in agreement with other theories of hippocampal function that propose a sparse representation in the hippocampus; this means that the number of commonly active neurons between similar memories is small, which indicates there is less interference, and in turn improves the reliability of storage and retrieval. At the same time this would permit the network to store large numbers of distinct representations [207, 208].

If neurogenesis plays a key role in learning and memory, then disruption of neurogenesis should adversely affect these processes. Likewise, if new neurons are integrated into networks that store memories, then elimination of these neurons should result in the erasure of memory. By blocking neurogenesis, a number of studies have demonstrated that retention of spatial memory over delays is impaired [209], as is hippocampal dependent learning [194, 210-212]. Furthermore, a recent study has shown that neurogenesis depletion may only affect stimuli that are spatially close but not when they are widely separated, suggesting neurogenesis may play a role in pattern separation [213]. Moreover, reduced neurogenesis seems to have a beneficial effect when it is advantageous to forget previous events in similar situations and respond to the current context [214].

Clearly it is very challenging to study neurogenesis in vivo in the human brain. Previous studies have found evidence for hippocampal neurogenesis by using a labelling technique pre-mortem in terminally ill patients and examining their brains post-mortem [181] (Figure 11); and by assaying n-acetyl aspartate (NAA) with MRI spectroscopy (NAA is an intracellular marker of neuronal function, and is employed as an indirect measure of neuronal integrity). An increase in the peak of NAA suggests an increase in functional neurons [144, 145]. Other evidence is even more indirect. Based on the assumption that stress and depression reduce neurogenesis, Becker et al. [206] compared depressed and non-depressed individuals on a range of memory tests, and in accord with animal findings, depressed individuals had a selective deficit on a delayed-matching-to-sample task. Becker et al. concluded that there may be an underlying failure to avoid memory interference from similar items when a reduced number of new neurons are available.

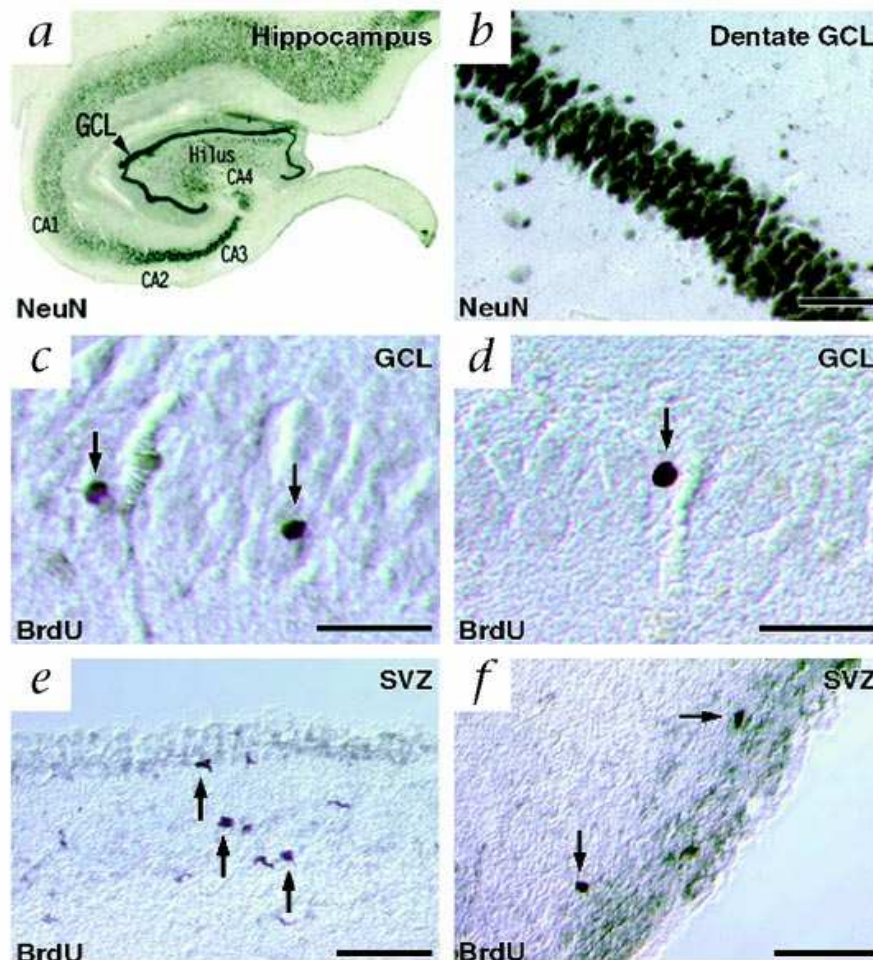


Figure 11. Newly generated cells can be detected in the adult human hippocampus in patients previously treated with BrdU.

a) The hippocampal region of the adult human brain immunoperoxidase-stained for the neuronal marker NeuN. b) The hippocampal dentate gyrus granule cell layer (GCL) visualized with immunoperoxidase staining for NeuN. c) BrdU-labeled nuclei (arrows) in the human dentate granule cell layer (GCL). d), BrdU-labeled nucleus (arrow) in the human dentate GCL. BrdU-positive nuclei have a rounded appearance and resemble the chromatin structure of mature granule cells and are found within the granule cell layer. e) BrdU-positive cells (arrows) in the subventricular zone of the human caudate nucleus. f) BrdU-positive cells (arrows) in the subventricular zone of the human caudate nucleus. All scale bars represent 50µm. (From Eriksson et al.[181]).

To summarize, it would seem that new neurons are born in the DG of the hippocampus, and make a relevant functional contribution to hippocampal processing [215, 216]. This probably occurs at a slow rate, but as the mossy fibres are scarce (of both old and new granule cells) the addition of even a small number of neurons can make a relatively large difference. It is likely that the function of new neurons is associated with peaks of information load over sustained periods that must be responded to by behavioural adjustments, which favour the recruitment of new neurons, thus it is part of a long-term adaptation process rather than an acute benefit

[217]. Neurogenesis seems to contribute to an increase in the whole granule cell number suggesting it modulates and refines existing neuronal circuits in the DG [209]; in rats a linear increase of the granular layer volume of the rat dentate gyrus throughout adulthood has been observed [218]. This in turn would increase the storage capacity within the hippocampus circuit. A number of studies have shown that after depletion of neurogenesis learning was not affected, however spatial memory was impaired, supporting the notion that neurogenesis may play a role in long-term spatial information storage [200]. Should there be hippocampal volume changes in response to acquiring and then using the spatial knowledge of London, then neurogenesis in taxi drivers' hippocampi may underpin this plasticity.

1.8 Functional differentiation along the long-axis of the hippocampus

The finding by Maguire et al. [147, 151] of increased grey matter volume in the posterior hippocampus and decreased volume in the anterior hippocampus of taxi drivers relative to control participants suggests differentiation along the long-axis of the hippocampus. During development hippocampal subregions have distinct developmental trajectories; over time the posterior region increases and the anterior decreases [219]. It is known that the lateral portion of the entorhinal cortex which projects to the dorsal (posterior in primates) hippocampus, receives the majority of its input from other neocortical areas, and it follows therefore that the posterior part of the hippocampus may be more involved in processing sensory information. The medial portion of the entorhinal cortex has higher levels of connectivity with areas such as the amygdalae, and therefore it is plausible that the ventral (anterior in primates) area of the hippocampus is more involved in processing interoceptive or emotional information [220]. In addition to the pattern of connectivity that points to a distinct roles for subregions of the hippocampus, differential genetic expressions also accompany this

neuroanatomical differentiation [221, 222]. Two broad subdivisions have been proposed where the dorsal is preferentially involved in spatial learning and memory, and the ventral is more involved in emotional processing [223]. Here I will briefly review empirical evidence for the functional heterogeneity of the hippocampus.

The majority of in vivo research on the rat hippocampus has focused on the dorsal part of the hippocampus. Moser et al. [223] trained rats with bilateral lesions either to the dorsal part of the hippocampus, or to the ventral part on a Morris water maze task. They found that damage to the dorsal third of the rat hippocampus was sufficient to cause severe impairments on the Morris water maze, but not lesions to the ventral two thirds. The degree of the impairment on this task has also been shown to correlate with the size of the lesion [223]. Similar findings have been found in rats using a T-maze task [224].

Differences in the spatial selectivity of neurons in the CA1 subfield of the dorsal and ventral hippocampus have been observed in rats during a random search task (food reward) in a chamber that contained simple visual landmarks. The dorsal hippocampus contained a higher number of place cells which at the same time showed higher resolution than cells in the ventral area of the hippocampus [69]. A more recent study employed a large scale environment (18 metre) and found that neurons in the rat dorsal hippocampus exhibited firing fields over lengths < 1 metre that increased linearly to firing fields that spanned up to 10 metres in the ventral part [70]. This evidence supports the idea that there is heterogeneity of function within the rat hippocampus and the dorsal part is necessary to sustain detailed high-resolution spatial learning and memory, and the ventral hippocampus could be more concerned with the general location i.e. which environment one is in. Furthermore, distinct firing patterns between ventral and dorsal CA3 neurons have shown - while dorsal CA3 neurons had firing fields that evenly covered the environment, ventral CA3 firing fields differed on open

and closed mazes and showed similar firing patterns with respect to goals on several arms of a radial maze, and during opposite journeys in a zigzag maze [225]. This suggests that not only do ventral CA3 neurons show an expansion of place fields, but include non-spatial firing patterns related to reward and emotional features. In this same study [225], theta rhythmic cells showed slower rates of precession in the ventral hippocampus than in the dorsal hippocampus suggesting deterioration of the spatial representation, as theta oscillations are considered a mandatory element of place cell activity. It may be that as the spatial representation decreases from dorsal to ventral, a parallel reduction of theta rhythm and increased representation of non-spatial information occurs. Although the power of the oscillations along the long-axis was reduced it remained coherent indicating that theta may serve as a mechanism to temporally bind spatial and non-spatial information [225].

Further evidence for this dissociation of function within the mammalian hippocampus comes from the study of non-human primates. Colombo et al. [226] trained monkeys on spatial and non-spatial delayed matching to sample tasks. Extracellular activity was recorded from 34 delay neurons (these are neurons that fire more during the delay period of a task than at baseline) in the ventral part of the hippocampus, 40 in the mid region and 20 in the dorsal region. The recording showed that spatial delay activity was predominantly located in the dorsal hippocampus. Non-spatial delay activity was spread evenly throughout the entire hippocampus. Moreover they found that the magnitude of delay activity increased as the delay progressed in the dorsal rather than in the ventral hippocampus.

Genetic studies have recently supplemented electrophysiological studies in rodents, offering findings that support a differentiation of gene expression along the long-axis of the hippocampus. Over 4000 genes are expressed in the mouse hippocampus. Gene expression correlations are observed between parts of the cortex that are innervated

by the dorsal and ventral hippocampus respectively [227]. Gene expression in the dentate gyrus [228] and subfield CA3 [221] also display similar regional specificity. The anatomically defined borders along the long-axis of the subfields also correlate with the regionalised gene expression [221, 222] (Figure 12). This evidence has been interpreted by some as evidence that distinct regions of the hippocampus are genetically connected to independently support different cognitive and emotional responses [220]. Nevertheless, how gene specificity and the connectivity of the hippocampus relate to each other and influence behaviour remains an unresolved issue for which evidence is slowly emerging with the advent of more refined techniques.

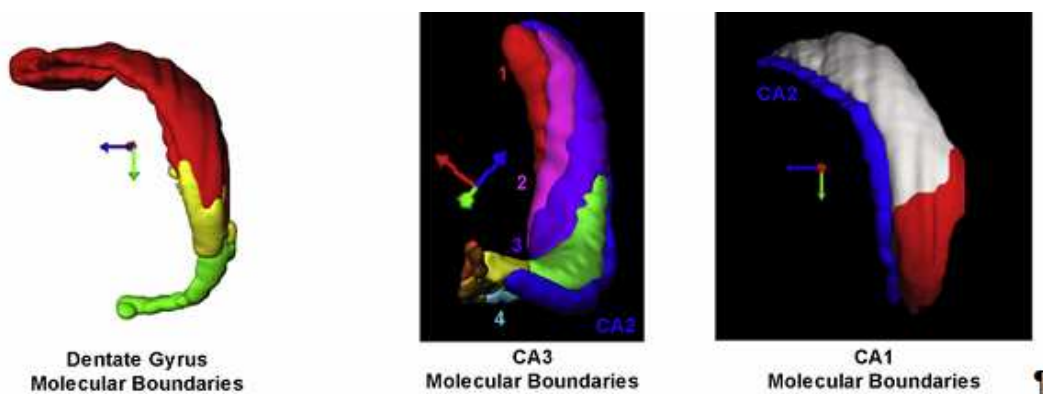


Figure 12. 3D modelling of molecularly defined subdivisions of DG, CA3, and CA1 delineated by gene expression.

Temporal CA1 is delimited by the gene *Dio3* (red), while the DG is divided into three regions based on the expression of the genes *Cyp7b1* (septal, red), *Trhr* (temporal, green), and their overlap (middle, yellow). (from Thompson et al. [171]).

Functional neuroimaging studies in humans have shown activations distributed along the rostrocaudal axis of the hippocampus. It has been suggested that remote autobiographical memories appear to be represented throughout the hippocampus, while recent memories are associated with activations clustered in the anterior hippocampus [115]. Other distinctions have been made where encoding was

associated with activation in the anterior hippocampus and retrieval with the posterior [229-232]. Other neuroimaging studies have reported regional differences in hippocampal functioning associated with the type of information processed, such as faces and names [233-235], and others associating the anterior hippocampus with novelty detection [233-235]. Overall, in terms of humans, the anterior/posterior distinction seems to be evident, but as yet there is no clear explanatory account.

Not only has a distinction between anterior and posterior hippocampus been proposed, but in humans a distinction between the functions of the left and right hippocampi has also been put forward. Notions of lateralisation within the medial temporal lobes, including the hippocampus, became prominent with the work of Milner and colleagues involving unilateral temporal lobectomy patients [239-241]. Numerous studies have over the years documented decreased verbal memory following left [236-238] and visuo-spatial memory [239-241] following right temporal lobe excisions. To some extent the cognitive map theory of hippocampal function [97] echoed this view, with the proposal that in humans the right hippocampus may support spatial representations of environments and the left language and temporal information. Functional neuroimaging gives some support for hippocampal lateralisation, with a predominance of right hippocampal activation in some spatial memory and navigation tasks [242-244], and left for verbal memory tasks [245]. However, the picture is not clear-cut. In numerous spatial memory and navigation studies the left hippocampus is often also active (e.g. [72, 244]). Moreover, left hippocampal activations are more common during the rich recollection of autobiographical memories [115], which would seem to be very 'visual' in nature. Even in previous structural MRI studies of London taxi drivers [147, 151] the effects, whilst often more evident in the right hippocampus, are also present in the left. In recent years study of laterality and memory has faded from prominence, and so the question of functional lateralisation in brain areas such as the hippocampus remains unresolved.

1.9 Thesis overview

In this chapter I outlined briefly the main themes that are germane to the research questions I seek to answer in this thesis. Using London taxi drivers as a model system, I performed six experiments aimed at developing a body of knowledge in relation to the effect of experience (operationalised as navigational expertise/skill) on the structure and function of the adult human hippocampus.

In Experiment 1 and 2, I characterised the neuropsychological profile of licensed London taxi drivers in detail, which included devising a number of new table-top associational memory tests. This was a vital starting point in order to appreciate possible cognitive effects (both positive and negative) of deploying spatial expertise.

In Experiment 3, I investigated the effects of navigational expertise in a more naturalistic setting, by examining how well taxi drivers could learn the layout of an unfamiliar town compared with a group of non-taxi drivers, and how effectively taxi drivers could integrate a new district into their existing spatial representation of London.

In the first three experiments, I focused on taxi drivers, whose expertise is clearly within the spatial domain. An important question is whether non-spatial expertise is associated with similar hippocampal grey matter volume patterns. In Experiment 4, I investigated the neuropsychological profile and assessed the hippocampal grey matter of medical doctors, whose knowledge was much less spatial than taxi drivers.

Experiment 1-4 were cross-sectional, and whilst suggesting hippocampal volume differences were associated with the development and use of a complex spatial representation, the evidence was still somewhat indirect. To definitively address the issue of hippocampal structural plasticity in the face of demands on spatial memory, a longitudinal study is required. In Experiment 5 I compared trainee London taxi drivers

before and after they completed their Knowledge training in order to assess, within subjects, the effect of their training on hippocampal grey matter volume and memory.

In addition, in order to fully characterise plasticity, in a preliminary study I also examined the effects of ceasing to use the Knowledge. In Experiment 6, retired London taxi drivers were examined for evidence of 'reverse' plasticity.

This thesis contains work reported in the following articles:

Woollett K, Maguire EA. Acquiring navigational expertise drives grey matter volume changes in the human hippocampus (in-preparation).

Woollett K, Maguire EA. Profiling anterograde associative memory in the context of navigational expertise (in-preparation).

Woollett K, Maguire EA. 2010. The effect of navigational expertise on wayfinding in new environments. *Journal of Environmental Psychology* (in press – available in early view).

Woollett K, Spiers HJ, Maguire EA. 2010. Talent in the taxi: a model system for exploring expertise. In *Autism and Talent*, Happe F and Frith U (Eds). Oxford University Press, pp 89-103.

Woollett K, Maguire EA. 2009. Navigational expertise may compromise anterograde associative memory. *Neuropsychologia* 47: 1088-95.

Woollett K, Spiers HJ, Maguire EA. 2009. Talent in the taxi: a model system for exploring expertise. *Philosophical Transactions of the Royal Society B Biological Sciences* 364: 1407-1416.

Woollett K, Glensman J, Maguire EA. 2008. Non-spatial expertise and hippocampal grey matter volume in humans. *Hippocampus* 18: 981-984.

Chapter 2

Materials and Methods

2.1 Introduction

Two methodological approaches were employed in this thesis. All six experiments involved neuropsychological assessments. In addition, Experiments 1, 4, 5 and 6 combined this neuropsychological testing with the acquisition and analysis of structural MRI brain scans. In this chapter I will give a general overview of the participants, the principles of MRI, the scanning parameters employed, the neuropsychological tests, and the data analysis techniques. Specific methodological details pertaining to individual studies, including the experimental neuropsychological tests developed during the course of this PhD, will be presented in the relevant subsequent chapters.

2.2 Participants

All participants gave informed written consent to participate in the studies in accordance with the local research ethics committee. All were healthy, with no history of neurological or psychiatric illness, and were not taking any medication.

In Experiments 1, 2 and 3, the experimental groups were licensed London taxi drivers. All had undergone the rigorous training, examinations and achieved the full green badge status described in Chapter 1. All were over the age of 21 (as required by law – see Chapter 1) and were under the age of 50 to eliminate possible confounds associated with structural brain changes during normal aging. The licensed London taxi driver profession is dominated by males. Therefore, for ease of recruitment, only males participated in all of the studies reported in this thesis. There are female taxi drivers, and it would be interesting in the future to examine whether the same effects noted here generalise to females also.

Taxi drivers were recruited in several ways. A number of the large licensed London taxi driver companies were contacted and agreed to put out calls over their in-taxi computers giving the basic details of an experiment. Interested taxi drivers then telephoned and were provided with further information and interviewed to ascertain their suitability to take part. In addition, flyers and posters were placed around London in the 'huts' (favoured cafés) that taxi drivers were known to use. Features on the laboratory's research in taxi driver publications and newspapers also encouraged taxi drivers to volunteer, as did word of mouth among the taxi drivers. In order to achieve the best working relationship with the taxi drivers, testing generally took place in two hour sessions. This was the maximum amount of time the taxi drivers could park in the taxi rank adjacent to Queen Square. The same set of taxi drivers (and control participants) took part in Experiments 2 and 3, otherwise, a new cohort of participants was recruited for each study.

In Experiment 4, the experimental group comprised medical doctors – full details are given in Chapter 5.

In Experiment 5, the longitudinal study, the experimental participants were trainee London taxi drivers. They were recruited through collaboration with the Public Carriage Office (PCO, the official UK taxi licensing body – see Chapter 1) and a number of "Knowledge schools", the taxi driver training schools in London. The PCO permitted me to make a brief presentation to their weekly intake of new trainees, in which the aims of the research were outlined. Flyers were distributed and trainees were able to declare their interest on the spot or were able to make contact at a later date. The Knowledge schools displayed posters and distributed flyers to their students and interested trainees telephoned and were interviewed to ascertain their suitability to take part. All trainees who participated were verified as officially registered 'Knowledge' candidates with the PCO. None had ever worked as mini-cab drivers.

Once recruited, a trainee was immediately invited for a structural MRI brain scan and neuropsychological testing (details in Chapter 6). Participants were issued with a wallet-sized laminated card with relevant laboratory contact details to enable them to keep in touch in case of changes in their circumstances relevant to their participation in the study. After this initial visit, contact was maintained through yearly telephone calls. These call-backs permitted me to monitor their contact details, the stage of their training, the number of appearances, any issues that had arisen that were relevant to their training, any illnesses that had occurred, and if they had decided to terminate training. All participants in the longitudinal study were also sent Christmas cards every year as another means to maintain contact and to remind them to keep in touch in case of any major change in their circumstances. Having monitored their progress, additional contact was made nearing the time they were hoping to qualify to remind them to advise me if they were successful. In fact, most contacted me voluntarily with the good news of their qualification, whereupon they were immediately invited back for another structural MRI brain scan and neuropsychological tests.

Experiment 6 involved retired licensed London taxi drivers and matched taxi drivers who were still working. The latter were recruited in the same way as the younger taxi drivers in Experiments 1, 2 and 3. Recruitment of the retired taxi drivers is described in Chapter 6. In this experiment, participants were older (typically over 60).

In every study a group of control participants was tested. The controls were also healthy males, none had ever been a licensed London taxi driver, or had been involved in taxi-driver training, nor had they worked as mini-cab drivers. Controls were age and IQ matched to the relevant experimental group. In all studies except Experiment 4 (see Chapter 5 for the reasons why), the controls also matched the experimental group in terms of years of education. In the longitudinal study of trainee taxi drivers, contact was maintained with the control participants as it was with the taxi trainees in order to

ensure their continued participation, and to monitor any changes in their circumstances (e.g. in case they decided to become taxi or mini-cab drivers – none did).

Controls were recruited from the Centre's volunteer database, from public service departments in local government offices, local banks, public libraries, and by means of online forums such as Gumtree. For Experiment 4, where control participants with high IQ's but no tertiary education were required, the assistance of the British branch of 'Mensa', a society for people with high IQs, was solicited. For Experiment 6, where elderly control participants were required, posters were displayed and flyers distributed in active retirement groups, adult education centres, bowls and golf clubs.

2.3 Biophysics of magnetic resonance imaging

In order to perform MRI, one needs a strong static magnetic field. The types of magnets used for MRI are generally one of three types: permanent, resistive and superconductive. In all experiments, a 1.5 Tesla (Sonata whole body scanner, Siemens Medical Systems, Erlangen, Germany) superconductive magnet was used, with a field strength approximately 30,000 times that of the earth. Superconductive magnets use liquid helium to keep the coils of wire at approximately 4 degrees Kelvin where there is no resistance.

2.3.1 Magnetic fields

Structural MRI has excellent spatial resolution and unlike other imaging techniques does not involve radiation. It is used to characterise anatomical differences in vivo and it can be applied to any body tissue, including the brain. MRI relies on the principle of nuclear magnetic resonance [246, 247]. All atomic nuclei possess the quantum property of spin. MRI techniques measure the effects of changing the spin of particular

atomic nuclei, such as ^1H and ^{13}C , which have an odd number of nucleons (protons plus neutrons).

In living tissue such as the brain, the most abundant source of protons is the hydrogen atom in the form of water. The hydrogen nucleus is positively charged and the spinning motion of this charge induces a local magnetic field. These hydrogen nuclei (or protons) therefore behave like small magnets, i.e. they have a magnetic moment. In the absence of a magnetic field, these individual spins are randomly orientated and the bulk material has no net magnetisation. However, in an externally applied magnetic field, B_0 , the individual spins align with the external magnetic field. If the spin is not completely aligned with the direction of the magnetic field B_0 , this causes the proton axis to revolve, or precess, around the field direction (Figure 13). The frequency with which the axis rotates around the field direction is called the resonance or 'Larmor' frequency, and is directly proportional to the field strength B_0 .

Quantum mechanics dictates that a spin can have different energies depending on the orientation of its magnetic moment with respect to the applied magnetic field: when the magnetic moment is aligned with the field, its energy will be lower than when it opposes the field. For the simple spin system of ^1H , the magnetic moment can have two orientations with respect to the magnetic field, either against it (high energy state) or along it (low, ground energy state). The amount of energy required to flip orientations is so small that the normal thermal energy available at room temperature is enough to flip spins. All of the signals generated are, therefore, based on small differences between these energy states. The fact that the energy differences are small is one reason why MRI techniques tend to be safe but also why they are typically limited by signal strength.

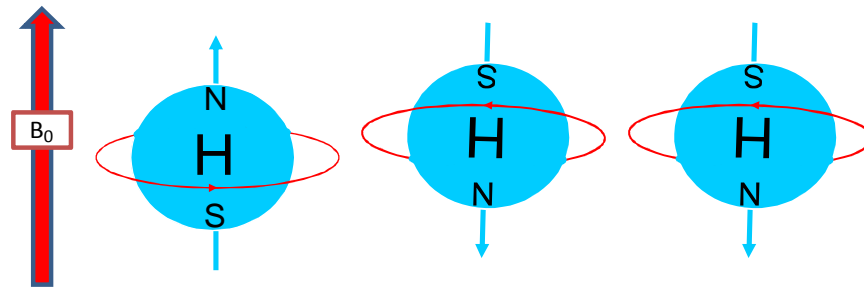


Figure 13. Protons align parallel or anti-parallel to the magnetic field (B_0).

The sum over all the nuclei in an object volume gives the net magnetisation for the body of tissue. Magnetic fields are described in a Cartesian co-ordinate system, with the z-axis being in the direction of the applied magnetic field. The resting net magnetisation is called the equilibrium magnetisation. In equilibrium, more spins are in the low than in the high energy state. Summing the contributions of individual magnetic moments will, therefore, give a net magnetic moment along the direction of the applied magnetic field. Even though a small part of the rotating magnetisation of each nucleus has a component projecting into the xy-plane, there is no equilibrium net 'transverse magnetisation' because the average magnetisation in this xy-plane over all nuclei cancels to zero.

2.4 Generating an MR signal

Spins can be excited from the low to the high energy state by the use of an oscillating radiofrequency (RF) pulse. This RF pulse is generated by applying an oscillating radiofrequency electromagnetic field (B_1) perpendicular to the main magnetic field (B_0). To achieve the most efficient transfer of energy, the oscillation frequency of the B_1 field should be the same as the spin Larmor frequency (i.e. 64 MHz for a proton in a 1.5 Tesla field, the field strength of the scanner used in my studies). The application of the RF pulse at the Larmor frequency therefore causes some of the protons in the low-energy state to absorb energy and transition into the high-energy state. This tips the

net magnetisation vector (M) of the tissue towards the transverse xy -plane (Figure 14). When sufficient energy is provided to produce a 90° change in the net magnetisation, the 'flip angle' of the RF pulse is said to be set at 90° , as was the case in all experiments in this thesis. Oriented perpendicular to B_0 is a receiver coil which measures, through current induction, the transverse magnetisation vector.

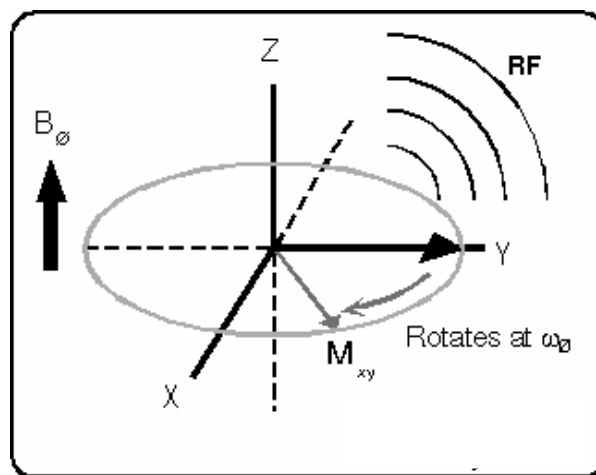


Figure 14. Generation of MR signal.

Given that the RF pulse is only transient, the spin orientations will tend to rebound to the direction of B_0 . This occurs through the protons losing the energy gained from the RF pulse to neighbouring tissue as heat (spin-lattice or T1 relaxation), for a given medium, the rate of spin-lattice relaxation is determined by a constant called T1. T1 varies for different elements: for example, the protons in water have a longer T1 than those in fat because the carbon bonds in fat resonate near the Larmor frequency, which facilitates the transfer of energy to the lattice. In the human brain, the different water content of grey and white matter (71% and 84%, respectively) means that T1 contrast can be used to provide contrast between these two tissues. White matter has a shorter T1 relaxation time than grey matter and hence appears brighter on T1-weighted images.

2.5 Tomographic image formation

Simply placing sample tissue (e.g. the brain) within a homogenous B_0 field will not produce tomographic MR images because all protons will experience (roughly) the same magnetic field and, hence, the frequencies of their emitted signal will all be identical. Additional manipulations are therefore needed to encode spatial information. In brief, this involves the application of a second magnetic field (gradient field) that allows slice selective excitation. During imaging acquisition, spatial locations are encoded by both the frequency and phase of the detected signal. Frequency encoding is determined by applying the frequency encoding gradient in the x-axis (conventionally) during acquisition. This is followed by phase encoding along a second orthogonal axis (y-axis).

The characteristics of the applied gradients determine the resolution of MRI images acquired. Step-wise increases in gradients partition the sample into small cubes, termed voxels (volume elements). Protons within a given voxel behave as a single entity given that they are subjected to the same frequency and phase encoding. The voxel size, therefore, determines the resolution of the MRI image and was set at 1x1x1 cubic mm for the experiments described in this thesis. Note that larger voxel sizes increase signal-to-noise ratios (SNR), and therefore sensitivity, but have the disadvantage of lumping protons with very different spin behaviour together which can produce a misleading signal.

2.6 Scanning parameters

Various scanner parameters can be chosen in order to optimise particular aspects of an experimental design including, for example, the signal-to-noise ratio, brain coverage, distortion, or voxel resolution.

Some of the main parameters are listed below: Repetition Time (TR): is the interval period between two consecutive 90° RF pulses and manipulates the effect of T1 relaxation.

Echo Time (TE): this is the time between the initial radiofrequency excitation and the readout of the MR signal which occurs following additional gradients (or RF pulses) that are applied to produce echoes and maximise the MR signal. The purpose of these echoes is to resynchronise the protons which have started losing coherence following the initial RF pulse, and therefore enhance the measured MR signal.

Field of view (FOV): is the square image area where the object of interest is located. The smaller the FOV, the higher the resolution of the image. This is because the magnetic field homogeneity decreases as the FOV increases which can affect the precessional frequencies across the volume of interest.

Flip Angle (α): The flip angle determines the fraction of magnetization tipped in the transverse plane (which produces the MR signal) and the quantity of magnetization left on the longitudinal axis. If the flip angle decreases, the residual longitudinal magnetization will be higher and the recovery of magnetization for a given T1 and TR will be more complete. The result of applying low-flip angle excitation is a faster recovery of longitudinal magnetization that allows shorter TR/TE and decreases scan time, and a stronger MR signal in the case of short TR.

Image contrast: is determined by a number of factors including the relaxation, and echo times. There are three main types of contrasts: T1-weighted, T2-weighted and spin density weighted. The different greyscale colouring of each tissue type reflects the variation in signal intensity between these (Figure 15). Each tissue type is derived from the sum over all the nuclei in an object volume which gives the net magnetisation for

the mass of tissue. T1 weighted images are generally used for anatomical images as they have higher spatial resolution and suffer less from artefacts.

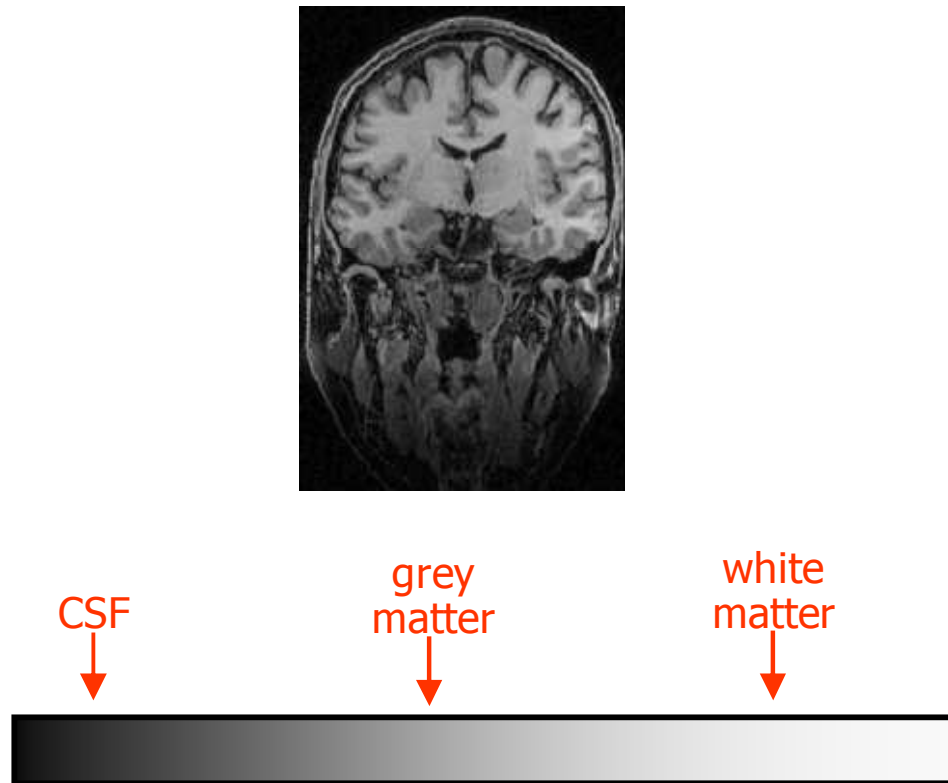


Figure 15. T1 weighted scan - Structural MRI image showing different tissue intensities.

2.7 Structural MRI sequences

Two Modified Driven Equilibrium Fourier Transform (MDEFT) sequences were employed in the experiments included in this thesis. MDEFT sequences permit the MRI signal to be acquired from the whole volume simultaneously [248]. The sequence entails preparation-acquisition cycles. During each cycle, the preparation aspect, the longitudinal magnetisation - is saturated by the 90° RF pulse, this pulse relaxes partially (during τ_1), and is then inverted by a 180° pulse, and relaxes again during the τ_2 , resulting in a T1 weighted longitudinal magnetization. It is represented by the following equation:

$$\tau_1 = \text{quot.} \times T_1 \quad \tau_2 = T_1 - \tau_1$$

Following τ_2 , gradient echoes are acquired with specific optimised parameters. Scanning of both sequences was performed using a whole-body coil for RF transmission and an 8-element phased-array head coil for signal reception. Array coils offer SNR gains over standard imaging coils. Placing an array of coils on the surface of a volume of interest allows the acquisition of images with high SNR near the surface, without sacrificing the SNR at the centre [249]. This increases the quality of the images which is particularly important when dealing with deep brain structures. The scanning parameters of the two MDEFT sequences were as follows:

For Experiments 4 and 6, parameters were optimised as described in the literature [249]: for each volunteer, 176 sagittal partitions were acquired with an image matrix of 256 x 224 (Read x Phase). Two-fold oversampling was performed in the read direction (head/foot direction) to prevent aliasing. The isotropic spatial resolution was 1 mm. Relevant imaging parameters were $TR/TE/TI = 14.59 \text{ ms}/4.15 \text{ ms}/650 \text{ ms}$, $BW = 85 \text{ Hz/Px}$, $\alpha = 20^\circ$. To increase the signal-to-noise ratio, an asymmetric position of the inversion pulse within the magnetisation preparation experiment (duration TI) was chosen, and the delay between the initial saturation and the inversion amounted to 40% of TI [249]. The total duration of the scan was 12 minutes. Special RF excitation pulses were used to compensate for B1 inhomogeneities of the transmit coil [250]. Images were reconstructed by performing a standard 3D Fourier transform, followed by modulus calculation. No data filtering was applied either in k space or in the image domain.

For Experiments 1 and 5, parameters were optimised as described in the literature [249-251]: for each volunteer, 176 sagittal partitions were acquired with an image matrix of 256 x 240 (Read x Phase). Two-fold oversampling was performed in the read direction (head/foot direction) to prevent aliasing. The isotropic spatial resolution was

1 mm. Relevant imaging parameters were TR/TE/TI = 14.59 ms / 3.4 ms / 650 ms, BW = 96 Hz/Px, $\alpha = 20^\circ$. To increase the signal-to-noise ratio, an asymmetric position of the inversion pulse within the magnetisation preparation (duration TI) was chosen, and the delay between the initial saturation and the inversion amounted to 40% of TI [249]. A fat saturation pulse was used to achieve fat suppression (see [251] for details). In addition, special RF excitation pulses were used to compensate for B1 inhomogeneities of the transmit coil [250]. Images were reconstructed by performing a standard 3D Fourier transform, followed by modulus calculation. No data filtering was applied either in k space or in the image domain. The total duration of the scan was 12 minutes.

The two scanning sequences were very similar, the main difference being the latter sequence was optimised to include fat saturation and fat selective excitation pulses which suppressed the high signal from the scalp, reducing motion sensitivity [251]. The latter sequence is an improvement of the MDEFT sequence, as such it was employed in the work included in this thesis as soon as it was developed.

Given that Experiment 5 was longitudinal, and involved scanning the same participants with a number of years elapsing between scans, it was important to monitor the scanner for stability over time. Quality assurance (QA) was carefully documented in a number of ways, including the use of a purpose-built T1 phantom with simulated cerebral spinal fluid (CSF), grey matter and white matter compartments, also a high resolution structural phantom, and standardised QA stability tests. There were no significant changes in the geometry of the images, in the signal-to-noise ratio, signal stability or signal drift. Thus the scanner was very stable over time.

2.8 Automated analysis of structural MRI scans

2.8.1 Standard voxel-based morphometry (VBM)

Data were analysed using the package Statistical Parametric Mapping (SPM; www.fil.ion.ucl.ac.uk/spm). In Experiments 1, 4, 5, and 6 SPM5 was used, and in Experiment 5, 6 SPM8 was also employed. VBM is an automated whole-brain analysis of brain structure that is implemented in SPM. It can be used to investigate differences in the composition of brain tissue on a voxel by voxel basis after deformations have been applied to spatially normalise the images in an unbiased manner. The fundamentals of this method have been described in detail elsewhere [11]. In brief, this involves a number of fully automated pre-processing steps including extraction of brain, spatial normalisation into stereotactic space, registration, segmentation into grey matter, white matter and CSF compartments, correction for volume changes induced by spatial normalisation, and smoothing. The final step in the VBM analysis after pre-processing is the statistical analysis, performed in SPM.

Spatial normalisation: In order to make comparisons between groups of individuals that will have brains of a variety of shapes and sizes it is necessary to normalise these to the same stereotactic space. This is done by registering each image to the same template image [252], thereby correcting for global brain shape differences. It is important that this registration is as accurate as possible in order to avoid partial volume effects, in which voxels contain a combination of different tissue types.

Segmentation: Following spatial normalisation the images are segmented into the different tissue types, namely grey matter, white matter and CSF. This is achieved by a tissue classification method combining *a priori* probability maps with a mixture model cluster analysis used to identify voxel intensity apportionment of particular tissue types.

This step also includes a correction for any image intensity non-uniformity. The resulting images contain voxels that will have values between 0 and 1. Most of these values will be close to either of the extremes and therefore the data is not strictly normally distributed. To be able to perform statistical analyses the segmented images need to be smoothed thereby rendering the data more normally distributed (see below).

Modulation: To identify regional grey matter volume differences it is important that the absolute volume be preserved. This is accomplished by modulating the images before smoothing. During the non-linear spatial normalisation, volumes in certain brain areas can expand and others shrink, applying a Jacobian modulation that involves multiplying the normalised grey matter by its relative volume before and after it has been spatially normalised ensures that the total amount of grey matter is the same before and after spatial normalisation [253].

Smoothing: Once the images have been segmented and modulated, they are smoothed with an isotropic Gaussian kernel (FWHM). This ensures that each voxel contains an average of the grey and white matter surrounding it and helps to compensate for any inaccuracies in the spatial normalisation, and ensures that the data are normally distributed. The size of the smoothing kernel should be comparable to the size of the expected regional differences between the groups of brains. Smoothing of 4mm and 8mm were used in the experiments included in this thesis. Smoothing at both of these kernels has been shown to be robust for VBM group comparisons [254].

From SPM5 onwards (and so used in the studies reported here) the spatial normalisation, bias correction and registration have been combined into a unified segmentation model. This methodology offers better results than serial applications of each component given that this is a circular procedure i.e. registration requires tissue classification, and tissue classification requires initial registration. This circularity is

resolved by combining these parameters into the same model and by alternating among the aforementioned parameters to obtain an improved segmentation [12].

2.8.2 Diffeomorphic Anatomical Registration using Exponentiated Lie algebra (DARTEL)

DARTEL is an intersubject registration model that is more sophisticated than the model employed by standard VBM [255]. It has been suggested that the precision of the standard intersubject registration is not refined enough to reflect the variability of brains found in the general population, and in some cases this may affect the results. DARTEL refines this process by creating a template that is the average of all the participants in the study. DARTEL registration uses the segmented grey matter and white matter images obtained using the unified segmentation process in SPM5 and SPM8 to compute a template based on the probability maps of all the participants, and then warps the corresponding tissues maps into increasingly accurate alignment with the template. The average template generated by DARTEL is not always well aligned with MNI space (the stereotactic space traditionally used in SPM) therefore a further spatial transformation is required to report the location of any differences. In order to achieve this, an affine spatial normalisation is used to match the grey matter of the template with the grey matter tissue probability map in MNI space (SPM Manual Chapter 26). DARTEL was used in some of the analyses in Experiment 5.

2.8.3 High Dimensional Warping

Standard VBM and DARTEL-VBM are used to examine grey matter volume differences in cross sectional studies, but should not be used for longitudinal assessment because of potential artefacts that could occur over time. Thus, in the analysis of the longitudinal data in Experiment 5, high dimensional warping (HDW) was implemented SPM8. HDW safeguards against non-specific subtle differences that may arise between the first and second scans within subjects (for full details see [256]). In brief,

a rigid registration of early to late scans is performed. This is followed by high dimensional registration which results in a detailed deformation field containing mapping from each point in the late image to its corresponding point in the early image for each subject. This is achieved by matching the images while concurrently maximising the smoothness of the deformations. The smoothness is incorporated into the registration using Bayesian statistics to optimise finding the maximum a posteriori (MAP) estimate of the deformation field of the parameters [257]. Grey matter changes are estimated individually before spatial normalization thus providing greater sensitivity. The arising deformation images were segmented and normalised.

2.9 Statistical inference

The main focus in the experiments in this thesis was on grey matter. Statistical inferences were made using a series of general linear model (GLM) contrasts in SPM to detect regions of grey matter volume that differed between groups of participants. The GLM framework allows a range of tests to be applied, including group comparisons, identifying regions of grey matter that are related to specified covariates. For example an F contrast of either [-1 1] or [1-1] will give an F-test equivalent to a two-tailed t-test. This procedure results in an 'image' of statistics (i.e. a statistical parametric map - SPM). Drawing statistical inferences from this SPM relies upon reliably locating voxels where an effect is present whilst controlling against the probability of false positives. The general convention, which was followed here, is to apply a statistical threshold of $p < 0.05$ corrected for multiple comparisons across the whole brain, and a threshold of $p < 0.001$ (uncorrected) where there are *a priori* hypotheses concerning specific brain areas [13]. When appropriate (and described where implemented), small volume correction was applied for the hippocampus, $p < 0.05$ corrected for the volume.

2.10 Neuropsychological tests

In each study a range of neuropsychological tests was employed. These included standardised tests with published normative data, and experimental tests devised specifically to address the research questions in this thesis. As different tests were used in different studies, I will describe the specific tests and their functions in subsequent chapters as they arise. Overall, the standardised tests were used to measure basic IQ and handedness, attention, perception, executive functions, memory – including recall and recognition, verbal and visuo-spatial, anterograde and retrograde; in addition there were several affective measures of anxiety and stress.

All participants had excellent spoken and written English. The vast majority of participants were native English speakers and as such, an estimate of verbal IQ was obtained using the Wechsler Test of Adult Reading [258]. In order to assess visual information processing and abstract reasoning skills, and also to gain a measure of intellectual functioning in participants who were not native English speakers, the Matrix Reasoning sub-test of the Wechsler Abbreviated Scale of Intelligence [259] was used. Handedness was assessed with the Edinburgh Handedness Inventory [260].

Data were analysed using SPSS version 17.0 for Windows. Basic group comparisons relating to participant characteristics, such as age and IQ, were made using two-tailed t-tests. For the main analyses, data were screened for outliers, homogeneity of variance, and to ascertain if the data were normally distributed. Multivariate analysis of variance (MANOVA - Hotelling's trace multivariate test) was employed using the general linear model with the significance threshold set at $p < 0.05$. Where a MANOVA indicated a significant effect, the between-subjects tests were employed to ascertain the source of the significance with a threshold of $p < 0.05$. Where appropriate,

correlations and partial correlations were used to examine relationships between the variables.

Chapter 3

The effects of navigational expertise: an in-depth analysis

Précis

Previous studies have documented the pattern of hippocampal grey matter volume in licensed London taxi drivers, and preliminary evidence was reported about their performance on some basic cognitive tasks. However, a detailed neuropsychological profile of taxi drivers is lacking. This is vital in order to appreciate possible cognitive effects (both positive and negative) of deploying spatial expertise, the implications this may have for understanding the effects of plasticity in the brain, and the functions of the human hippocampus in particular. In this chapter I describe two experiments where an extensive battery of both standardised and newly-developed experimental tests was administered in order to address this issue.

3.1 Introduction

As outlined in Chapter 1, grey matter volume increases in various parts of the brain have been identified in a number of skilled groups such as musicians [42, 44, 261], mathematicians [48], bilinguals [49], jugglers [52, 55], and medical students [53], typists [46], golfers [47], dancers [50]. As well as exploring the grey matter substrates of a skill itself, a related and important question is whether a skill and its associated neuroanatomy confer broader cognitive advantages or indeed costs. This question has particular significance for the domains of rehabilitation and education. Only a limited number of studies have considered the neuropsychological consequences of expertise. Professional musicians with increased grey matter volume in Broca's area were found to show enhanced judgements of line orientation and three-dimensional mental rotation ability [43, 215]. This was attributed to their musical sight-reading and motor sequencing expertise. Professional musicians also show a greater increase in grey matter volume in motor and auditory areas. However, this

expertise can come at a cost, with some musicians suffering focal dystonia, a loss of control and degradation of skilled hand movements [41].

The consequences of skill acquisition have also been investigated in another group of experts, London taxi drivers [147, 151]. The volume of the hippocampus in some non-human species has been reported to vary as a function of the demands placed on spatial memory (see Chapter 1; refs [118, 121, 262]). Similar effects were also found in licensed London taxi drivers, with greater grey matter volume in posterior hippocampi and less grey matter volume in anterior hippocampi compared with control subjects. In addition, posterior hippocampal grey matter volume correlated positively, and anterior hippocampal grey matter volume negatively, with the number of years spent taxi driving [147, 151]. Related to this, Maguire et al. [151] confirmed that London taxi drivers performed significantly better than control subjects (London bus drivers) on tests assessing knowledge of London landmarks and their spatial relationships. Interestingly, in the same study taxi drivers were found to be significantly worse than control subjects on the delayed recall of the Rey-Osterreith complex figure, a test of anterograde visuo-spatial memory. It was suggested that this below average score could be related to their reduced anterior hippocampal grey matter volume.

Unlike many other experts such as professional musicians and bilinguals, London taxi drivers acquire their knowledge in adulthood. In addition, taxi drivers are the only group where expertise has been consistently associated with both increases and decreases in grey matter volume. When considered along with the preliminary evidence of positive and negative functional consequences of their navigation expertise [151], this makes taxi drivers a particularly interesting model for studying the effects of expertise on the brain.

Moreover, given the focal nature of the grey matter volume changes in the taxi drivers, they offer another line of evidence to complement neuropsychological and functional neuroimaging studies in helping to understand the role of hippocampus in memory and navigation.

While Maguire et al.'s [151] preliminary neuropsychological findings in taxi drivers are intriguing, the battery of tests they employed was brief and did not permit a wide-ranging examination of the cognitive and memory functions of taxi drivers. In the current study, therefore, I sought to replicate and extend the previous findings by undertaking a comprehensive neuropsychological evaluation of a new sample of taxi drivers and matched control participants. As well as assessing basic cognitive and affective functions, twenty two different memory measures were taken. These tasks assessed a broad range of memory types and processes including visual and verbal memory, recall and recognition memory, single item and associative memory, and anterograde and retrograde memory. The main aim was to ascertain the effect of spatial expertise on the memory profile of taxi drivers, examining whether any specific memory type or process was implicated. In so doing I hoped to contribute new information to debates about the functions of the human hippocampus specifically, and the benefits and costs of skill acquisition more generally.

3.2 Experiment 1

3.2.1 Materials and methods

Participants

Forty healthy male volunteers participated in the study. Of these, 20 were licensed London taxi drivers, and 20 were control subjects. The background details of the two groups are shown on Table 1. All taxi drivers had completed “The Knowledge” training, had passed the necessary Public Carriage Office examinations, and obtained a full (green badge) licence. The taxi drivers and control participants did not differ in terms of age ($t(38)=1.81$; $p=0.07$), handedness ($t(38)=0.90$; $p=0.9$), or age at which they left school ($t(38)=1.51$; $p=0.1$). As all participants were native English speakers, an estimate of verbal IQ was obtained using the Wechsler Test of Adult Reading [258]. IQ estimates for both groups were in the average range, and did not differ significantly ($t(38)=0.57$; $p=0.5$). Visual information processing and abstract reasoning skills were assessed using the Matrix Reasoning sub-test of the Wechsler Abbreviated Scale of Intelligence [259]. The mean scaled scores for both groups were comparable and did not differ significantly ($t(38)=0.36$; $p=0.7$).

Table 1 - Participant characteristics

Participant characteristics	Taxi drivers Mean (SD) (n=20)	Controls Mean (SD) (n=20)
Age (years)	43 (3.46)	40 (6.54)
Education (age left school, years)	16.30 (0.80)	16.70 (0.86)
Estimated verbal IQ (WTAR)	98 (5.10)	98.90 (4.59)
Matrix reasoning scaled score (WASI)	11.75 (1.68)	12 (2.55)
Handedness – laterality index ¹	78 (45.91)	79 (19.32)
Years experience taxi driving	12.35 (6.85)	-

WTAR=Wechsler Test of Adult Reading; WASI=Wechsler Abbreviated Scale of Intelligence;
¹Edinburgh Handedness inventory

Structural MRI brain scan

A whole brain structural MRI scan was acquired for each participant as described in Chapter 2, using the MDEFT sequence that included fat suppression.

Standardised neuropsychological tests¹

A test battery was employed to assess a range of basic cognitive, memory and affective functions. The majority of tests were widely-used standardised instruments and questionnaires with published normative data. I summarise these tests and their purpose below.

Basic cognitive functions were assessed with the following tests:

Digit span: Wechsler Adult Intelligence Scale III (WAIS III; [263]). This test assesses working memory and attention. Sets of digits are read out aloud and subjects have to repeat the numbers in the same order. The trials start with two numbers, and then rise incrementally to contain 9 numbers. The second part of the test involves the same scenario but this time subjects have to say the numbers in the reverse order that they hear them. Raw scores were converted to scaled (standardised) scores in line with the test instructions.

Spatial span: Wechsler Memory Scale III (WMS III; [264]). This test assesses visual working memory and attention. A set of cubes is laid out on a plastic board placed in front of the participant. The experimenter taps the cubes in a sequence. Subjects then have to tap out the same sequence. The trials start with sequences of two and then rise incrementally to sequences of 9. The second part of the test involves the same scenario but this time subjects have to

¹ Examples of standard neuropsychological tests cannot be included due to copyright restrictions.

tap the sequences in the reverse order to that presented. Raw scores were converted to scaled (standardised) scores in line with the test instructions.

Verbal fluency [265]: This test evaluates spontaneous production of words under restricted conditions. The participants are instructed to say as many words as possible, excluding proper nouns, that begin with a letter of the alphabet. The letters F, A, S are then presented in that order. The words produced during a 60-second period for each letter are recorded. The score is the number of words produced. This phonetic version of the fluency test is sensitive to frontal lobe dysfunction.

Block design: Wechsler Abbreviated Scale of Intelligence (WASI; [259]). This is a timed test of visuo-spatial ability and motor skills. Participants view abstract designs and are asked to reproduce these using coloured blocks. The designs increase in difficulty as the test progresses by using 4 or 9 blocks.

Brixton test [266]: The Brixton test is a test of executive function and involves the presentation of cards containing an array of numbered circles. A blue circle moves from one card to the next following an undeclared set of rules and participants must predict where the blue circle will appear next.

Visual Object Space Perception Battery (VOSP; [267]): This is a test of object and spatial perception. Three of its sub-tests were administered. 1. Object decision: this assesses the integrity of object recognition. On each trial a subject is shown 20 silhouettes of four objects, three of which are nonsense figures and one of which is a real object. The task is to indicate which is the real object in each array. 2. Cube analysis: this assesses visuo-spatial perceptual abilities. On each trial a card is presented depicting a three-dimensional

drawing. The participant has to indicate how many blocks appear in each drawing, taking account of the fact that some blocks are 'hidden' and have to be assumed to exist behind or underneath other blocks. 3. Number location: this assesses spatial perception. On each trial a card is presented containing two squares, one above the other. The top square has the numbers one to nine randomly distributed within it, and the other square has only one black dot. The participant has to identify which number corresponds exactly to the position of the black dot.

Memory was assessed with the following standardised tests (see later section for the experimental memory tests):

Doors and People test [268]: This test is designed to evaluate visual and verbal episodic memory, including measures of recall and recognition that are matched for difficulty.

Warrington Recognition Memory for Words [269]: This test is designed to assess recognition memory for words. Participants are shown single words one at a time for three-seconds each and asked to say if the word is "pleasant" or "unpleasant" to aid encoding. Immediately after presentation, the participants are given a list of word pairs and instructed to indicate which one of each pair they have seen in the previously presented set. A score of 1 is awarded for correct responses, and the total possible score is 50.

Warrington Recognition Memory for Faces [269]: This test is designed to assess recognition memory for faces. Participants are shown single faces one at a time for three-second each and asked to say if the face is "pleasant" or "unpleasant" to aid encoding. Immediately after presentation, the participants are shown

pairs of faces and instructed to indicate which one of each pair they have seen in the previously presented set. A score of 1 is awarded for correct responses, and the total possible score is 50.

Autobiographical Memory Interview (AMI) [270]: This test assesses participants recall of facts from their own past life, and recall of specific events. Both types of memory are assessed across three broad time bands: childhood, early adult life, and recent times.

Rey Auditory Verbal Learning Test (RAVLT) [265]: This is a test of verbal learning and memory. A list of 15 words is read aloud at the rate of one word per second. The participant must repeat all the words he can remember, in any order. This procedure is carried out a total of five times. Then, a second list of 15 words is presented once. Immediately following this, the participant is asked to remember as many words as possible from the first list.

Logical memory (WMS-III) [264]: This sub-test comprises a short story, and is designed to assess the free recall of verbal material. A participant listens to the story, and is then asked to recall as much of it as he can remember. In addition, a delayed recall test is also performed after a thirty-minute interval.

Verbal paired associates (WMS-III) [264]: This is a test of verbal learning and memory. A list of 8 pairs of words is read aloud at the rate of one pair per second. The experimenter then gives the participant one word from each pair and the participant must give the missing word to form the pair. This procedure is carried out a total of four times. Then, after a delay of 30 minutes it is repeated once.

Rey-Osterrieth Complex Figure [271]: This test is designed to assess perceptual organization and the free recall of visual material. Participants are instructed to copy a complex figure onto a blank sheet of paper. Thirty minutes later they are asked to recall the figure and draw what they can remember onto another blank sheet of paper. (Figure 16).

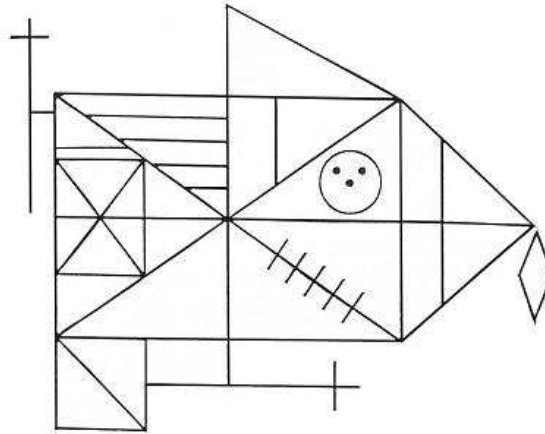


Figure 16. The Rey-Osterrieth complex figure.

Taylor Complex Figure [272]: This test is an alternative version of the Rey complex figure. The administration procedure is identical.

Anxiety and stress were assessed with the following instruments:

Perceived Stress Scale (PSS) [273]: This questionnaire, comprising 10 questions, is designed to measure the degree to which situations in life are seen as stressful within the last month. The questions are of general nature and easy to comprehend. Each question can be awarded a score from 0 to 4, 0=never up to 4=very often.

State-Trait Anxiety Inventory (STAI) [274]: This questionnaire is divided into two parts; the first assesses the subject's anxiety right at the moment of testing (state), while the second part assesses the subject's anxiety generally in daily

life (trait). Each part comprises 20 statements. The subject must describe the intensity of feelings. Each question can be awarded a score from 1 to 4, 1=not at all up to 4= very much so.

Life stress and job stress ratings: Participants were asked to rate the level of current life stress and job stress respectively, on a scale of 1-10 (low-high).

Experimental memory tests

A number of additional non-standard memory tests were also administered and are detailed below.

Public events test: Retrograde memory for public events was evaluated using 28 black and white photographs of well known public events spanning four decades (1974-2005). Participants were required to name the event, give the date of the event (+/- 3 years) and name the location where the event took place. Only events where all 3 details were correct achieved a score of 1.

London landmark recognition memory test: Recognition memory for London landmarks was assessed by showing participants 48 colour photographs of landmarks one after the other. Half of the pictures were of famous London landmarks and half were distractor landmarks that were neither famous nor in London, but were visually similar to the London landmarks (Figure 17). Target and distractor landmarks were randomly intermixed. The format was a yes/no recognition test where subjects were asked to state whether they recognised each landmark as a famous London landmark or not. The test was not formally timed, however subjects on average took 2-5 seconds per photograph.



Figure 17. Example of London landmark recognition test stimuli.

The top photograph is a famous London landmark, the Royal Albert Hall. The bottom photograph is a visually similar non-target distractor landmark.

London landmark proximity judgements Subjects' knowledge of the spatial relationships between London landmarks was tested using a proximity judgements task. Stimuli were 10 colour photographs each depicting a famous London landmark. On each trial, subjects had to judge which of two other London landmarks was closer (as the crow flies) to the target London landmark (see Figure 18 for an example stimulus). There were 10 trials. The test was not formally timed, however subjects on average took 5-8 seconds per photograph.



A



B

Figure 18. Example of London proximity judgements test stimuli.

The top photograph is the reference landmark, in this example, the Millennium Wheel/London Eye. Subjects had to indicate whether the London Aquarium (photograph A) or the Millennium Dome (photograph B) is closer to the London Eye as the crow flies.

Recognition memory tests for unfamiliar landscapes and buildings: Visual memory was further probed using two recognition memory tests of unfamiliar topographical stimuli, one of which involved landscapes, and the other buildings. The presentation and testing procedures for both the landscapes and buildings tests were identical to that of the standard Warrington Recognition Memory Test (see previous section; [269]; see also [275]). The target stimuli in each test were 50 black and white photographs of buildings, and 50 black and white photographs of landscapes. House numbers, people, animals and vehicles were all excluded from the photographs. The distractor items were chosen based on extensive pilot testing to be visually similar to the targets. As with the standard

Recognition Memory Test, during the study phase the participant was instructed to decide if each landscape (or building) was pleasant or unpleasant. This was followed by a test phase in which the target and distractor items were randomly intermixed and presented in pairs. The subject was instructed to indicate which one of the items from each pair had been presented during the study phase.

Object-place associations test: The ability to form and recall object-place associations was examined using a table top array (similar to those employed in [240, 276]). This comprised 16 coloured photographs of single objects placed on a white board measuring 64cm x 48cm (Figure 19). Subjects were given one minute in which to name each object and study their positions. Immediately after, the array was removed and a blank board of equal dimensions was introduced. Subjects were given the 16 object photographs and asked to place them in the correct locations. Locations were noted by the examiner before the next learning trial. Subsequent study periods lasted 30s and after each study period participants were immediately asked to reproduce the array as described above. No feedback was given during the test. The study-recall procedure was repeated until a criterion of 100% correct object placements was reached. Subjects were also asked to reproduce the array after a delay of 30 minutes. The number of correct and incorrect positions was determined using a transparent template grid comprising squares of equal size (8x6cm each). The exact position of correctly and incorrectly placed objects was recorded. A correct score of 1 was given to objects placed within 3cm in any direction of their original place on the board. Two measures were derived, the number of learning trials required to achieve criterion, and the number of correct object-place associations recalled following the delay.

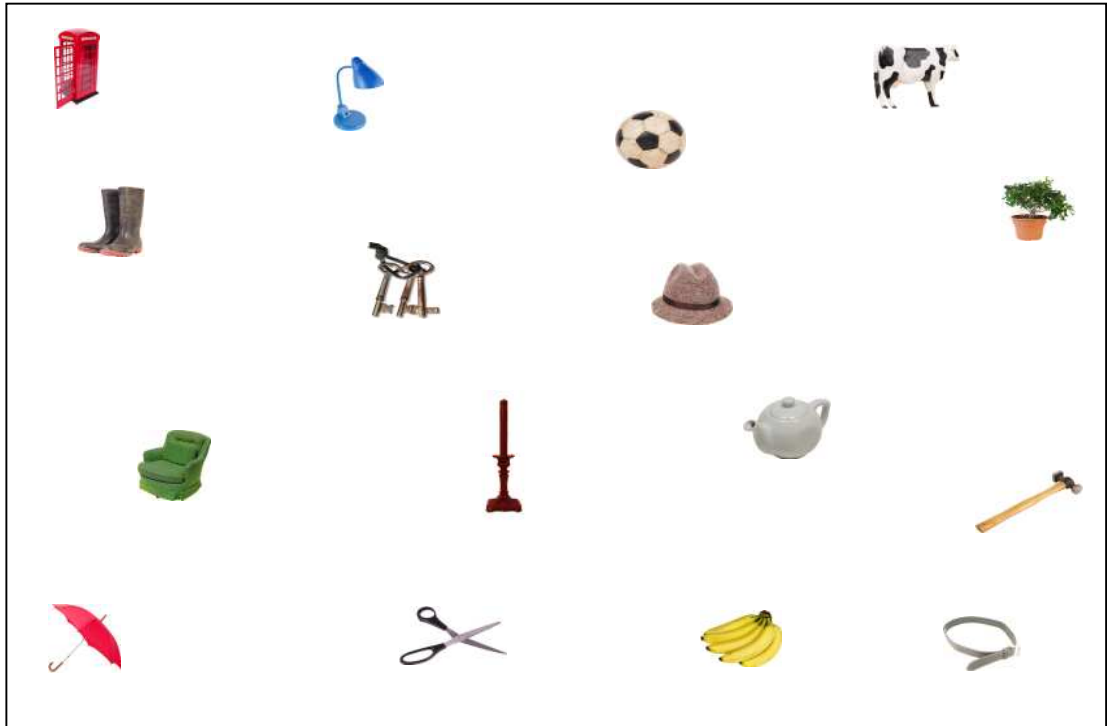


Figure 19. Object–place associations test.

The table top array used in the object–place associations test

3.2.2 Procedure

Each participant was tested individually during two sessions each lasting approximately two hours. The two sessions were at least one week apart and no more than three weeks apart. The order of neuropsychological tests within and across sessions was carefully balanced to ensure that similar tests were not administered in close proximity (for example, the Rey and Taylor complex figure tests were always administered in separate sessions). The order of neuropsychological testing and MRI scanning was random across subjects.

3.2.3 Data analysis

MRI images were analysed using voxel-based morphometry (VBM) implemented in the Statistical Parametric Mapping software (SPM5, Wellcome

Trust Centre for Neuroimaging, London, UK – see Chapter 2). Smoothing with an 8mm full width at half maximum isotropic gaussian kernel was used. The two groups (taxi drivers, control subjects) were compared using a two sample t-test to investigate differences in grey matter volume. In addition, parametric effects on grey matter volume of participant characteristics and behavioural performance were also examined. The effects of global grey matter volume and subject age were excluded by modelling them as confounding variables. Given our a priori interest in the hippocampus, the significance level was set at $p < 0.05$ corrected for the volume of the hippocampus using a sphere of 4mm. The significance level for the rest of the brain was set at $p < 0.05$ corrected for multiple comparisons across the whole brain. Our interest was in anterior and posterior hippocampal regions. These were delineated as outlined in Chapter 2. Of note, the anterior and posterior peaks in this study were clearly distinct in the anterior-posterior ('y') direction and separable by substantially more than the smoothing kernel.

Behavioural data were analysed as outlined in Chapter 2.

3.3 Results

3.3.1 MRI data

The main focus of this study was to investigate the neuropsychological profile of licensed London taxi drivers. However, we first sought to establish if the patterns of hippocampal grey matter volumes observed in previous studies of taxi drivers [147, 151] were also true of the present sample. In particular, we were interested to know if they too would show a decrease in anterior hippocampal

volume relative to control subjects, and a negative correlation between years taxi driving and grey matter volume in this region.

In the first instance the two groups were compared to assess differences in grey matter volume. Greater grey matter volume was found in the control subjects compared to the taxi drivers in the right anterior hippocampus (peak (x, y, z) 34, -14, -14; $z=3.31$; extent anteriorly in the y direction to -9mm and -22mm posteriorly; see Figure 20a). Taxi drivers had greater grey matter volume in the left posterior hippocampus (-24, -34, 6; $z=2.70$) compared to control subjects. Taking into account the smoothing kernel (8mm), these findings are comparable with those reported previously [146, 150]. No significant effects were apparent anywhere else in the brain.

We next examined the effect of navigation experience on grey matter volume by entering number of years taxi driving in London as a covariate of interest in the VBM analysis (controlling for subject age). As in the previous studies, grey matter volume in the right anterior hippocampus (34, -4, -20; $z=3.41$) was found to decrease the longer taxi drivers had been navigating in London (Figure 20b). Right posterior hippocampal grey matter volume increased with number of years taxi driving (although this just failed to reach statistical significance, 12, -34, 2; $z=2.36$; $p=0.06$). No significant effects were apparent anywhere else in the brain.

3.3.2 Neuropsychological data

Having established the pattern of hippocampal grey matter volume was comparable with previous taxi drivers findings, we next turned to the main aim of the study, namely to establish if there were neuropsychological consequences

of being a licensed London taxi driver. Mean performance scores are shown on Tables 2-4. Three separate MANOVAs were used to interrogate the data.

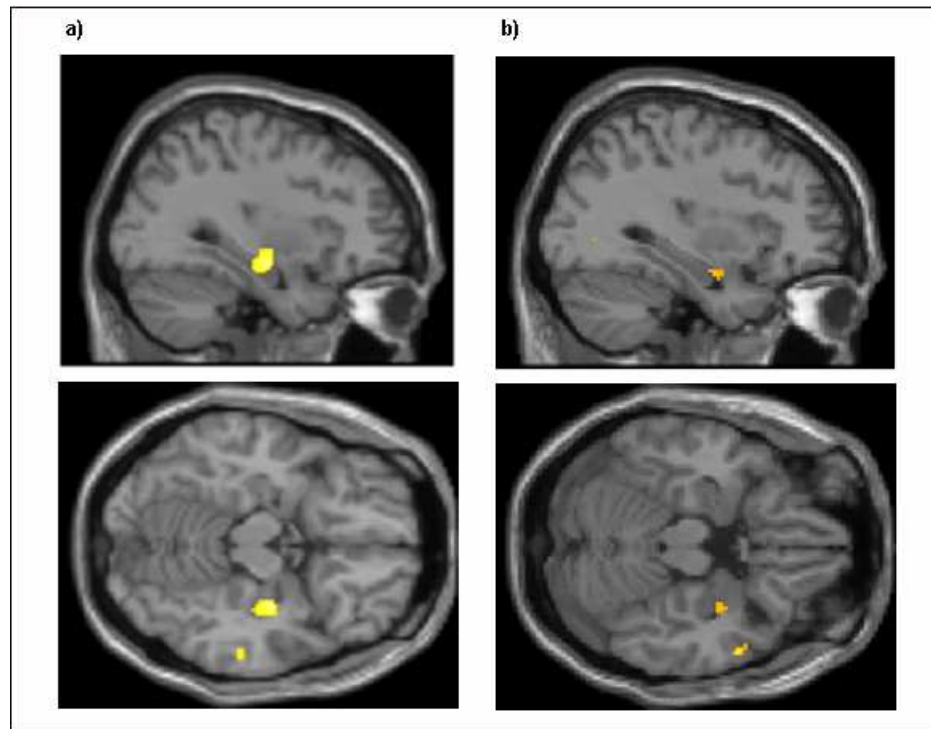


Figure 20. Anterior hippocampal grey matter volume differences between taxi drivers and control subjects.

a) Anterior hippocampal grey matter volume was decreased in licensed London taxi drivers compared with matched control subjects. Data are shown on sagittal (upper panel) and axial (lower panel) sections from the canonical SPM5 template. b) Anterior hippocampal grey matter volume was negatively correlated with navigation experience, with less grey matter volume the more years spent taxi driving. Data are shown on sagittal (upper panel) and axial (lower panel) sections from the canonical SPM5 template.

A MANOVA was first performed on the 8 basic cognitive measures listed in Table 2. There was no overall difference between the groups ($F(8,31)=1.22$; $p=0.32$). This suggests that attention, working memory, spatial span, executive functions, construction abilities and perception are not variables that distinguish between the groups, and are presumably unlikely to explain the hippocampal grey matter volume differences that were observed.

Table 2 - Basic cognitive measures

Basic cognitive measures	Taxi drivers Mean (SD) (n=20)	Controls Mean (SD) (n=20)
Digit span scaled score (WAIS-III)	12.25 (2.57)	11.45 (2.01)
Spatial span scaled score (WMS-III)	12.90 (2.67)	12 (2.40)
Verbal fluency - FAS (total score)	43.60 (10.01)	47.40 (12.43)
Block design scaled score (WASI)	8.85 (1.81)	9.45 (2.43)
Brixton Test (6-7=average-high average)	6.60 (1.14)	7.20 (1.19)
VOSP Object decision (/20)	17.85 (1.78)	16.85 (1.72)
VOSP Cube analysis (/10)	9.70 (0.51)	9.60 (0.59)
VOSP Number location (/10)	9.65 (1.13)	9.75 (0.55)

WAIS-III=Wechsler Adult Intelligence Scale; WMS-III=Wechsler Memory Scale; WASI=Wechsler Abbreviated Scale of Intelligence; VOSP=Visual Object and Space Perception Battery

A MANOVA performed on the 5 stress/anxiety measures listed in Table 3, showed taxi drivers and control subjects did not differ significantly on this set of measures either ($F(5,34)=0.89$; $p=0.49$).

Table 3 - Stress measures

Stress measures	Taxi drivers Mean (SD) (n=20)	Controls Mean (SD) (n=20)
Perceived Stress Scale	12.75 (7.30)	11.65 (7.13)
State-Trait Anxiety Inventory (State)	31.45 (9.13)	28.10 (5.49)
State-Trait Anxiety Inventory (Trait)	38.10 (10.33)	35.50 (11.98)
Life stress rating ^a	4.40 (1.75)	4.80 (2.41)
Job stress rating ^a	5 (1.91)	4.50 (2.03)

^aRatings from 1(no stress)...10(very high stress)

A MANOVA was then performed on the memory measures listed on Table 4. On this occasion there was a significant difference between the two groups ($F(22,17)=3.26$; $p=0.01$). The source of this difference was investigated using the tests of between-subjects effects produced by MANOVA. There were 8 main effects. Taxi drivers were significantly better than control participants on the London landmark recognition memory test ($F(1,38)=4.08$; $p=0.05$) and the

London landmark proximity judgements test ($F(1,38)=13.02$; $p=0.001$; see also Figure 21). They also performed better than control subjects on the semantic section of the Autobiographical Memory Interview (AMI; $F(1,38)=1.34$; $p=0.01$). By contrast, taxi drivers scored significantly worse than control subjects on the verbal-paired associates test both at immediate ($F(1,38)=7.09$; $p=0.01$) and delayed recall ($F(1,38)=4.51$; $p=0.04$), and the delayed recall of the Rey complex figure ($F(1,38)=9.08$; $p=0.005$). Taxi drivers also took significantly longer to reach criterion on the object-place association test (see also Figure 22) ($F(1,38)=18.25$; $p=0.0001$), and recalled fewer of the object-place associations after a delay ($F(1,38)=9.36$; $p=0.004$).

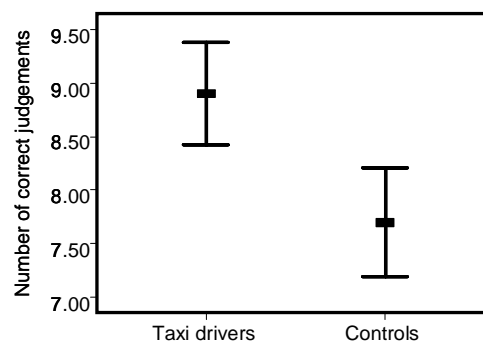


Figure 21. London proximity judgements.

Taxi drivers were significantly better than control subjects at making proximity judgements. Bars represent ± 2 standard errors.

Table 4 - Memory measures

Memory measures	Taxi drivers Mean (SD) (n=20)	Controls Mean (SD) (n=20)
Doors and People Test (overall scaled score)	9.70 (3.26)	9.95 (2.81)
WRMT – Faces scaled score	10.30 (3.11)	10.80 (3.25)
WRMT – Words scaled score	12.10 (1.83)	12.85 (1.22)
Autobiographical Memory Interview (semantic) ^T	53.10 (1.52)	49.02 (6.55)
Autobiographical Memory Interview (autobiographical)	22.25 (3.29)	23.20 (1.60)
Rey auditory verbal learning test – IR (Σ for 5 trials /75)	54.90 (6.92)	57.10 (5.32)
Rey auditory verbal learning test – DR (/15)	10.55 (2.56)	12.00 (2.10)
Logical memory – IR (WMS-III) percentile	69.25 (19.47)	70.05 (22.67)
Logical memory – DR (WMS-III) percentile	68.70 (14.91)	69.05 (20.95)
Verbal paired associates – IR (WMS-III) scaled score ^C	10.85 (2.47)	12.80 (2.14)
Verbal paired associates – DR (WMS-III) scaled score ^C	10.75 (2.38)	12.10 (1.55)
Rey-Osterreith complex figure copy (/36)	35.65 (0.98)	36 (0)
Rey-Osterreith complex figure – DR (/36) ^C	19.72 (5.07)	24.85 (5.65)
Modified Taylor complex figure copy (/36)	35.75(0.91)	35.90 (0.44)
Modified Taylor complex figure – DR (/36)	19.35 (6.02)	22.35 (5.25)
Public events test (/28)	20.05 (10.49)	17.00 (5.08)
London landmark recognition memory test (/48) ^T	41.10 (3.71)	37.60 (6.40)
London landmark proximity judgments (/10) ^T	8.90 (1.02)	7.70 (1.08)
Recognition memory test for unfamiliar landscapes (/50)	41.60 (5.28)	43.85 (4.83)
Recognition memory test for unfamiliar buildings (/50)	44.05 (3.01)	45.75 (2.80)
Object-place associations – number of trials to criterion ^C	4 (0.79)	2.95 (0.75)
Object-place associations – DR (number correct /16) ^C	9.40 (1.46)	10.85 (1.53)

WRMT=Warrington Recognition Memory Test; IR=immediate recall; DR=delayed recall after 30 minutes); WMS-III=Wechsler Memory Scale; The Autobiographical Memory Interview is the test devised by Kopelman et al. [270]; ^TTaxi drivers significantly better than control subjects; ^CControl subjects significantly better than taxi drivers

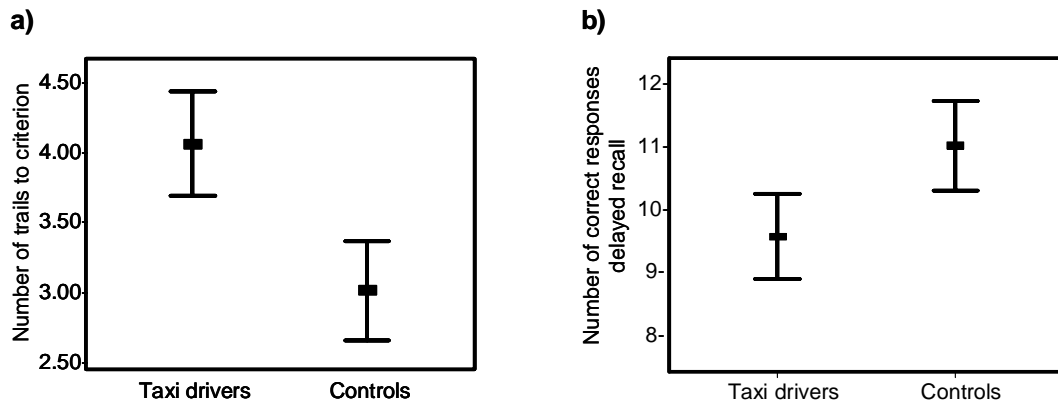


Figure 22. Object-place test.

a) Taxi drivers required more learning trials to reach criterion than control subjects. b) Taxi drivers recalled fewer of the object-place associations than control subjects following a delay. Bars represent +/- 2 standard errors.

Correlations between MRI and neuropsychological data

Taking the memory measures that distinguished between the two groups, we next examined whether performance on these tests correlated with grey matter volume. In taxi drivers, performance on the London landmark proximity judgements test correlated positively with grey matter volume in the posterior hippocampi (30, -30, 0, $z=3.59$, $r=0.65$, $p<0.002$; -28, -30, 0; $z=2.87$; $r=0.52$, $p<0.02$). In control subjects, performance on the delayed recall of the Rey complex figure correlated positively with hippocampal grey matter volume (36 - 24 -8, $z=2.90$, $r=0.60$, $p<0.005$; -24, -32, 6, $z=2.98$, $r=0.62$, $p<0.003$), as did immediate (-36 -22 -14, $z=3.57$, $r=0.7$, $p<0.0001$) and delayed recall (30 -18 -20, $z=3.11$, $r=0.57$, $p<0.009$) of the verbal paired associates test. Finally years experience taxi driving did not correlate significantly with any of the memory scores.

3.4 Discussion

In this study I found that licensed London taxi drivers, whilst being experts in their knowledge of London's layout, were deficient compared with matched control subjects at acquiring and retaining certain types of new information. Specifically, they were poorer at learning object–place and word-pair associations. After a delay they also recalled less of this associative information, and fewer elements of the Rey complex figure. By contrast their learning of and recognition memory for individual items was comparable with control subjects, as were retrograde memory for autobiographical and semantic information, executive and perceptual functions, working memory and levels of stress and anxiety.

That taxi drivers were significantly more knowledgeable about London landmarks and their spatial relations than the control subjects replicates a finding reported in a previous study [151]. The same is true for their poorer delayed recall of the Rey complex figure [151]. In the current study I extended this line of enquiry further by employing a broader range of memory tests. I found that in addition to poorer performance on delayed recall of the Rey complex figure, the taxi drivers took significantly more trials to learn the locations of 16 objects on a table top array. Despite reaching criterion during the learning phase, they also recalled fewer of these object–place associations after a delay. The combination of these findings suggests that taxi drivers are poorer than control subjects at acquiring and retaining associations between objects (or lines in the case of the Rey complex figure) and locations. However, taxi drivers were also worse at associating pairs of words, both at immediate and delayed recall. This could suggest a broader associative deficit within visual and verbal domains. Whilst this may be the case, I note that the words used in the WMS-III

verbal paired associates test are highly imageable. In my view it is likely that this test loads heavily on the visual domain and that overall the taxi drivers' can be characterised as having poorer anterograde visual associative memory. Support for this comes from their performance on another verbal test that would seem to depend on associative processing. Their scores on the immediate and delayed recall of short stories (logical memory, WMS-III) were comparable with the control subjects. In addition, their learning over five trials of a list of words (RAVLT) was also similar to the control participants.

Whilst I suggest the neuropsychological findings point to a reduced facility for anterograde visual associative memory in taxi drivers, it is necessary to consider some additional issues. One of the subtests of the Doors and People Test requires subjects to associate the names and professions of four people [268]. Taxi drivers and controls subjects did not differ on this task (see Table 4) . This subtest, with just four associative items, whilst it may be sensitive to pathology, is not a difficult test for healthy individuals and is considerably easier than the other associative tests in our battery (such as the 16 item object–place associations test and the 18 item Rey complex figure test). I believe that the ease of this subtest meant that in this particular case it was not sensitive enough to differentiate between these two healthy age and IQ-matched groups.

In addition to the Rey complex figure, we also included the modified Taylor figure [277], to examine if poor performance extended to other complex figures. However, there was no difference between the groups on the delayed recall of the modified Taylor figure. Whilst it has been suggested that the two figures are comparable [278], Casey et al. [279] found that the Taylor figure lends itself to a verbal strategy to a greater degree than the Rey complex figure. That the taxi drivers were able to make associations within the verbal domain as noted

above, may explain the lack of difference between the two groups on this test, and suggests that the Rey complex figure may be a purer measure of visuo-spatial memory than the Taylor figure.

Having established that taxi drivers show reduced performance on several anterograde visual associative memory measures, why might this be? It could be associated with the reduced grey matter volume in their anterior hippocampi. The anterior hippocampus in taxi drivers may be less efficient at forming associations. Even when this information is learned to criterion, there may be limited capacity for further consolidation and storage in the posterior hippocampus, given its involvement in supporting the complex spatial representation of London. The evidence for this, however, is not clear-cut. Performance of taxi drivers on the test of London landmark spatial relations correlated with right posterior hippocampal grey matter volume. However, their scores on the anterograde associative tasks on which they performed more poorly than the control participants did not correlate with grey matter volume. This echoes a finding from the previous study where taxi drivers were also poorer at the delayed recall of the Rey complex figure and their scores did not correlate with grey matter volume [151]. Similarly, in this and the previous study the relationship between years experience taxi driving and performance on the memory tests also failed to reach statistical significance.

Performance on the delayed recall of the Rey complex figure and the verbal paired associates test did, however, correlate with mid-hippocampal grey matter volume in control subjects in the current study. The peak voxels in several of these grey matter-memory correlations are within the spatial extent of the area showing decreased grey matter volume in taxi drivers relative to controls. However, the correlation peaks are generally more posterior. The reasons for

the puzzling correlations are not clear, but are typical of the literature on this issue where Van Petten [280] reports difficulties in finding consistent hippocampal volume correlates of standardised memory tests in healthy individuals in the age-range of our participants. Indeed, even in the case of patients, the literature reports only modest relationships between hippocampal volume and severity of memory deficit [281]. In our case, the anterior hippocampal volume measures in taxi drivers did not have a limited range nor were they skewed. We speculate that the lack of correlation in taxi drivers might reflect an early and adverse impact of taxi training on anterograde visuospatial memory, and so a categorical difference rather than a linear relationship with grey matter volume results. Of note the variance was low on these tasks. The correlation data, however, highlight the difficulty of extracting a clear message from this aspect of my study.

What mechanisms might underpin the reduced associative memory processing observed here? In some ways it can be likened to a model of hippocampal processing in aging where it has been proposed that prior memories become the dominant pattern of the hippocampus to the detriment of the ability to encode new information [282]. Whilst the taxi drivers in the current study were relatively young, the extraordinary amount of information they have acquired, information that is known to be dependent on the hippocampus [113], may mimic the effects of an aged hippocampus, with concomitant changes in the balance of information processing. Alternatively, in vivo studies have shown that widespread synaptic strengthening on a population of neurons can cause a shut-down in long-term potentiation (LTP), and may reduce the information storage capacity of the neuronal circuit [283]. In the case of taxi drivers, experience driven grey matter volume changes may push the neuronal circuit

close to its capacity limits and so alter the way the hippocampus processes new information.

The poorer learning ability in licensed London taxi drivers documented in this and the previous study [151], suggests these findings are robust. Whilst the results are intriguing, I acknowledge they are not easily explained and the theoretical implications that can be drawn are somewhat limited at this time. Further studies are required to explore the nature of associative memory in taxi drivers in more depth. This should include establishing if indeed the problem is definitively within the visual domain or whether it might be better characterised as difficulty with between-domain associations [104]. My next experiment aimed to address these issues.

3.5 Experiment 2

Experiment 2, a behavioural study, led on directly from Experiment 1 and involved an entirely new cohort of taxi drivers and control participants. Associative memories may differ depending on whether the items to be associated come from one domain (within domain), whether they comprise items from distinct sensory modalities (between domain), and whether they can be unitized (i.e. perceived as one item) or not. It has been proposed that factors such as these may give rise to neurofunctional differences [104, 284]. It has been shown that after hippocampal damage, recognition of within domain or unitized pairs is largely spared whilst recollection is compromised across all three categories, thereby suggesting that other regions within the medial temporal lobes make contributions to recognition memory [89, 285-288]. In addition, presenting objects within scenes has been suggested to increase the

rate of successful retrieval, this increase is thought to be associated with activity in the parahippocampal cortex that reinstates the visual context to mediate the successful retrieval of the previously learnt objects [360]. In the preceding study I showed that while excelling at London-based tests, taxi drivers were significantly worse than controls on several tests of anterograde associative memory, two of which consisted of making object-place associations. In this experiment I devised a broader range of associative memory tests with the aim of determining specifically which types of associative memory were affected by navigational expertise. The associative memory tests included, visual, verbal, within and between-domain material.

3.5.1 Material and Methods

Participants

Thirty eight healthy male volunteers participated in the study. Of these, 20 were licensed London taxi drivers, and 18 were control participants. The background details of the two groups are shown on Table 5. All taxi drivers had completed “The Knowledge” training, had passed the necessary Public Carriage Office examinations, and obtained a full (green badge) licence. All control participants were resident in greater London. Taxi drivers and control participants did not differ in terms of age ($t(36)=1.85$; $p=0.07$), handedness ($t(36)=0.24$; $p=0.8$), or age when they left school ($t(36)=0.7$; $p=0.4$). Visual information processing and abstract reasoning skills were assessed using the Matrix Reasoning sub-test of the Wechsler Abbreviated Scale of Intelligence [259]. The mean scaled score for both groups did not differ significantly ($t(36)=0.7$; $p=0.4$). An estimate of verbal IQ was obtained using the Wechsler Test of Adult Reading [258]. Data for two taxi drivers and two control participants were not obtained – although they were

very proficient in English, it was not their first language, a requirement of the test. Verbal IQ estimates for both groups were in the average range, and did not differ significantly ($t(32)=1.05$; $p=0.3$).

Table 5 - Participant characteristics

Participant characteristics	Taxi drivers Mean (SD (n=20)	Controls Mean (SD) (n=18)
Age (years)	42.10 (5.37)	38.72 (5.85)
Education (age left school, years)	16.45 (0.94)	16.72 (1.22)
Estimated verbal IQ (WTAR)	98.66 (3.91)	100.31 (5.17)
Matrix reasoning scaled score (WASI)	8.90 (1.88)	8.38 (2.54)
Handedness – laterality index ¹	87.05 (40.46)	83.83 (39.18)
Years experience taxi driving	13.27 (7.86)	-

WTAR=Wechsler Test of Adult Reading; WASI=Wechsler Abbreviated Scale of Intelligence;
¹Edinburgh Handedness inventory

Neuropsychological tests

A test battery was administered that comprised the WMS-III verbal paired associates test (described previously), and number of non-standard associative memory tests which were developed after pilot testing:

Visual paired associates: Stimuli were pairs of unrelated objects presented on a white background (see Figure 23a). The test comprised 16 pairs of objects that were presented on a colour computer monitor every 3 seconds. Having seen all the pairs, there was then a cued recall test where one item from each pair was shown and the participant was required to say the associated object (Figure 23b). There were six encoding-recall trials in total. Delayed recall was assessed after a 30 minute delay. One point was given for each correct answer.

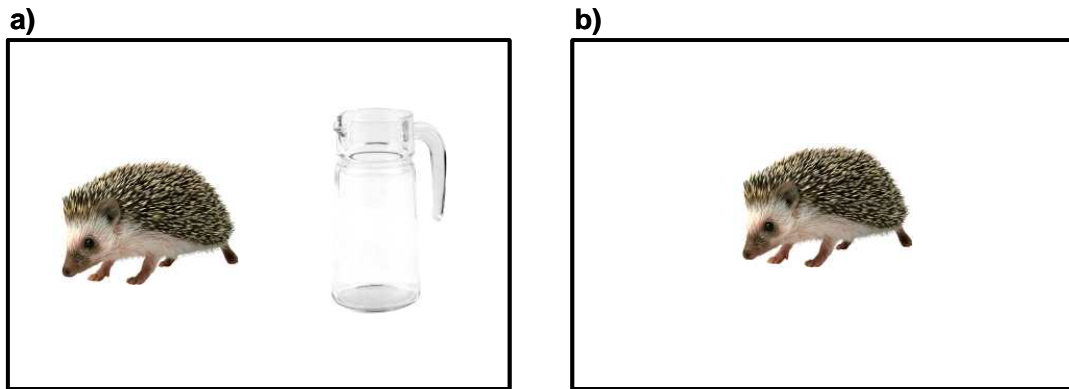


Figure 23. Visual paired associates test.

a) An example stimulus pair of objects. b) An example stimulus from the cued recall. Participants had to name the object that made up the pair.

Visual paired associates within a scene context: Stimuli were pairs of unrelated objects presented on a scene background that could be associated with both objects independently (see Figure 24a). The test comprised 16 pairs of objects that were presented on a colour computer monitor every 3 seconds. Having seen all pairs, there was then a cued recall test where one item from each pair was shown and the participant was required to say the associated object (Figure 24b). There were six encoding-recall trials in total. Delayed recall was assessed after a 30 minute delay. One point was given for each correct answer.

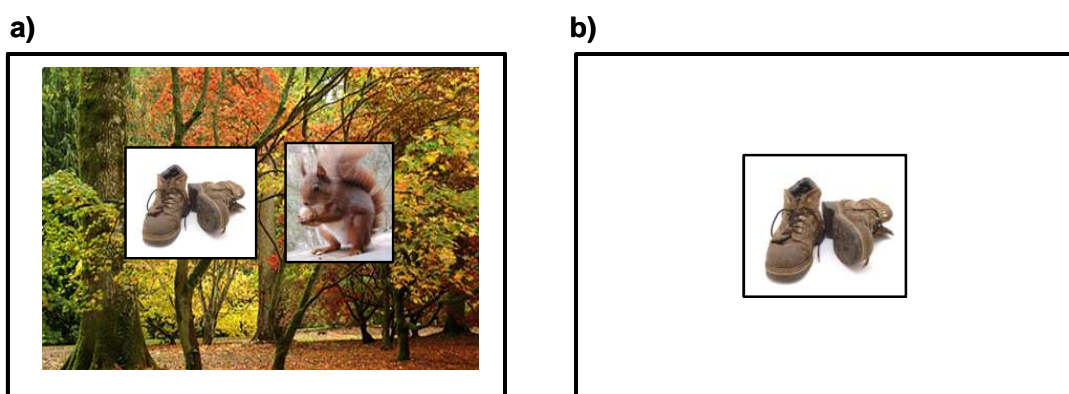


Figure 24. Visual paired associates within a scene context test.

a) An example stimulus pair of objects. b) An example stimulus from the cued recall. Participants had to name the object that made up the pair.

Face-name associations: Stimuli were 16 black and white photographs of a male faces with a name (first and surname) displayed underneath (see Figure 25a). The face-name pairs were presented centrally on a computer monitor every 3 seconds. Having seen all pairs, there was then a cued recall test where the photograph alone was presented (Figure 25b) and the subject was required to give the name of the person in the photograph. One point was awarded for correct first name, one for the correct surname and three points if both were correctly recalled. There were six encoding-recall trials in total, thus each trial had a maximum score of 48 points. Delayed recall was assessed after a 30 minute delay.

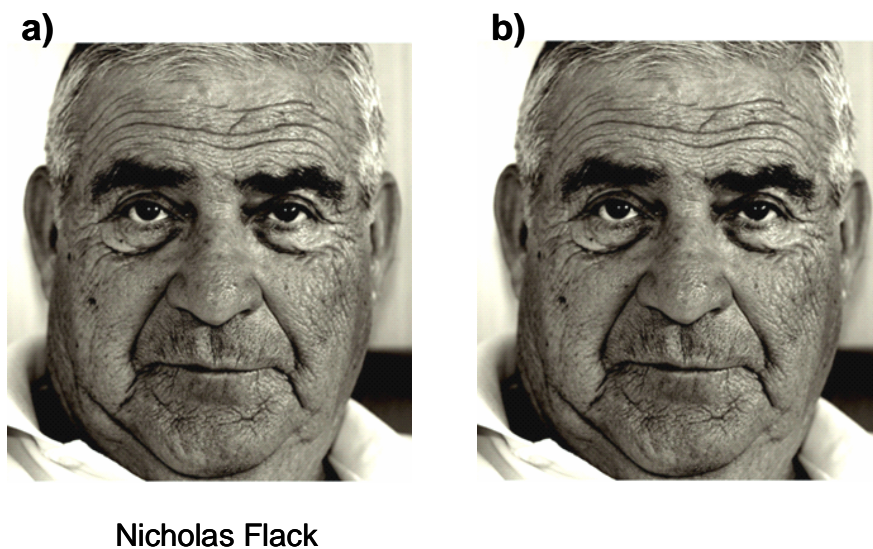


Figure 25. Face-name associations test.

a) An example stimulus of a face-name pair to be learnt in the face-name association test; b) An example stimulus from the cued recall. Participants had to name the person in the photograph.

Object-sound associations: Stimuli were 16 pairs comprising a colour picture of an everyday object presented for three seconds on a computer screen with a white background, and each paired with an unrelated sound presented simultaneously via headphones (see Figure 26a). Having been exposed to all pairs, there was then a cued recall test where the sound alone was presented

(Figure 26b) and the subject was required to name the object they had seen while hearing that sound. There were six encoding-recall trials in total. Delayed recall was assessed after a 30 minute delay. One point was given for each correct answer.

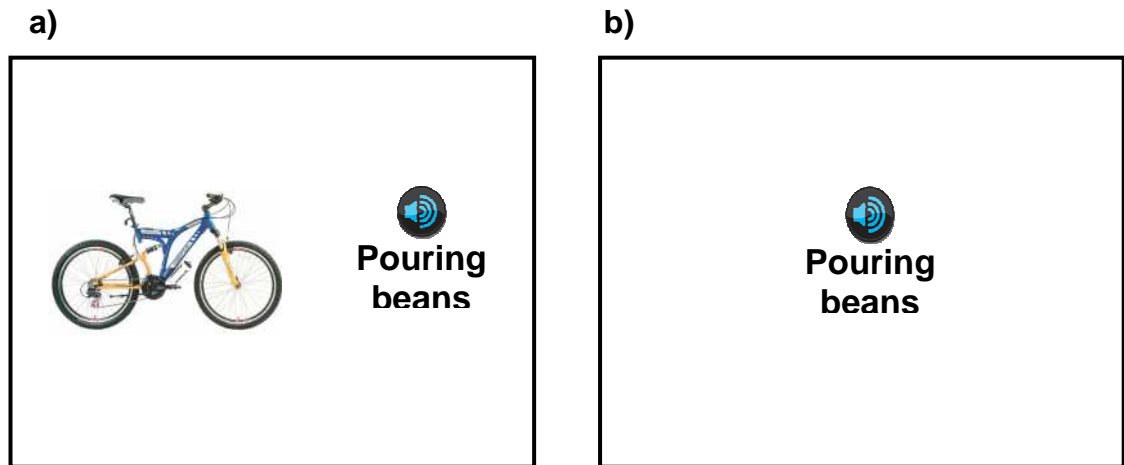


Figure 26. Object-sound associations test.

a) An example stimulus of an object-sound pair to be learnt in the object-sound association test; b) For illustrative purposes, an example sound stimulus from the cued recall (Participants saw a blank screen and heard the sound). Participants had to name the object paired with the sound.

Verbal paired associates test (experimental): The presentation and testing procedures for this test were designed to be identical to that of the standard verbal paired associates test (WMS-III), i.e. there were 8 pairs of words. However, in this new version, the imageability score for all words was equal to zero on the Colheart norms [289], and the frequency for the words included was in the low range (1-116) according to the Kucera-Francis (1967) written frequency count. The norms were obtained from the MRC Psycholinguistic Database (www.psych.rl.ac.uk/MRC_Psych_Db_files/mrc2). The idea with this test was to try to lessen the extent to which the words were imageable. As with the standard test, during the study phase the participant was instructed to remember each pair of words. This was followed by a test phase in which one word out of each pair was presented with the requirement to state the

associated word, receiving one point. There were six encoding-retrieval trials (Figure 27).

Pairs	Recall (cue)
Advice – decrease	Advice
Certainty – begin	Yearly
Verify - impatient	Litigate
Yearly - honesty	Relief
Void – novel	Verify
Litigate - Buffer	Certainty
Relief – Discovery	Void
Logical – deny	Logical

Figure 27. Verbal paired associates.

The list of pairs of words used in the verbal paired associates test

Object-place associations test: This test is described above in Section 3.2.1 and was included in this experiment to assess if the results would be replicated in a new cohort of participants.

3.5.2 Procedure and data analyses

These were the same as for the behavioural tests in Experiment 1.

3.5.3 Results

A MANOVA was performed on the associative memory measures listed on Table 6. There was a significant difference between the two groups ($F(1,36)=2.57$; $p=0.02$). The source of this difference was investigated using the tests of between-subjects effects produced by MANOVA. There were two main effects. Taxi drivers took significantly longer to reach criterion on the object-place association test ($F(1,36)=9.70$; $p=0.004$), and recalled fewer of the object-place associations after a delay ($F(1,36)=8.84$; $p=0.005$).

Table 6 - Associative memory measures

Associative memory measures	Taxi drivers Mean (SD) (n=20)	Controls Mean (SD) (n=18)
Visual paired associates IR (Σ for 6 trials /96)	87.55 (5.09)	85.22 (6.55)
Visual paired associates DR (/16)	14.75 (.91)	15.11 (.96)
Visual paired associates within a scene context IR (Σ for 6 trials /96)	91.85 (3.93)	90.05 (3.45)
Visual paired associates in context DR (/16)	15.30 (1.03)	15.33 (.90)
Face-name associations IR (Σ for 6 trials /288)	110 (52.07)	125.44 (55.64)
Face-name associations DR (/48)	27.10 (12.80)	30.22 (10.66)
Object-sound associations IR (Σ for 6 trials /96)	80.95 (6.77)	79.22 (10.69)
Object-sound associations DR (/16)	13.65 (1.53)	14.22 (1.66)
Verbal paired associates – IR (Σ for 6 trials /48)	22.95 (7.81)	28.00 (8.78)
Verbal paired associates – DR (/8)	4.20 (1.39)	5.55 (1.38)
Object-place associations – number of trials to criterion ^c	2.95 (0.82)	2.11 (0.83)
Object-place associations – DR (number correct /16) ^c	12.65 (1.53)	14.27 (1.84)
Verbal paired associates – IR (WMS-III) scaled score	12.85 (2.49)	11.88 (2.21)
Verbal paired associates – DR (WMS-III) scaled score	11.80 (1.60)	12.00 (1.74)

IR=immediate recall; DR=delayed recall after 30 minutes ^cControl subjects significantly better than taxi drivers

3.5.4 Discussion

This experiment was designed to follow-up on the results on Experiment 1, in exploring in a more in-depth fashion the apparent anterograde associative memory problems shown by taxi drivers. I found that in fact taxi drivers were unimpaired relative to control participants on a range of associative memory tasks, except for object-place associations.

The results of Experiment 1 showed that, as well as being impaired on the delayed recall of the Rey complex figure, taxi drivers were also significantly poorer than controls at learning the locations of objects in a table-top array, and

also at recalling the object-place associations after a delay. These effects appear to be robust, having now been replicated here in a new set of participants. In the previous experiment, taxi drivers were also impaired on the verbal paired associates test from the WMS III. However, in Experiment 2, this effect did not replicate, with no difference between the scores of the taxi drivers and control participants. Moreover, an alternative version of this test that I designed to be less reliant on visual imagery also failed to distinguish between the two groups. Thus, Experiment 2 allows me to conclude that associative memory within the verbal domain is not affected by navigational expertise, and the verbal paired associates effect found in Experiment 1 is not reliable.

However, what about associative memory within the visual domain? Two tests of visual associative memory were employed, one involved two unrelated objects presented centrally on a white background, and the other involved objects, whilst unrelated to each other, they were presented centrally on a related background scene context. The latter was included in case taxi drivers were impaired on the former, the intention being to examine whether the addition of a background context might aid performance. The findings showed the taxi drivers were unimpaired on both tasks, suggesting that visual associative memory is intact in taxi drivers.

Given the deficit on object-place associations, I also investigated whether they might have a more general deficit when stimuli spanned more than one domain. Face-name associations, and semantically unrelated object-sound associations comprised two further tests. In neither case was the taxi driver group impaired relative to the control participants, suggesting that a general problem with between domain associative memory is not a pervading factor in their memory profile.

Overall, the results of Experiments 1 and 2 which involved an extensive range of memory tests and associative memory measures, permit a number of conclusions. First, the pattern of hippocampal grey matter volume documented initially by Maguire et al. [147] and replicated in Maguire et al. [151] is truly robust, with yet another replication in Experiment 1. Second, taxi drivers, unsurprisingly, show a consistent and replicable advantage over controls on tests of London-based topographical knowledge. Third, taxi drivers' extensive spatial London knowledge appears to come at something of a cost, in significantly poorer performance on two tests of visuo-spatial memory. Specifically on the delayed recall of the Rey complex figure, and the learning and delayed recall of object-place associations, with these findings now replicated across different cohorts. The data show that these deficits are not due to problems with memory per se, or with anterograde associative memory in general, or to problems with visual, or between domain associative memory. Rather, the deficits seem to be specific to these two tasks, the Rey complex figure and object-place associations.

Both of these tasks involve visuo-spatial associations, so perhaps it is the spatial aspect that is particularly sensitive to and constrained by navigational expertise. Unlike the Rey and the object-location tests, none of the other tasks in Experiments 1 and 2 involved a number of spatial locations in the stimuli/elements. As considered in the Discussion of Experiment 1, the extreme demands in object-location associations placed on taxi drivers through their training and everyday navigation around London, may reduce the resources specifically for the processing of new object-place associations. Further evidence relating to the issue of space and navigation will be reported in the next two chapters, and I will then consider this issue further, and the overall pattern of spared and impaired function in taxi drivers in Chapter 7.

Chapter 4

The effects of navigational expertise on wayfinding in new environments

Précis

In Experiments 1 and 2, I focused on examining the neuropsychological, and in particular the memory profile of taxi drivers using standardised and table-top tests. However, the scale of the tasks used is arguably quite different from the kind of spatial tasks on which the taxi drivers excel, namely navigating in large-scale space. It is therefore important to extend consideration of the effect of navigational expertise further by examining whether or not taxi drivers can integrate new information into their established representation of London's layout, and also if they can learn the layout of an entirely new town or city. Examining these issues was the aim of Experiment 3.

4.1 Introduction

Finding your way around a large city can be challenging. All the more so if the layout of the city in question is complex [290]. While cities in North America are often laid out in a regular grid pattern, by contrast many European cities have chaotic, unpredictable layouts, a prime example being London (UK). Despite its complexity, licensed London taxi drivers are expert at navigation around central London. In Chapter 3 (see also [291]) I documented taxi drivers' expertise in London knowledge compared with control participants, including London bus drivers [146, 150], who also drive customers along London's streets, but using a constrained set of routes. Notably, however, the wayfinding expertise of taxi drivers seemed to come at a cost. They performed more poorly than control participants on several memory tests involving new visuo-spatial materials. Specifically, they recalled fewer details of a complex figure (Rey-Osterrieth complex figure [271]) after a delay [151, 291]. Similarly, they took longer to

learn associations between sixteen objects and locations on a table-top array, and also had poorer recall of the object-location pairs after a delay [291].

Thus, taxi drivers who excel at wayfinding around a complex city were poor at acquiring some types of new visuo-spatial information that control participants could learn with ease. However, it could be argued that the types of table-top tasks employed in previous studies did not directly assess the spatial processing that taxi drivers typically engage in when immersed in a large-scale complex space such as an urban environment. As such, a key question remains unanswered, namely, does being a very skilled navigator in one environment confer an advantage when learning the layout of a new and unfamiliar environment? Or, do the decrements observed on some table-top spatial learning and recall tasks in taxi drivers suggest that the ability to learn a new spatial layout might be compromised in the context of navigational expertise? Both outcomes have implications for understanding potential mechanisms of wayfinding, and the effect of navigational expertise on memory and hippocampal grey matter volume. For instance, if wayfinding expertise facilitates learning a new environment, this could suggest that expertise might in part be underpinned by the development and use of generalisable and effective strategies for wayfinding. By contrast, if learning a new environment is constrained, this may indicate that expertise and detailed existing environmental knowledge limits the capacity for processing or storage of new spatial layouts, or results in a bias towards existing knowledge, similar to some models of ageing [282].

The central aim, therefore, of this study was to ascertain if taxi drivers' navigational expertise in London affected their ability to learn the layout of a new environment. A range of tests used in previous wayfinding studies [291-298] was employed to assess knowledge of environmental features and

topographical details, spatial relationships between landmarks, planning and execution of routes, and whether an overall survey-like representation had been acquired [291, 297, 299]. I examined the acquisition of new environmental knowledge in two ways; first, by testing how well taxi drivers could learn the layout of a previously unfamiliar town compared with a control group of non-taxi drivers. Second, I investigated how effectively taxi drivers could integrate a new district into their existing spatial representation of London. In this latter case, it was not possible to test control participants, as they simply did not have a comparable knowledge of London's layout as a basis for integration. Consequently, the key comparison here was between taxi drivers' performance in the entirely new town and their performance in the new area of London.

4.2 Experiment 3

4.2.1 Materials and methods

Participants

The subjects who took part in Experiment 2, also participated in this experiment. A total of forty healthy male volunteers participated. Of these, 20 were licensed London taxi drivers, and 20 were control subjects. – see Chapter 3 for their full details.

Environments and tests

The effect of navigational expertise on environmental learning and knowledge was assessed in two ways, firstly using an unfamiliar town - "New Town", and secondly using modifications to the familiar environment of London (UK) – "London".

New Town

Film footage was acquired of navigation along two routes through an urban environment (see Figure 28). Stimulus material was adapted from that used in [296], and featured a town called Blackrock which is south of Dublin City, Ireland. There was no footage in common between the two routes except for brief navigation across one point of overlap at a central road junction. Footage was shot in colour with a wide angle lens, at eye-level, and at average walking pace. The camera panned from side to side, to simulate natural viewing and in order to include the salient features along each route such as prominent buildings and shop fronts. When a road junction was reached, the pace slowed and the camera panned down all the elements of the road junction before moving on. The two routes were presented one after another on a computer screen, with a total viewing time of 4 minutes 53 seconds. As the primary domain of interest in this study was visual, sound was not included. None of the participants was familiar with the environment. The film was shot in the main shopping area of the town. Landmarks were defined to participants as prominent buildings and distinctive shops/businesses. The following instructions were also given at the start of the test:

You are going to see films of navigation along two overlapping routes through a town. It will proceed at a brisk walking pace and the camera will pan and move as if you're looking around while walking along. You should try to remember as many of the landmarks (buildings, shops, etc) as possible. Ignore cars, buses and people as they are not important. After each viewing of the footage, you will see a series of very short film clips – you will need to indicate if they were part of the routes you have just seen or not.

Participants viewed the footage four times. To ensure they paid attention and to check that learning was occurring, after each viewing, participants were shown four clips lasting 3 seconds each and asked to indicate if they formed part of the route they had just seen or not. Two were actual clips from the routes, and two

were similar but never-seen distractor clips. There were new clips for each viewing (i.e. each test clip was shown only once during learning).

After participants had viewed the footage four times, they then completed a number of tests designed to assess their knowledge of the environment they had just learned.

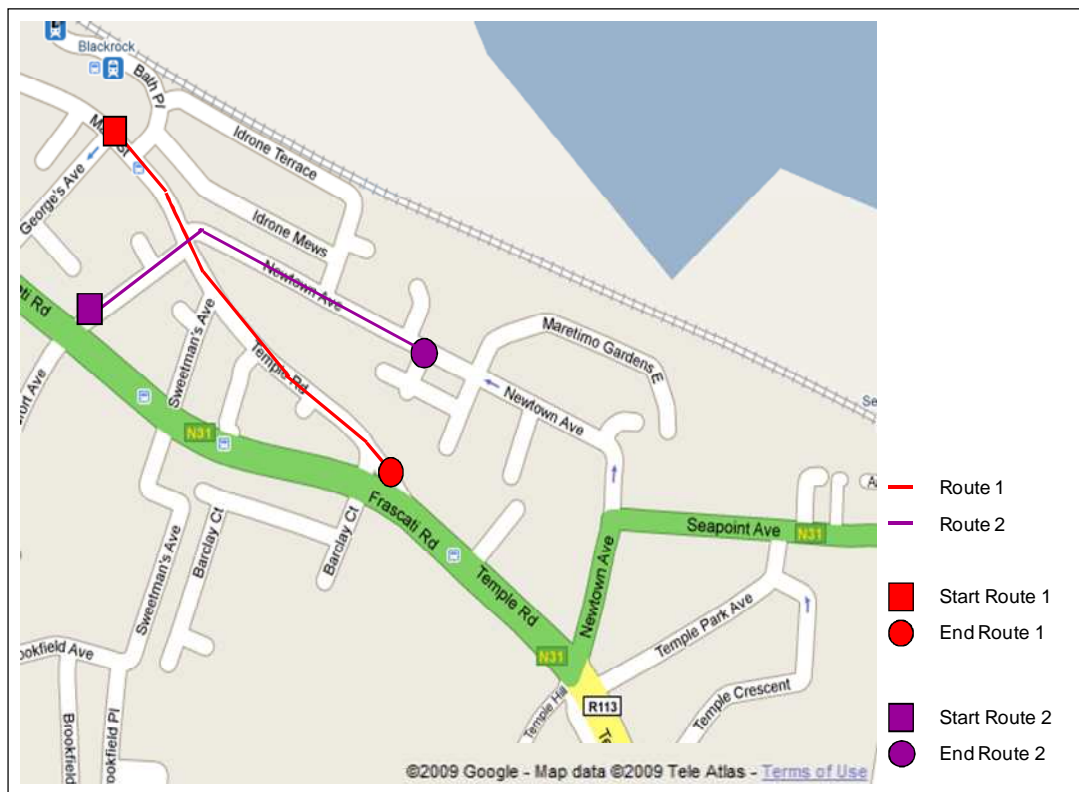


Figure 28. Map of New Town.

The two overlapping routes are shown. Note that participants never saw this map. © Google Maps.

New Town scene recognition memory test: Recognition memory for environmental features and topographical details was assessed by showing participants 32 colour photographs of scenes. Twelve were scenes from route 1 in New Town, twelve were scenes from New Town's route 2, and eight were distractor scenes that were not from New Town, but were visually similar. Target and distractor scenes were randomly intermixed. The format was a yes/no

recognition memory test where participants were asked to state whether they recognised each scene as a part of New Town or not. The test was not formally timed, however subjects took on average 2-5 seconds per photograph.

New Town proximity judgements: Participants' knowledge of the spatial relationships between landmarks in New Town was tested using a proximity judgements task. Stimuli were colour photographs each depicting a New Town landmark (see example in Figure 30A). On each trial, subjects had to judge which of two other New Town landmarks was closer (as the crow flies) to the target landmark. There were ten trials of which three trials comprised landmarks solely within route 1, three trials where landmarks were all within route 2, and four trials where landmarks were taken from both routes. The test was not formally timed, however participants took on average 5-8 seconds per trial.

New Town route execution: To test the ability to plan and 'navigate' along routes, participants were given a photograph of a New Town landmark labelled 'start point' and another labelled 'end point'. Six additional landmark photographs were then supplied, with the instruction to place these in the correct order that they would be passed en route between the start and end points. There were four route execution trials, one where all landmarks were in route 1, another where all landmarks were in route 2, and two that spanned both routes (i.e. started in one route and ended in the other). Each correctly placed photograph was given a score of 1. The mismatch between the presented and correct order was derived by calculating the vector distance $[\sum(x-y)^2]$ between the position presented for each photograph (y) and what the correct position for that photograph should have been (x). Thus a score of zero indicates a perfect match between the presented and actual placements for a given trial.

Of note, a previous study found no performance differences when within and between route trials were compared on either the proximity judgements task or the route execution task [296]. Consequently, in this study, analysis was collapsed across within and between trial types.

New Town sketch map: Participants were asked to draw a sketch map of the routes seen in the film footage, including any landmarks they could recall. It was made clear that drawing ability would not be assessed. Standard A3 size white paper (297x420mm/11.7x16.5inches) and pencils were used. Participants were not provided with erasers. If they wished to restart, they were permitted to do so on the reverse of the sheet, although this never occurred. In two cases, more paper was requested as the map exceeded the space on one page. The following variables were examined:

- number of road segments, where a segment referred to a section of road between road junctions
- number of road junctions
- number of correct landmarks
- landmark placement, with a maximum of three points per landmark, where one point was given if the landmark was on the correct side of the road, one point if it was correctly placed in relation to nearby road junctions, and one point if it was in the correct sequence relative to preceding and subsequent landmarks
- orientation score – an experimenter rating that assessed how the roads and layout were orientated, on a scale of 1 to 5, where 1 was poor and incorrect...5 was good and accurate orientation
- overall map categorisation score – an experimenter rating on a scale of 1-6.

The map categorisations for 'New Town' were based on the range of previous data obtained using the same stimuli [296]. As such, the categories represented distinct progressions in the quality of map coherence.

- The two routes were merged into one
- There were two routes, but they were perceived as separate
- The two routes were close together, but not joined accurately
- Some elements across the two routes were joined up, but integration was broadly lacking
- The two routes were integrated, but there were some inaccuracies in layout
- Correct integration across the two routes, survey-like map, accurate and easy to follow

Twenty percent of the sketch maps were also scored by a second rater, blind to group membership; the inter-rater correlation was 0.99.

London

In the first instance we sought to verify if taxi drivers had significantly more knowledge about London's layout than the control participants. To do this we used a test shown in several previous studies to be a reliable measure of topographical knowledge ([147, 151, 291]; see also Chapter 3). The London landmarks proximity judgements test comprised colour photographs each depicting a famous London landmark. On each trial, subjects had to judge which of two other London landmarks was closer (as the crow flies) to the target London landmark. There were 10 trials. The test was not formally timed, however subjects on average took 5-8 seconds per trial.

Having assessed how well taxi drivers could learn a novel environment (see New Town above), we were also interested in whether they could integrate new environmental knowledge into their existing cognitive representation of London. To examine this, we devised the London test. The London test was only given to taxi drivers, as control participants could not be compared to taxi drivers in their basic level of London knowledge.

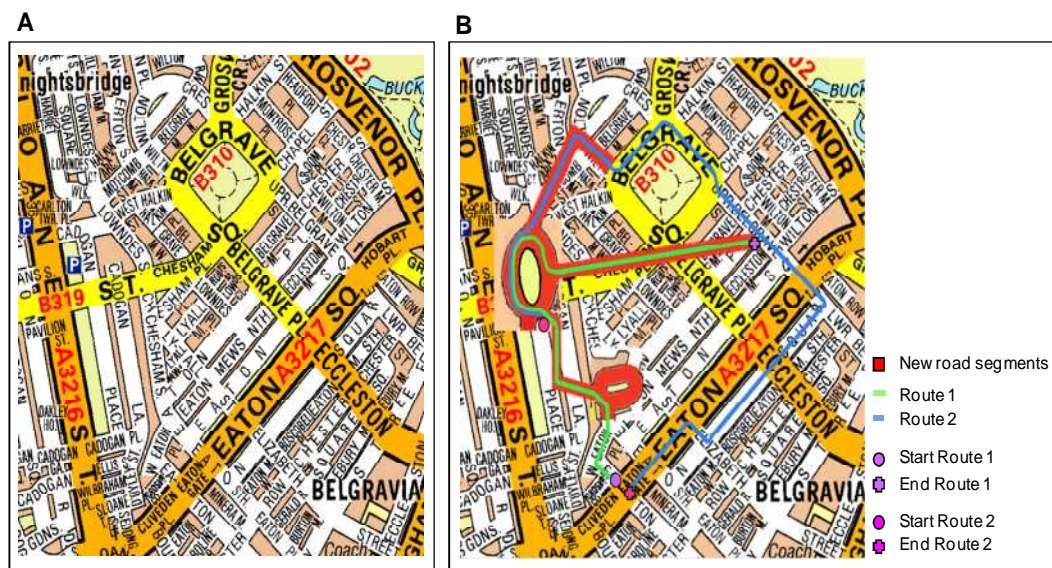


Figure 29. Map of London.

A) London as it is normally. B) A map showing existing London integrated with 'new' London, where modifications are depicted in red. Note that participants never saw these maps. Maps reproduced by permission of Geographers' AtoZ Map Co. Ltd. - Crown Copyright 2005. All rights reserved. Licence number 100017302.

Footage was presented on a computer screen and comprised continuously updating colour photographs; each photograph was on the screen for 2 seconds before the next photograph in the sequence appeared. Pilot testing determined this was a comfortable pace for viewing. The photographs (see Figure 30B) were taken at eye level, with a wide angle, and in an evenly-spaced fashion to mimic walking, giving the impression of navigation along routes. Participants (and piloting) confirmed that this readily conveyed the sensation of walking through the environment. The routes depicted in the photographs were made

up of an existing part of London with some modifications. Modifications involved diverting participants from existing London into new areas they were not familiar with – ‘new London’ (see Figure 29). Photographs depicting these new areas were shot in another city (Bath, UK) with buildings of broadly similar character and historical period as the existing part of London (see Figure 30C). None of the participants was familiar with this other city. Several criteria guided the development of the London task: first, I needed to match the overall appearance and architecture of that part of existing London with somewhere similar (but not too distinctive in and of itself). Next, I did not want to have too many links between existing and new London, but rather a sufficient number to test the taxi drivers whilst preserving the logic of connections between existing and new, without too much confusion or a sense of ‘weirdness’. Finally, I wanted to approximately equate the amount of information in the London task with that in New Town.

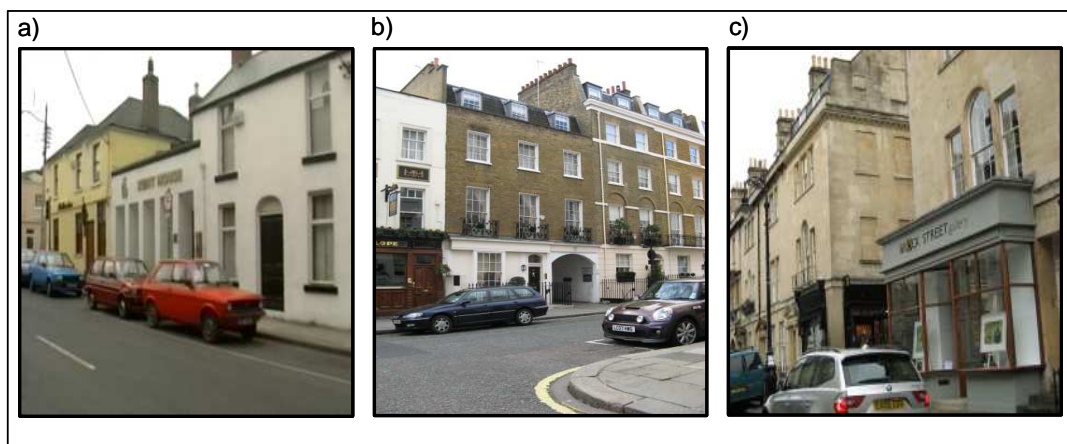


Figure 30. Example views from the three environments.

a) Photograph taken in New Town. b) A view from existing London. c) A photograph from new London.

As with the New Town task, there were two overlapping routes. Each route was made up of segments from existing London and new London. Overall, the London test comprised 47% existing London and 53% new additions. I elected

to use photographic stimuli in this instance, rather than film as used for the New Town task, because this enabled a seamless transition between existing and new London. All transitions between existing and new London involved a turn (left or right) in order to avoid transition points appearing unrealistic or confusing. Text instructions appeared on the screen indicating that a turn was coming up. The difference between the film and photographic presentations was never raised as an issue by participants or during piloting. Previous work has suggested a relatively high correspondence in response to colour photographs and on-site presentations [300-302]. Whilst I cannot rule out a potential effect of presentation mode, given the naturalistic feel associated with both types of stimuli, I do not believe it significantly affected the results. The following instructions were given at the start of the London test:

You are going to see footage of navigation along two overlapping routes through London. The area will seem familiar but some of it has been rebuilt after a large number of buildings collapsed, and some of the roads have also been replaced. It will proceed at walking pace and the camera will pan and move as if you're looking around while walking along. You should try to remember as many of the landmarks (buildings, shops, etc) as possible. Ignore cars, buses and people as they are not important. After each viewing of the footage you will see a series of very short film clips – you will need to indicate if they were part of the routes you have just seen or not.

Learning and memory was assessed in the same way as described above for New Town. There were four exposures to the footage with four short test clips after each viewing, one from existing and one from new London, plus two distractor clips. There were new clips for each viewing (i.e. each clip was shown only once during learning). Testing then proceeded with a scene recognition memory test, a proximity judgements test, a route execution test, and the drawing of a sketch map. For the London test, three additional adjustments were made. (1) The London scene recognition test, just like the New Town test, comprised twelve scenes from each route and eight distractor scenes. Ten of

the target scenes were from the new additions to London, while fourteen scenes were from existing London. (2) The sketch maps were scored in two ways. Initially the complete sketch maps were analysed in an identical fashion to New Town. In order to examine the taxi drivers' knowledge of specifically the new additions to London in the context of existing knowledge, the aspects of the map relating to existing and new London were also scored separately. (3) An additional sketch map experimenter rating was included for the London test in order to capture how well the existing and new parts of London were integrated, on a scale of 1-5, where 1=little or no integration... 5=good integration. Twenty five percent of the New London sketch maps and experimenter ratings were also scored by a second rater; the inter-rater correlation was 0.99.

4.2.2 Procedure

Each participant was tested individually. The New Town and London tests were administered to the taxi drivers in separate sessions that were at least 1 week and no more than 3 weeks apart (these sessions also included the associative memory tests described in Experiment 2). Participants were debriefed following the two sessions in order to obtain feedback on how they found the New Town and London tests, and to make comparisons between the two learning experiences.

4.2.3 Data analysis

As described in Chapter 2, group comparisons relating to participant characteristics were made using two-tailed t-tests. For the main between-group analyses MANOVAs were employed. For the main within-group analyses (taxi drivers only), data were screened for outliers, homogeneity of variance, and to ascertain if the data were normally distributed. One-way analysis of variance

(ANOVA) was employed with the significance threshold set at $p < 0.05$. Two separate one-way ANOVAs were performed. In the first, town (New Town, London) was the independent variable, and the main environmental knowledge measures were the dependent variables. In the second ANOVA, environment type (New Town, existing London, new London) was the independent variable, and a range of the environmental knowledge measures were the dependent variables. Where ANOVA indicated a significant effect, post-hoc tests using Bonferroni correction were employed to ascertain the source of the significance with a threshold of $p < 0.05$. In addition, where variables were not included in the ANOVAs, two-tailed paired t-tests were employed (see further details in Results). To reduce the chance of Type I error, Bonferroni correction ($p = 0.05/n$, where n is the number of t-tests) was applied. Effect size was calculated using Cohen's d , and is reported where relevant (i.e. for significant differences between two means). Correlations were also performed between the number of years taxi driving and all of the environmental knowledge measures, although none were significant.

4.2.4 Results

New Town

Mean test scores are shown in Table 7. There was no difference between taxi drivers and control participants in their ability to recognise the film clips that were shown during the initial viewing phase ($t(36) = 0.88$; $p = 0.38$). A MANOVA was then performed with Group (taxi drivers, control participants) as the independent variable, and the main New Town environmental knowledge measures (scene recognition, proximity judgements, route execution, sketch map number of road segments, sketch map number of road junctions, sketch map number of landmarks, sketch map landmark placement) as the dependent variables. This revealed a significant difference between the groups

($F(1,30)=2.52$; $p=0.03$). The source of this difference was investigated using the tests of between-participant effects produced by MANOVA. There were five main effects. Taxi drivers were significantly better than control participants on route execution ($F(1,36)=5.72$; $p=0.02$; $d=0.77$), on sketch map number of road segments ($F(1,36)=13.96$; $p=0.001$; $d=1.22$), sketch map number of road junctions ($F(1,36)=8.64$; $p=0.006$; $d=0.95$), sketch map number of landmarks ($F(1,36)=6.11$; $p=0.01$; $d=0.82$), and sketch map landmark placement ($F(1,36)=5.60$; $p=0.02$; $d=0.77$). In addition, taxi drivers were also rated better on the sketch map orientation score ($t(36)=2.52$; $p=0.01$; $d=0.82$) and on the overall sketch map categorisation score ($t(36)=3.27$; $p=0.002$; $d=1.07$) (see Figure 31 for example sketch maps).

Table 7 - Performance of both groups on the New Town tests

New Town	Taxi drivers Mean (SD) (n=20)	Controls Mean (SD) (n=18)
Learning		
Short film clip recognition (/16)*	15.80 (.52)	15.61 (.77)
Environmental knowledge		
Scene recognition (/32)*	22.3 (2.95)	22.1 (3.19)
Proximity judgements (/10)	7.1 (1.44)	6.5 (1.29)
Route execution (vector distance, where 0 is perfect performance, and a larger score is poorer) ^{TD}	46.25 (31.32)	70.22 (30.30)
Sketch map number of road segments (/16) ^{TD}	9.05 (3.25)	5.44 (2.61)
Sketch map number of road junctions (/8) ^{TD}	4.30 (1.55)	2.61 (1.97)
Sketch map number of landmarks (/28) ^{TD}	12 (4.63)	8.72 (3.35)
Sketch map landmark placement (/84) ^{TD}	27.65 (14.22)	17.83 (10.89)
Ratings		
Sketch map orientation (scale 1-5) ^{TD}	3.55 (.99)	2.72 (1.01)
Sketch map overall map categorisation (scale 1-6) ^{TD}	4.05 (1.39)	2.66 (1.18)

*Includes correct detections and correct rejections. ^{TD}Taxi drivers significantly better than control participants.

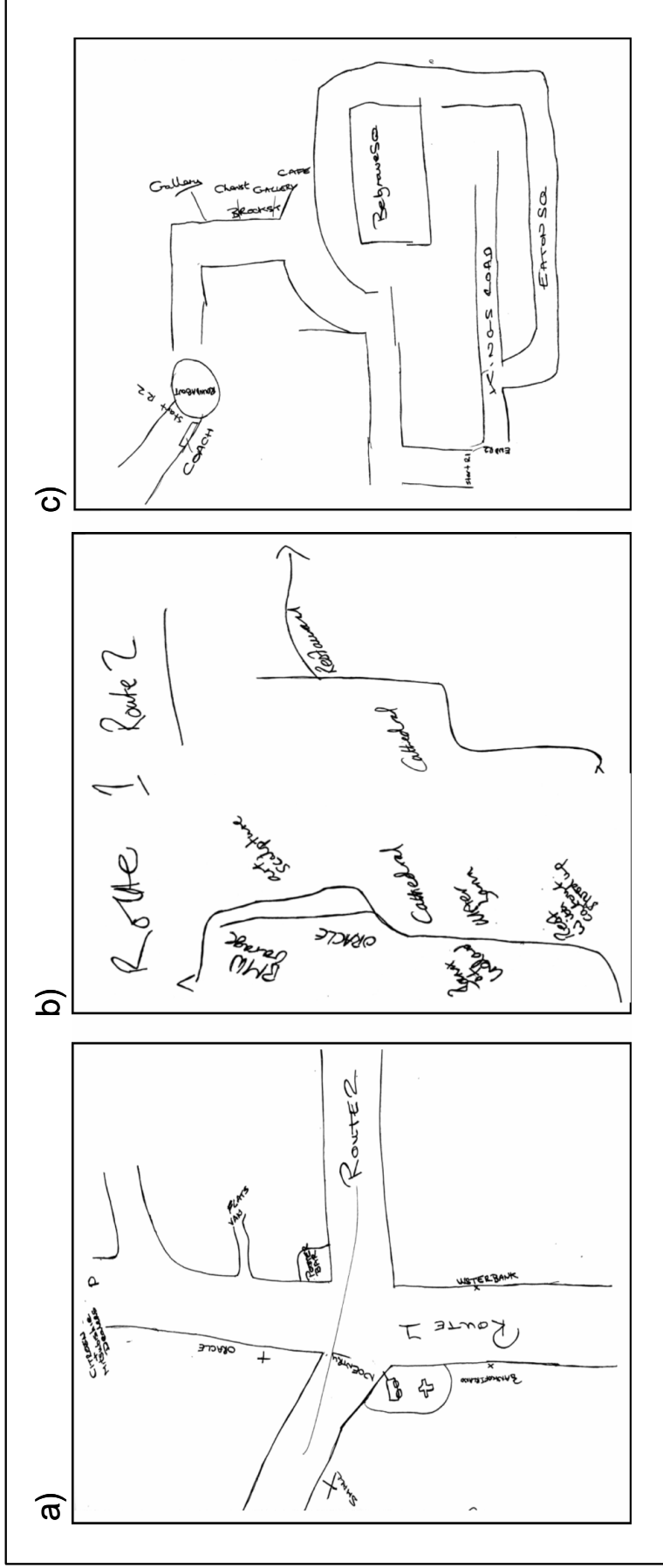


Figure 31 - Example sketch maps.

a) A taxi driver's sketch map of New Town. b) A control participant's sketch map of New Town. c) A map of existing and new London, as drawn by the same taxi driver whose map of New Town is shown in A.

London

I first sought to examine whether taxi drivers had significantly better knowledge about London's layout than the control participants; this was measured using the London landmarks proximity judgements test. Taxi drivers mean score on this test was 8.75/10 (SD 0.91) compared with 7.61 (1.19) for control participants, with taxi drivers performing significantly better ($t(36)=3.32$; $p=0.02$; $d=1.08$).

I then compared taxi drivers' performance on the London test with their performance in New Town (see mean London scores on Table 8). The raw scores were converted to proportion scores to make comparison possible across the two towns.

There was no difference in the ability to recognise the film clips that were shown during the initial viewing phase of New Town and London ($t(19)=0.32$; $p=0.74$). A one-way ANOVA was used to test for differences between London and New Town on seven environmental knowledge measures (scene recognition, proximity judgements, route execution, sketch map number of road segments, sketch map number of road junctions, sketch map number of landmarks, sketch map landmark placement). There were two main effects. Scene recognition was better for London compared with New Town ($F(1,38)=31.29$; $p=0.001$; $d=1.77$), while sketch map number of landmarks was better for New Town ($F(1,38)=7.25$; $p=0.01$; $d=0.85$). There was no effect for proximity judgements ($F(1,38)=0.11$; $p=0.73$), sketch map number of road segments ($F(1,38)=0.15$; $p=0.69$), sketch map number of road junctions ($F(1,38)=2.34$; $p=0.13$), sketch map number of correctly placed landmarks ($F(1,38)=2.38$; $p=0.13$), or route execution ($F(1,38)=3.23$; $p=0.08$). Experimenter ratings of sketch map orientation ($t(19)=1.63$; $p=0.12$) and overall sketch map categorisation ($t(19)=0.96$; $p=0.35$) did not differ between the two environments.

Table 8 - Performance of taxi drivers on the London tests

London (overall)	Taxi drivers Mean (SD)
Learning	
Short film clip recognition (/16)*	15.75 (.55)
Environmental knowledge	
Scene recognition (/32)* ^L	27 (2.31)
Proximity judgements (/10)	6.95 (1.35)
Route execution (vector distance, where 0 is perfect performance, and a larger score is poorer)	62.85 (26.91)
Sketch map number of road segments (/25)	8.57 (2.34)
Sketch map number of road junctions (/23)	7.8 (3.69)
Sketch map number of landmarks (/19) ^{NT}	4.47 (2.34)
Sketch map landmark placement (/57)	11.75 (7.95)
Ratings	
Sketch map orientation (scale 1-5)	3.15 (1.08)
Sketch map overall map categorisation (scale 1-6)	3.8 (1.54)
Sketch map integration of existing and new London (scale 1-5)	2.6 (1.14)

*Includes correct detections and correct rejections. ^LSignificantly better performance for London compared with New Town. ^{NT}Significantly better performance for New Town compared with London.

My main interest was in examining specifically the acquisition of new environmental knowledge on its own (New Town), and in the context of existing knowledge (new additions to London). A one-way ANOVA was used to test for differences among three environment types, New Town (NT) existing London (EL) and new London (NL) on five environmental knowledge measures (scene recognition, sketch map number of road segments, sketch map number of road junctions, sketch map number of landmarks, sketch map landmark placement sketch map scores). In order to directly compare performance across the three environment types raw scores were converted to proportional scores (see mean proportional scores on Table 9). It was not possible in this instance to include sufficient trials to compare New Town and new London on the proximity and

route execution tests. There were four main effects across the three environment types, scene recognition ($F(2,57)=26.18$; $p=0.001$), sketch map number of road segments ($F(2,57)=4.32$; $p=0.001$), sketch map number of road junctions ($F(2,57)=8.22$; $p=0.001$), and sketch map number of landmarks ($F(2,57)=4.77$; $p=0.012$). There was no effect for number of correctly placed landmarks ($F(2,57)=2.35$; $p=0.105$).

Bonferroni post-hoc comparisons of the three environment types showed that the number of correctly recognised scenes in EL was significantly higher than the number of correctly recognised scenes in NL ($p=0.001$; $d=1.65$), and in NT ($p=0.001$; $d=2.50$). There was no difference in the number of correctly recognised scenes between NL and NT ($p=0.25$) (Figure 32).

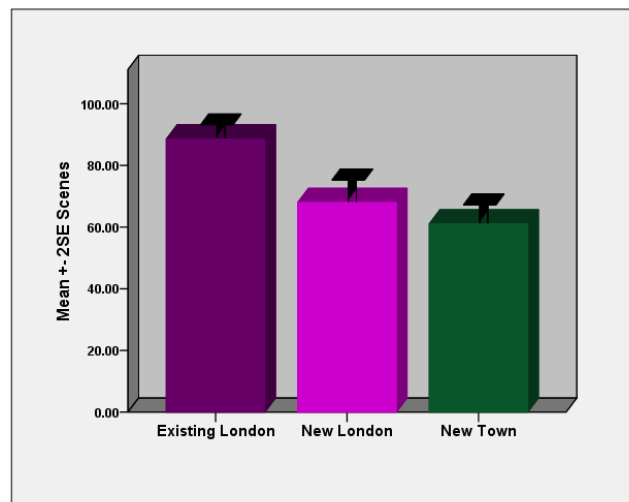


Figure 32. Number of correctly recognised scenes across environments

The number of road segments recalled for EL was significantly higher than the number of road segments recalled for NL ($p=0.001$; $d=1.64$), while the number of road segments recalled for NT was significantly higher than the number of

road segments recalled for NL ($p=0.004$; $d=1.08$). There was no difference in the number of road segments recalled for EL and NT ($p=0.198$) (Figure 33).

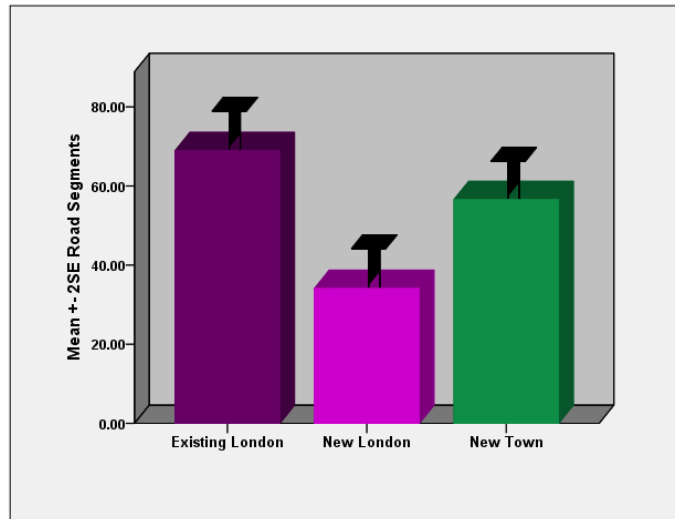


Figure 33. Number of road segments recalled across environments

The number of road junctions recalled for EL was significantly higher than the number of road junctions recalled for NL ($p=0.002$; $d=1.12$), while the number of road junctions recalled for NT was significantly higher than the number of road junctions recalled for NL ($p=0.004$; $d=0.98$). There was no difference in the number of road junctions recalled for EL and NT ($p=1.00$) (Figure 34).

The number of landmarks recalled for NT was significantly higher than the number of landmarks recalled for NL ($p=0.009$; $d=1.08$). There was no difference in the number of landmarks recalled for EL and NT ($p=0.28$), or for EL and NL ($p=0.51$).

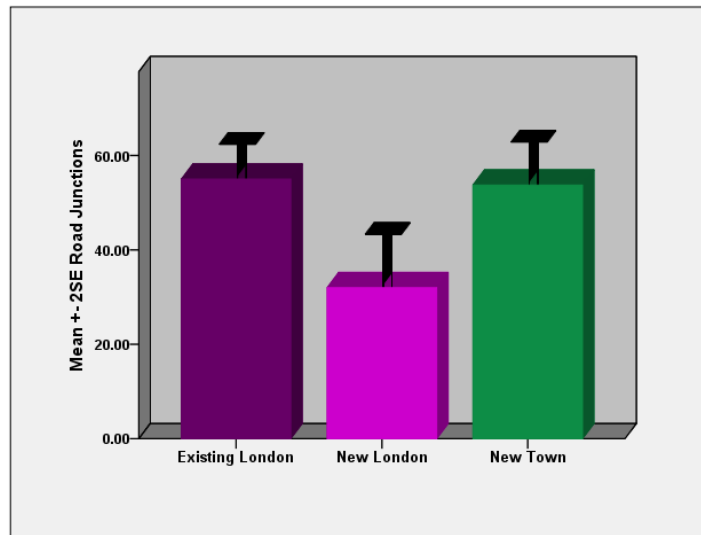


Figure 34. Number of road junctions recalled across environments

Experimenter ratings of sketch map orientation ($t(19)=8.54$; $p=0.001$; $d=2.28$) and overall sketch map categorisation ($t(19)=6.96$; $p=0.001$; $d=1.65$) were significantly better for New Town compared to new London.

Table 9 - Performance of taxi drivers on tests of existing and new London (scored separately), and New Town.*

	Existing London Mean (SD)	New London Mean (SD)	New Town Mean (SD)
Scene recognition– targets (/14) ^{EL}	88.57(9.09)	68.00(15.07)	61.04(12.62)
Sketch map number of road segments (/14) ^{EL/NT}	68.92(21.02)	34.09(21.24)	56.56(20.32)
Sketch map number of road junctions (/13) ^{EL/NT}	55.00(15.64)	32.00(24.40)	53.75(19.49)
Sketch map number of landmarks (/9) ^{NT}	37.22(24.25)	28.50(16.31)	48.00(18.53)
Sketch map landmark placement (/27)	31.66(24.72)	22.83(17.71)	36.86(18.96)

*Mean scores are presented as proportional scores. ^{EL}Significantly better performance for existing London. ^{NT}Significantly better performance for new Town.

Considering integration between the existing knowledge of London’s layout and the new environmental information, several other scores are relevant. There were three major integration points (road junctions) between existing and new London. On average, taxi drivers included 1 (SD (0.77)) on their sketch maps.

The London sketch maps were also scored by the two raters for integration on a scale of 1-5, where 1=little or no integration... 5=good integration. The average integration score was 2.6 (SD 1.14). In a debriefing session taxi drivers were also asked how they found doing the London task relative to New Town. The overwhelming response was that they found the London task more difficult than learning New Town. Some of the reasons given included: *“I couldn’t link the old with the new”*; *“It would have been easier if I had to learn just the new London bits. It was confusing to add them to the old London”*; *“Very hard to remember the new parts”*; *“It was confusing to put the new parts into London”*; *“I found it a lot harder than New Town. I found it hard to integrate the new parts into London”*; *“I couldn’t link the new parts into London properly”*.

4.2.5 Discussion

In this study, I examined whether being a very skilled navigator in one environment had any effect on learning the layout of an entirely new environment. I found that it did, with licensed London taxi drivers significantly better than control participants at executing routes through a new town, and representing this environment at an overall map-like survey level. However, the benefits of navigational expertise were not universal. Compared with their performance in the new town, taxi drivers were significantly worse when tested on their knowledge of the layout of a new area that had to be integrated into their existing representation of London. I discuss each of these results in turn.

Both control participants and taxi drivers learned New Town at the same rate, were comparable in their ability to recognise scenes from New Town from among similar-looking distractors, and were able to make proximity judgements

between landmarks to an equivalent level. Thus, the significantly better performance by taxi drivers in executing routes through New Town and across all the sketch map variables was not attributable to differences in basic topographical knowledge. Rather the advantage for taxi drivers was in being able to plan and execute routes, and in possessing a survey-like representation of the town as exemplified in sketch maps. Examination of the representative sketch map in Figure 31A, illustrates that taxi drivers appreciated how the two routes overlapped, and were able to integrate them into a reasonably coherent spatial map of the area. By contrast, the control participants, whilst acquiring knowledge of the two routes separately, were generally less able to make the discrete representations cohere (Figure 31B).

Overall, this finding suggests that navigational expertise in a specific environment is not only a matter of accruing a large amount of information about the layout and content of that environment. The fact that experts are benefited when learning a new environment means that aspects of navigational expertise generalise. What might these aspects be? Taxi drivers undergo years of training where they have to pay close attention to multiple complex routes as well as salient landmarks, and learn how the routes across a large city relate to each other. Similarly in their job, day in day out, they are required to plan and execute routes. Clearly these general attentional, learning and memory mechanisms are finely-tuned and readily called upon when they are required to learn a new town. In the current task, and in line with their training and experience, it is likely that they paid more attention than control participants to the content of the two routes, when salient landmarks appeared and in what order, and in particular to how the routes fitted together.

As well as possessing well-honed cognitive strategies when actively learning a spatial layout, another factor should be considered. Licensed London taxi drivers have been shown to have a different pattern of grey matter compared with control participants in the hippocampus. They have more grey matter in the posterior part of the hippocampus and less grey matter in the anterior hippocampus, and this correlates with the amount of time spent taxi driving [147, 151, 291]. Although it is not known what underpins the greater grey matter volume in the posterior hippocampus of taxi drivers, it may be related to recruitment of new neurons following neurogenesis [303, 304] that are pressed into the service of spatial processing, or the development of greater communication between neurons in the form of increased synaptogenesis [305]. This pattern of hippocampal grey matter volume may mean that there is simply a larger neural capacity for assimilating new spatial layouts in taxi drivers compared with control participants. Whilst this may be the case, the findings from my second test suggest that this explanation may not be sufficient.

While taxi drivers are highly skilled at wayfinding around London, and were advantaged when learning a new environment, acquisition of novel spatial information in the context of an existing knowledge domain was significantly poorer in comparison. Thus the cognitive mechanisms that were at play in New Town were seemingly not in operation to the same degree in the new parts of London. For example, a number of the sketch map measures such as number of road segments, number of road junctions, overall orientation and overall map categorisation scores were significantly lower for new London suggesting the taxi drivers failed to integrate existing and new parts of London into a holistic representation. Further evidence for this lack of an overall survey-like representation was in the low ratings for integration between the two parts of London, the fact that taxi drivers on average only recalled one of the three major

junctions between existing and new London, and most clearly of all in the direct comments of the taxi drivers in the debriefing. They unanimously found the London task much more challenging than learning a new town from scratch. Why might this be?

There may be parallels between my findings and those of previous research in the realm of expertise more generally. While it has been shown that experts' performance generally surpasses that of novices in domains such as chess, bridge, sport, music, and physics [306-309], such experts can be prone to making mistakes and be more errorful than novices in some circumstances [310-313]. Luchins [314] was one of the first to demonstrate this experimentally by showing the Einstellung effect in a water jug experiment. Participants were trained to solve a series of problems using a fixed solution. When tested on new problems that were similar to the ones used during the training phase, the majority of participants failed because they applied the fixed solution instead of finding the most appropriate solution for a given trial. Luchins argued that experts' knowledge can make them unable to adapt to new task demands thereby missing the optimum solution to a problem. This effect has also been found within specific domains of expertise, where experts were reported to show increased accuracy, speed and capacity compared to novices, but they made more errors when presented with new problems that closely resembled their knowledge base [315-317].

Ericsson [2] suggested that after a period of training, performance becomes automated and experts may lose conscious control over the execution of skills, making intentional modifications difficult, resulting in errors within the domain in which they have their expertise. This suggests that performance is negatively affected when there is increased similarity between old and new information, or

where there are overlapping demands between old and new stimuli [315]. Memory in aged adults has also been shown to be negatively affected when there is greater similarity between contexts and objects [318, 319]. I suggest that the Einstellung effect in experts can be regarded as a form of memory interference. Thus, when operating in a new environment which is distinct from the environment where they have their expertise, the finely-honed strategies (and possibly the increased posterior hippocampal grey matter volume) of London taxi drivers offers them a distinct advantage. By contrast, when presented with new information to learn that is similar to their existing knowledge, their poorer performance may reflect expert inflexibility and an inability to inhibit access to existing (and now competing) memory representations.

I speculate that this latter effect could be exacerbated by the decreased volume of the anterior hippocampus in taxi drivers, perhaps limiting the ability to integrate the old representations of London and the new. Alternatively, it could relate to neuronal 'rigidity' akin to that observed in place cells of aging rats. These cells maintain their original place fields despite changes in the surrounding cues, especially when the changes are subtle [320]. Although the taxi drivers in this study were healthy and relatively young, perhaps the high demand on spatial memory and navigation leads to an adaptive reorganization of the associational circuits of the hippocampus that produce changes to the way in which they process information that are comparable to those reported in aged animals.

When working in London, taxi drivers are required to incorporate new information into their representation of the city on a daily basis, such as new traffic flows and road works information. If they had difficulty with assimilating

this new information then we might expect a large proportion of taxi drivers to have problems carrying out their job. But this is not the case; taxi drivers in London usually stay working successfully in the profession for decades. It is likely, however, that the changes to the environment that they encounter are incorporated gradually, with difficulties overcome by repeatedly travelling along the modified routes. In our task, they had only four exposures in which to acquire the novel information. Interestingly, in the debriefing session, a number of taxi drivers noted the parallel between the difficulties they had with our new London task, and a situation that arose a number of years ago with the development of a new area in London. When the Canary Wharf district was opened, taxi drivers had great difficulty getting to know the area, and said that it took a long time before they were able to navigate successfully there. It may be that our sudden introduction of radical changes caused the taxi drivers to stop using their normal survey representation that allows for flexible wayfinding, and instead they reverted back to route-based strategies relying on rigid route representations [321-324]. Strategy choice usually depends on the information available to the navigator, the reliability of this information [364, 365]. Nevertheless, individuals that are best at navigation appear to be able to alternate between strategies, depending on what is optimal for that specific situation [366]. Thus, in this instance it may have been overburdened by the amount of change and the mismatch between expected, well-known sequences of intersections and the new experiences.

Previous studies of London taxi drivers documented their poor performance on table-top tests of visuo-spatial memory. Specifically, taxi drivers recalled fewer details of a complex figure (Rey-Osterrieth complex figure; [271]) after a delay compared with control participants [151, 291]. Similarly, they took longer to learn associations between sixteen objects and locations on a table-top array, and

also had poorer recall of the object-location pairs after a delay [291]. It is not immediately clear why deficits should be seen on these tests which involved exclusively new material, and in the current London task which involved the integration of new with existing knowledge in large-scale space. In Chapter 7 I will consider this issue in more detail.

In conclusion, the main aim of this study was to assess whether navigational expertise in one urban environment had any effect on learning a new environment, and when incorporating novel information into an extant spatial representation. I have demonstrated that expertise is coupled with an advantage over novices for large-scale spatial layouts that are distinct from existing knowledge, whilst at the same time placing limits on experts' performance within their specific domain of expertise.

Chapter 5

Non-spatial expertise and hippocampal grey matter volume

Précis

In the previous three experiments, my focus was on taxi drivers, whose expertise is clearly within the spatial domain. An obvious and important question is whether non-spatial expertise is associated with similar hippocampal grey matter volume patterns. I investigated this in Experiment 4.

5.1 Introduction

While grey matter volume differences have been described relative to control subjects for numerous groups with expertise, such as musicians, bilinguals, and mathematicians (e.g. [42, 48, 49] ; see also Chapter 1), licensed London taxi drivers are distinct from these experts in several ways. Taxi drivers acquire their knowledge in adulthood, unlike other groups who often begin to acquire or practise their skills in childhood. There is also an intensive period of knowledge acquisition over several years, and then continued use and development of that knowledge thereafter [151]. The knowledge that taxi drivers acquire is spatial, with knowledge elements highly interconnected, giving rise to one large and complex mental representation of London's layout. Any of these factors could contribute to the hippocampal grey matter volume patterns previously observed.

In this study I sought to examine one of these factors, namely whether the hippocampal grey matter effects noted in taxi drivers are specifically associated with intensive and extensive knowledge acquisition in the spatial domain, or whether knowledge that is less or non-spatial would be associated with similar patterns of hippocampal grey matter volume. Terrazas and McNaughton [325] questioned whether other groups such as medical doctors or lawyers, who also have high memory demands placed upon them, might also show similar

hippocampal grey matter volume patterns as London taxi drivers. Their question is reasonable given the intimate link between memory and the hippocampus [73], and long-standing debates about whether or not the hippocampus is preferentially engaged by spatial stimuli [97, 102, 106, 326, 327].

In order to examine this issue, I needed to identify a group of participants who are similar to taxi drivers in terms of how their knowledge acquisition occurs, but whose knowledge is less spatial. As suggested by Terrazas and McNaughton [325], we focussed on medical doctors as an appropriate group to test. They also acquire their knowledge in adulthood over a number of years of initial intensive medical training, and they then deploy and develop it subsequently. Grey matter volume has been examined previously in medical students where knowledge acquisition occurred over a three month period [53]. However this short time scale is not comparable with that of taxi drivers, and thus our interest was in testing medical doctors whose experience ranged over much longer timescales. In this context, the relationship between medical knowledge acquisition and grey matter volume is unknown.

In group comparison studies it is important that groups are matched on key variables such as gender, age and IQ. IQ scores, for example, have been reported to correlate with hippocampal volume in healthy adults [328]. While previous work has shown taxi drivers and their control subjects have average IQs [151]; see also Chapters 3 and 4), our expectation was that medical doctors would have above average IQs. Thus, while ideally one would like to directly compare medical doctors and taxi drivers, it would be difficult to match the groups for age, IQ, and socio-economic background. Instead we sought control subjects matched to the doctors on these variables, but who had not undergone university education or periods of intense learning. In this way we ensured that

differences between the medical doctors and their control subjects were not due to disparity on basic factors such as intellectual ability.

5.2 Experiment 4

5.2.1 Material and methods

Participants

Forty six healthy male volunteers participated in the current study. Of these 23 were medical students/qualified doctors, and 23 were control subjects. The background details of the two groups are shown in Table 10. The medical doctors varied in their amount of medical experience ranging from 0.5 to 22.5 years, with a mean of 12.07 years (SD 5.43). Interestingly, this is comparable to the navigation experience of the licensed London taxi drivers I tested in Chapters 3 and 4. Of those who had progressed beyond general medical training and into specialization, we excluded radiologists and surgeons, given the increased potential for their reliance on spatial processing.

The control subjects had not received formal education after the age 18 and none had studied at university. We also excluded from the control group anyone who had engaged in substantial non-university courses, or who had hobbies that involved developing a high level of expertise on a particular topic. Control participants were recruited from our Centre's volunteer database, from local government offices, and local banks. In addition, I solicited the assistance of the British branch of 'Mensa', a society for people with high IQs. Careful screening of volunteers from Mensa yielded a number of participants who met the criteria for inclusion. Medical doctors and control subjects did not differ in terms of age ($t(44)=1.31$; $p=0.19$) or handedness ($t(44)=1.57$; $p=0.12$). All participants were highly proficient in English. Given that both groups comprised participants with

English as their second language (10/23 medical doctors; 5/23 control subjects) it was not appropriate to use vocabulary-based tests to estimate IQ. Instead, like Maguire et al. [151] we assessed visual information processing and abstract reasoning skills using the Matrix Reasoning sub-test of the Wechsler Abbreviated Scale of Intelligence [259]. The mean scaled score for both groups was, as expected, above average, and the groups did not differ significantly ($t(44)=-1.00$; $p=0.31$).

Table 10 - Participant characteristics

Group	Medical Doctors Mean (SD) (n=23)	Controls Mean (SD) (n=23)
Mean age (years)	31.48 (5.42)	34.13 (8.01)
Handedness - laterality index	70.90 (34.25)	83.43 (6.76)
Mean scaled score on Matrix Reasoning	14.04 (1.39)	13.65 (1.22)
Mean medical experience (years)	12.07 (5.43)	

Structural MRI brain scan

A whole brain structural MRI scan was acquired for each participant as described in Chapter 2, using the MDEFT sequence without fat suppression.

Neuropsychological tests

In the first instance the medical doctors made some ratings (1-10, low-high) and answered a range of questions about their medical knowledge. These ratings were employed to ascertain the degree of connectedness of medical knowledge elements, whether this knowledge is mainly verbal or visual and when the amount of medical knowledge was greatest.

A battery of standardised neuropsychological tests was also administered to participants primarily examining memory, given my particular interest in the hippocampus. The tests are listed on Table 11, and detailed in Section 3.2.1 of Chapter 3. In addition, a range of anxiety and stress measures were taken – see Table 12 – these are also described in Section 3.2.1 of Chapter 3.

5.2.2 Procedure

Each subject was tested individually during one session of approximately 1.5 hours. The order of neuropsychological testing and MRI scanning was random across subjects.

5.2.3 Data analyses

MRI scans were analysed using VBM implemented in SPM5 as described in Chapter 2, with a smoothing kernel of 10mm [12]. The significance level for the hippocampus was set at $p < 0.001$ uncorrected for multiple comparisons given our a priori interest in this region. The significance level for the rest of the brain was set at $p < 0.05$ corrected for multiple comparisons.

For the behavioural data, basic group comparisons relating to participant characteristics were made using two-tailed t-tests. For the main analyses, data were analysed using MANOVA (see Chapter 2), with group (medics, controls) as the independent variable, and the neuropsychological/stress measures as the dependent variables.

5.2.4 Results

My goal in testing medical doctors was to assess a group of subjects whose acquired knowledge was less visual/spatial than that of taxi drivers. Responses (see Figure 35) confirmed that a low percentage of doctors regarded their

knowledge as purely visual. Instead most thought their knowledge was verbal or a mix of verbal and visual. Moreover, when asked to rate how much they used imagery when retrieving medical information from memory, the average rating was quite low at 5.56 out of 10 (SD 2.44). Other ratings showed that, aside from this difference in domain, there were reassuring parallels between medical doctors and those of taxi drivers. Doctors rated the extent to which their career in medicine involved acquiring a lot of knowledge as high (mean rating 8.87 out of a possible 10; SD 1.12). In addition, nearly three quarters of doctors described their knowledge as comprising elements that were well or highly connected. As with taxi drivers (e.g. see [151]) the ratings showed that a high proportion of doctors perceived their knowledge as continuing to increase well beyond their primary medical degree. These results support my expectation that medical doctors were an appropriate group in which to examine the relationship between hippocampal grey matter volume and expertise that is similar but less visual/spatial than that of London taxi drivers.

Neuropsychological data

Next I compared medical doctors and control participants in terms of their neuropsychological profiles. A MANOVA was performed with group (medical doctors, control subjects) as the independent variable. Scores from seven cognitive measures served as the dependent variables, namely: digit span, word recognition and face recognition, logical memory immediate recall, logical memory delayed recall (see Table 11). No significant differences were found between the groups ($F(7,38)=1.64$; $p=0.16$).

Table 11 - Memory measures

Memory measures	Medical Doctors Mean (SD) (n=23)	Controls Mean (SD) (n=23)
Digit Span – scaled score	12.96 (3.23)	14.30 (4.06)
WRMT – Faces scaled score	12.35 (2.36)	11.09 (2.79)
WRMT – Words scaled score	12.83 (0.38)	12.83 (1.15)
Logical memory (WMS) – IR	86.87 (13.31)	87.57(15.54)
Logical memory (WMS) – DR	81.04 (15.58)	87.00 (15.85)
Rey complex figure copy (/36)	35.52 (1.37)	35.43 (1.08)
Rey complex figure – DR (/36)	24.54 (4.92)	23.02 (5.09)

WRMT=Warrington Recognition Memory Test; IR=immediate recall; DR=delayed recall after 30 minutes). Standard deviations are shown in parenthesis

An additional MANOVA was performed using five anxiety/stress measures: Perceived Stress Scale, State Trait Anxiety Inventory (state); State Trait Anxiety Inventory (trait); life stress rating and job stress rating (see Table 12). No significant differences were apparent between the groups ($F(5,38)0.34$; $p=0.89$).

Table 12 - Stress measures

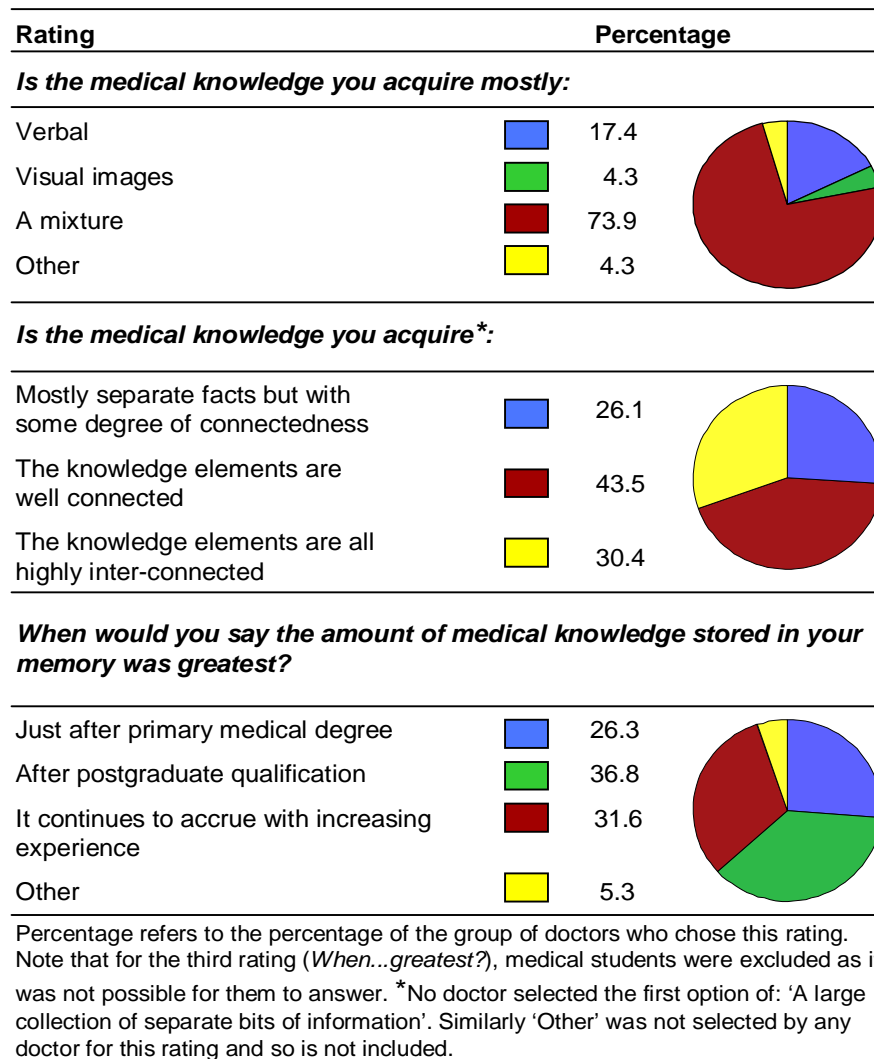
Group	Medical doctors Mean (SD)	Controls Mean (SD)
Perceived Stress Scale	12.91 (6.09)	13.39 (7.05)
State Trait Anxiety Inventory -1	31.43 (7.86)	29.96 (8.86)
State Trait Anxiety Inventory -2	37.74 (9.13)	36.39 (11.66)
Life stress	5.04 (2.05)	4.96 (2.09)
Job stress	5.00 (1.93)	4.65 (2.01)

MRI data

Using VBM I examined whether there were any grey matter volume differences between the doctors and control participants anywhere in the brain. No significant differences in grey matter volume were found. To limit the chances of false-negatives, a number of other analyses were undertaken. The standard VBM analysis was repeated but this time with significance levels at $p<0.005$

uncorrected for multiple comparisons. Even at this liberal threshold no significant difference in hippocampal grey matter volume was observed between the groups. While a smoothing kernel of ~10mm is normally employed in VBM, it might be suggested that a smaller smoothing kernel is more appropriate for the hippocampus, such as the 4mm used by Maguire et al. [147] and Gadian et al. [329]. Thus images were smoothed at both 8mm and at 4mm and the analyses described above were repeated, again with liberal statistical thresholds, but no significant differences were found between the two groups.

Figure 35 – Knowledge in medical doctors



I also examined whether any of the neuropsychological or stress scores correlated with grey matter volume anywhere in the brain. I found no significant correlations for either group. Finally, I assessed whether grey matter volume anywhere in the brain correlated with the number of years of medical experience in the doctors. No significant correlations were found. In particular, no correlations between years medical experience and hippocampal grey matter volume were apparent even at the liberal threshold of $p < 0.005$ uncorrected for multiple comparisons.

5.2.5 Discussion

In this study I used medical doctors as a model of knowledge acquisition and use that is less spatial than that of taxi drivers. I made strenuous efforts to match the control participants to the doctors on their background characteristics, except for the crucial experimental manipulation of interest, namely university medical training. Even when liberal statistical thresholds were employed, I failed to find any evidence for grey matter volume differences between the two groups anywhere in the brain, including the hippocampus. Moreover, years medical experience did not correlate with grey matter volume in the hippocampus or elsewhere in the brain.

Naturally caution is required in interpreting the absence of positive findings. I acknowledge that the absence of positive findings is not evidence for the absence of an effect. However, by the same token, one cannot rule out the possibility that there is no effect to be found, particularly given that previous studies using the same techniques and protocols, and involving similar subject numbers have identified differences in the spatial domain. I therefore consider some implications of the latter possibility as suggested by the current findings.

The main difference between the two groups was that one group of subjects acquired a large amount of knowledge over many years, whilst the other group did not. Despite this difference, the groups were comparable in terms of memory performance, and grey matter volume in the hippocampi and indeed elsewhere in the brain. Furthermore, years medical experience did not correlate with grey matter volume. The question, therefore, is why does spatial expertise and years navigating in London by taxi drivers correlate with grey matter volume ([147, 151]; see also Chapter 3), while medical expertise and experience does not. It has been suggested that the hippocampus is particularly specialised for supporting allocentric spatial representations of the type that permits flexible navigation in large-scale space [87, 330]. In non-humans the volume of the hippocampus is known to vary as a function of the demands placed on spatial memory (see [117]; see Chapter 1). The lack of positive findings in medical doctors lends further support to the view that the hippocampus may be preferentially engaged by spatial memory [97, 102, 106, 326, 327].

Whilst the less spatial nature of their knowledge is an obvious reason for the lack of positive findings in medical doctors, there may be another pertinent factor to consider. The knowledge elements for taxi drivers (e.g. roads, landmarks, road junctions, etc) are multiply connected with each other, and the resultant overall mental representation is a large, highly integrated coherent whole. Whilst medical doctors also acquire vast amounts of knowledge elements, and rated them as well or highly connected, it may be they are better characterised as sets of associations, with fewer multiple connections. It is also arguable whether their medical knowledge is represented as a single coherent entity, perhaps being better characterised as a reservoir of multiple sets of knowledge elements. One might argue, therefore, that one should test another group of subjects whose knowledge is non-spatial like medical doctors, but that

is more like taxi drivers in its nature. However, it is difficult to think of another instance of expertise, besides London taxi drivers, where knowledge is so vast and multiply connected, and forms one coherent single representation.

Notwithstanding this issue, the current findings allow me to conclude that intensively acquiring a large amount of knowledge over many years is not invariably associated with hippocampal grey matter volume differences relative to matched control participants. Instead it would seem that hippocampal volume effects are more likely to be observed when the knowledge acquired concerns a complex and detailed large-scale spatial layout.

Chapter 6

Hippocampal plasticity: a longitudinal study

Précis

In the four experiments so far, I have explored two key issues, namely what are the effects of navigational expertise on memory and the hippocampus, and are the effects of intensive knowledge acquisition, spatial or non-spatial, invariably associated with a particular pattern of hippocampal grey matter volume. However, all of these studies were cross-sectional, and even though the hippocampal grey matter volumes correlated with time taxi driving (in Experiment 3; see refs. [147, 151]) suggesting hippocampal structural plasticity was associated with the development and use of a complex spatial representation, the evidence is still somewhat indirect. To definitively address the issue of hippocampal structural plasticity in the face of demands on spatial memory, a longitudinal study is required - Experiment 5 constitutes such a study. Here I compared trainee London taxi drivers before and after they completed their Knowledge training in order to assess, within subjects, the effect of their training on hippocampal grey matter volume and memory.

To examine plasticity fully, I reasoned that it would also be important to consider not only the acquisition and use of the Knowledge, but also the effects of ceasing to use the Knowledge. Thus in Experiment 6, I conducted a preliminary investigation of retired London taxi drivers to ascertain whether there was evidence for 'reverse' plasticity, that is, do memory and the pattern of hippocampal grey matter volume associated with taxi drivers 'normalise' in this context?

6.1 Introduction

As outlined in Chapter 1, volumetric analyses of the hippocampal region in birds and other small mammals show that the size of the hippocampus varies depending on whether or not the animal engages in tasks that place a demand on spatial memory. For example, animals that store food and have to remember the location of their caches have larger hippocampi than their non-food storing counterparts (see [117] for a review; [325]). This volume difference is present even between animals of the same species, depending on the degree to which an animal exhibits storing behaviour [118]. Even more striking are the hippocampal volume differences evident in some species specifically during seasons where the demand for spatial memory is greatest i.e. the memory load during food-storing periods coincided with an increase in hippocampal volume [119, 120, 326]. Neurogenesis in the hippocampus of non-humans (see Chapter 1) may be related to the volume changes observed.

The obvious question is whether the volume of the hippocampus in humans is also susceptible to change in response to the memory demands placed upon it. Damage to the hippocampus, and the consequent volume reduction, is known to adversely affect memory performance [128]. However, the relationship between hippocampal volume and memory in healthy individuals is more uncertain ([280]; see Chapter 1). Perhaps one of the most consistent findings has been of increased posterior and decreased anterior hippocampal grey matter volume in licensed London taxi drivers relative to control participants [147, 151]. Moreover, correlations were found in taxi drivers between grey matter volume and time spent navigating, with posterior hippocampal volume increasing and anterior volume decreasing the longer spent navigating around London ([147, 151] – see Figure 9). While findings such as these suggest that hippocampal structural

plasticity is associated with the development and use of a complex spatial representation, the evidence is still somewhat indirect. What is required is a longitudinal study that investigates, within subjects, whether intensive spatial navigation and the formation and use of a complex spatial representation induce structural changes in the healthy adult human hippocampus.

Whilst longitudinal studies of grey matter atrophy, including in the hippocampus, in the context of progressive pathologies are well documented, longitudinal assessments of grey matter volume in the healthy brain are much fewer. In one study, Hyde et al. [56] examined the effect of musical training on the grey matter volume of children's brains. They found increased volume in motor and auditory areas in those children who received training. Whilst this is undoubtedly interesting, the fact that the training took place in children, whose brains are still plastic and developing, means it is impossible to know whether training an adult brain would have similar effects on grey matter volume.

In three other studies, this time in adults, participants were trained for a number of weeks on a visuo-motor task – juggling [52, 55, 327], and in both instances grey matter volume increased in the training group in parietal regions. Again, this is suggestive of plasticity in the adult cortex, but the low-level motoric nature of the training does not speak to whether higher cognitive functions such as memory could affect grey matter volume. Draganski et al. [53] studied young medical students over several months whilst they revised for an examination and found grey matter volume increases in parietal and medial temporal regions associated with revising. This does seem to suggest that intensive use of higher cognitive functions may affect grey matter volume in adults, although as in the previous motor studies, the timescales involved were relatively short. It is also notable that in Experiment 4, when I examined medical doctors whose

knowledge was acquired and used over much longer timescales, and compared them with IQ-matched controls, I found no differences in grey matter volume anywhere in the brain. Thus, the effects observed by Draganski et al. [53] may only be short-term and associated with the initial phase of learning. To the best of my knowledge, there are no published studies that have examined the impact of a higher cognitive function such as learning and memory, on the grey matter volume of adults' brains over a timescale of years, despite the obvious importance of knowing whether grey matter volume increases (or indeed decreases) persist in the long term.

Finally, as far as I am aware, no one has examined the effect on grey matter in the adult brain of ceasing a long-term activity that was previously associated with changes in grey matter volume. Again, a better understanding of this could provide clues about the extent of structural brain plasticity, and inform issues such as 'use it or lose it'.

I therefore set out in Experiment 5 to examine trainee London taxi drivers as they embarked upon acquiring the Knowledge – learning over a number of years the layout of ~25,000 streets and the locations of thousands of places of interest – all tested and monitored by the Public Carriage Office (PCO). Participants were recruited in an unselected way as they arrived at the PCO; I then scanned and tested them at this point (Time 1), maintained contact with them over the subsequent 3-4 years, and then scanned and tested them again once qualified (Time 2). Thus, by taking adults with average IQs, who were engaged in intense spatial learning in a naturalistic setting, I could examine the effects, within subjects, on hippocampal (and indeed whole brain) grey matter and memory performance. The PCO advised that approximately 50-60% of trainees fail to qualify. If this failure rate was maintained in my sample, then I anticipated

having three subject groups, trainees who qualified, those who trained but failed to qualify, and the control participants.

There were a number of possible outcomes that I could foresee. Perhaps there would be no evidence of hippocampal structural plasticity associated with acquiring the Knowledge. However, this outcome is not necessarily easy to interpret in light of previous findings. It could mean that despite the cross-sectional findings, in particular the correlations between grey matter and years navigation, that in fact there is no evidence of structural plasticity in the human hippocampus in this context. However, a negative result could also mean that the timescales of structural plasticity are longer than the 3-4 years. Indeed the graphs of the navigation-hippocampal grey matter relationships in previous studies (e.g. see [151]) suggest that the longer the time post-qualification, the larger the grey matter volume in posterior hippocampus and the smaller the anterior volume. It could therefore be the case that testing immediately post-qualification is not the optimal time to appreciate hippocampal structural plasticity, and in fact re-testing should be at least several years after qualification. Another possible outcome could be that those who succeeded in qualifying were in fact already different to those who failed to qualify, even at Time 1. Perhaps those with larger posterior hippocampal grey matter volumes are predisposed to succeed on this challenging spatial learning task. My design would allow me to examine this possibility directly.

It is also possible that I would observe hippocampal grey matter volume changes within subjects as a consequence of the acquisition of the Knowledge. This would imply that, whilst volume changes continue post-qualification due to the continued use and development of the mental map of the city, there is also a step increase in posterior hippocampal grey matter volume associated with the

initial learning and formation of the spatial presentation of London. This was my favoured hypothesis, and would provide compelling evidence for plasticity in the adult human brain related to the use of a higher cognitive function. Moreover, if spatial memory can induce structural changes in the hippocampus, it may indicate a fundamental link between space and the human hippocampus [97]. Finally, a finding of hippocampal plasticity in response to environmental stimulation could allow us to start thinking about the boundaries within which human memory operates, and the scope for improving or rehabilitating memory in educational and clinical contexts.

To examine plasticity fully, I reasoned that it would also be important to consider not only the acquisition and use of the Knowledge, but also the effects of ceasing to use the Knowledge. As alluded to above and in Chapter 1, hippocampal volume can increase and also decrease in non-humans depending on the demands placed on spatial memory [119, 120, 122, 123, 331]. The question I therefore asked was what happens to the hippocampal grey matter volume patterns and memory performance of taxi drivers once they retire. Perhaps nothing changes, and once the hippocampal grey matter volume pattern is established during training and working as a taxi driver, it is unchanging even after retirement. Alternatively, and given the data from non-humans, perhaps once they cease to use their navigational expertise, the pattern of grey matter volume 'normalises', that is the posterior hippocampal volume decreases and the anterior increases to be more like control participants. Similarly, perhaps performance on tests such as the delayed recall of the Rey complex figure, shown repeatedly to be impaired in taxi drivers relative to controls, improves once taxi driving ceases.

6.2 Experiment 5

6.2.1 Materials and methods

Participants

In Section 2.1 of Chapter 2 I described how the participants for this longitudinal experiment were recruited. In the first phase (Time 1) 110 healthy male volunteers participated, of which 79 were trainee London taxi drivers, and 31 were control participants. The background details of the two groups are shown on Table 13. A MANOVA was performed to examine whether trainee taxi drivers and control participants were matched for the participant characteristics included in Table 13. This confirmed the two groups were not significantly different in terms of age, number of years in education, handedness, and visual information processing and abstract reasoning skills as measured by the Matrix Reasoning sub-test of the Wechsler Abbreviated Scale of Intelligence [259] ($F(4, 105)=2.30$; $p=0.063$). An estimate of verbal IQ was obtained using the Wechsler Test of Adult Reading [258]. Data for five trainee taxi drivers and two control participants were not obtained – although they were very proficient in English, it was not their first language, a requirement of the test. The two groups were very similar and in the average range for estimated verbal IQ, although the control group was significantly better than the trainee taxi drivers ($t(101)=2.86$; $p=0.005$) when formally tested with a t-test.

MANOVA is considered to be robust when comparing groups that are unequal in size. Nevertheless, any adjustments made to account for differences in group size will diminish the statistical power of the test. Therefore the participant characteristics of a group of 31 randomly selected trainee taxi drivers were compared to the control participants to rule out Type I and Type II errors. I did this several times, with different groups of 31 trainees chosen at random. No

significant differences were noted at any point between the trainees and control groups across the background characteristics (e.g. no differences in terms of age, number of years in education, handedness, and matrix reasoning ($F(4,57)=1.58$; $p=0.19$), or on the Wechsler Test of Adult Reading ($t(55)=1.90$; $p=0.062$).

Table 13 - Time 1 Volunteer characteristics

Memory measures	Trainee taxi drivers Mean (SD) (n=79)	Controls Mean (SD) (n=31)
Mean age (years)	38.22 (7.71)	35.00 (8.99)
Education (age left school, years)	16.60 (1.22)	16.77 (1.30)
Estimated verbal IQ (WTAR)*	97.23 (6.21)	100.79 (3.79)
Mean scaled score on Matrix Reasoning	11.72 (2.15)	11.83 (2.39)
Handedness - laterality index	85.15 (31.49)	71.74 (38.69)

WTAR= Wechsler Test of Adult Reading; Matrix reasoning= this sub-test of the Wechsler Abbreviated Scale of Intelligence. Standard deviations in parentheses. * Not all subjects were included – see text.

In the second phase of this study (Time 2) 90 of the original 110 participants returned. Of these, 39 were now fully qualified licensed London taxi drivers (QTD), 20 were non-qualified trainee taxi drivers (NQTD) - they had ceased training or had failed their appearances, and 31 were control participants (C). Of the 20 trainees that did not return during at Time 2, two had moved away, two qualified but did not wish to complete the study, the remaining 16 had decided to stop training/made no appearances but were unable to complete the study due to family and work commitments. A MANOVA confirmed the three groups were not significantly different in terms of age, number of years in education, handedness, and matrix reasoning ($F(8,166)=1.81$; $p=0.07$). An ANOVA employed to identify any differences between the groups on the Wechsler Test of Adult Reading just reached significance ($F(2,80)=3.08$; $p=0.051$). The source of this difference was investigated using the post-hoc tests produced by ANOVA

(Bonferroni corrected) and showed that the controls performed better than the QTD group ($p=0.048$), although all groups performed within the average range and with similar group means. This shows that prior to training, there were no substantial differences in basic background characteristics between trainees who went on to qualify and those trainees who failed to qualify.

Of note, my data show that 49% of trainees qualified, which is in line with the estimates of the Public Carriage Office. The 39 qualified taxi drivers on average took 38.84 (SD 7.02) months to qualify and 15.64 (3.66) appearances.

Table 14 - Time 2 Volunteer characteristics

	Qualified trainee taxi drivers Mean (SD) (n=39)	Non-qualified trainee taxi drivers Mean (SD) (n=20)	Controls Mean (SD) (n=31)
Memory measures			
Mean age (years)	37.97 (7.96)	40.50 (5.27)	35.00 (8.99)
Education (age left school, years)	16.66 (1.32)	16.75 (1.40)	16.77 (1.30)
Estimated verbal IQ (WTAR)*	97.72 (6.29)	98.66 (3.49)	100.79 (3.79)
Mean scaled score on Matrix Reasoning	11.89 (2.06)	12.20 (1.98)	11.83 (2.39)
Handedness - laterality index	87.97 (24.09)	88.10 (21.50)	71.74 (38.69)

WTAR= Wechsler Test of Adult Reading; Matrix reasoning= this sub-test of the Wechsler Abbreviated Scale of Intelligence. Standard deviations in parentheses. Not all subjects were included – see text.

Test materials

A range of basic cognitive, memory and affective functions were assessed. As longitudinal studies involve repeat testing, practice effects may hamper interpretation when identical tests are administered, for this reason two test batteries were used here, one administered at Time 1, at the start of training, and the second at Time 2. Most of the tests have already been described in Chapter 3 – the digit span subtest from the WAIS III, the Warrington recognition

memory test for faces and for words, the Rey complex figure, and the London landmarks proximity judgements test. At Time 2, parallel versions of these tasks were included – the modified Taylor complex figure [277], a second version of the Warrington recognition memory test developed locally at the National Hospital for Neurology and Neurosurgery Queen Square, an alternative version of the digit span test taken from Lezak [265], and a second version of the London landmarks proximity judgements tests that I developed. The anxiety and stress measured used my previous experiments were also included in this study: the perceived stress scale, the State-Trait Anxiety Inventory (state, trait) and ratings of life stress and job stress. At Time 2 the Object-place associations test was also included. In addition to these tests, several others were included:

Adult Memory and Information processing Battery (AMIPB) Story Recall [328]:

The story recall subtest is very similar to that of the logical memory subtest of the Wechsler Memory Scale, and is designed to assess the free recall of verbal material at immediate and 30 minutes delayed recall. There are two forms of the AMIBP, and each has one story – for story 1 (used at Time 1) two points are awarded for each idea unit correctly recalled; 1 point is awarded for each partially recalled idea. The total number of idea units for the story is 28, thus the total possible score is 56. Story 2 (used at Time 2) has 30 idea units and a maximum score of 60.

Little Man Test [329]: This test assesses spatial orientation and was used at Time 1. Participants are shown cards of a silhouette figure 'Little Man' with a black disc in one of his hands. There are four positions the man can appear in each of which are shown eight times in random order. The task is to state

whether the black disc is in the figure's right or left hand. A score of 1 is awarded for correct responses, and the total possible score is 32.

Flags test [330]: This is also a test of spatial orientation and was used at Time 2. Each target item (flag) has a rectangular geometric design. In addition on each trial there are six flags that either mirror the target design, or differ in their spatial rotation. The participant has to indicate which of the six flags show the same design as the target. A score of 1 is awarded for each correct response, and the total possible score is 126.

Structural MRI brain scan

Each participant had two whole-brain structural MRI brain scans, one at Time 1 and another at Time 2, using the same scanner and the same scanning sequence (with fat suppression) as described in Chapter 2.

6.2.2 Procedure

Testing at both Time 1 and Time 2 took place in one session of approximately 2.5 hours. The order of the MRI scan and testing was random across subjects.

6.2.3 Data analyses

Behavioural data were analysed as outlined in Chapter 2, with the additional use of repeated measures MANOVAs when comparing the same subjects at different time points.

MRI images were analysed using standard VBM implemented in the Statistical Parametric Mapping software as described in Chapter 2, with a smoothing kernel of 8mm full width at half maximum. Analyses focussed on grey matter.

The effects of global grey matter volume and subject age were excluded by modelling them as confounding variables. Given our *a priori* interest in the hippocampus, the significance level was set at $p < 0.001$ corrected for the volume of the hippocampus using a sphere of 4mm. The significance level for the rest of the brain was set at $p < 0.05$ corrected for multiple comparisons across the whole brain. DARTEL and high dimensional warping (HDW) were also used.

6.2.4 Results: Time 1

MRI data

Two types of analyses were undertaken (see also Chapter 2). In the first instance standard SPM8-VBM was used to analyse the data. In a subsequent analysis DARTEL was used.

The trainee taxi drivers ($n=79$) and the control participants ($n=31$) were compared using SPM8-VBM. No differences in grey matter volume were found between the groups. To limit the chances of false-negatives, the above analysis was repeated with a much more liberal statistical threshold of $p < 0.005$ uncorrected for multiple comparisons. There were no significant differences in grey matter volume between the groups. Then the images were smoothed using a 4mm kernel and the analyses described above were repeated. Again, no significant differences were found between the groups. When DARTEL was employed, and the analyses repeated, no significant differences between the groups were apparent.

I next repeated the above analyses, this time considering the Time 1 scans for three groups: those trainees who ultimately qualified ($n=39$), those who did not ($n=20$), and the control participants ($n=31$). Standard VBM and the use of DARTEL failed to identify any differences between the groups either in the

hippocampus or elsewhere, at 4 or 8mm, even using a liberal statistical threshold.

These analyses clearly show that at Time 1, before training, the trainees and the controls did not differ. Most importantly, the trainees who went on to successfully qualify did not differ in terms of grey matter volume from those trainees who failed to qualify. This shows that qualification cannot be explained by a pre-existing pattern of hippocampal grey matter volume.

Neuropsychological data

Having established that there were no grey matter volumes differences, I next investigated if there were neuropsychological differences between the groups at the start of training. In the first instance I compared the trainees (n=79) and the control participants (n=31) on the 9 cognitive measures: digit span, Warrington recognition memory for faces, Warrington recognition memory for words, story recall at immediate and delayed recall, Rey complex figure copy and delayed recall, London landmarks proximity judgements, and Little Man test – see Table 15. There were no significant differences between the groups ($F(9,100)=1.46$; $p=0.17$).

I next repeated the above analyses, this time considering the Time 1 scores for three groups: those trainees who ultimately qualified (n=39), those who did not (n=20), and the control participants (n=31) - see Table 16. There were no significant differences between the groups ($F(18,156)=1.46$; $p=0.10$).

Table 15 - Time 1 Memory measures

Memory measures	Trainee taxi drivers Mean (SD) (n=79)	Controls Mean (SD) (n=31)
Digit Span – Scaled score	12.00 (3.13)	12.19 (2.52)
WRMT – Faces scaled score	11.60 (3.37)	10.29 (3.29)
WRMT – Words scaled score	12.45 (1.67)	12.38 (1.62)
AMIPB Story recall Form1 - IR (/56)	41.16 (6.96)	43.41 (8.56)
AMIPB story recall Form 1 - DR (/56)	39.20 (6.95)	40.90 (8.29)
Rey-Osterreith complex figure copy (/36)	35.75 (0.78)	35.87 (0.49)
Rey-Osterreith complex figure – DR (/36)	22.75 (5.51)	25.01 (5.98)
London landmark proximity judgments (/10)	7.65 (1.28)	7.38 (1.43)
Little Man (/32)	30.88 (2.40)	30.67 (2.22)

WRMT=Warrington Recognition Memory Test; IR=immediate recall; DR=delayed recall (after 30 minutes).

Next, I examined the five stress measures, first in the main trainee group (n=79) compared with the control participants – see Table 17. The two groups did not differ significantly ($F(5, 104) = 1.42; p=0.2$).

Table 16 - Time 1 Memory measures

Memory measures	Qualified trainee taxi drivers Mean (SD) (n=39)	Non-qualified trainee taxi drivers Mean (SD) (n=20)	Controls Mean (SD) (n=31)
Digit Span – Scaled score	12.15 (3.26)	12.80 (3.31)	12.19 (2.52)
WRMT – Faces scaled score	11.71 (3.75)	12.10 (2.88)	10.29 (3.29)
WRMT – Words scaled score	12.51 (1.53)	13.20 (1.36)	12.38 (1.62)
AMIPB Story recall Form1 – IR (/56)	40.74 (6.38)	40.95 (8.35)	43.41 (8.56)
AMIPB Story recall Form1 – DR (/56)	39.51 (6.35)	38.60 (8.38)	40.90 (8.29)
Rey-Osterreith complex figure copy (/36)	35.51 (1.07)	35.80 (0.61)	35.87 (0.49)
Rey-Osterreith complex figure – DR (/36)	22.39 (5.99)	23.35 (4.20)	25.01 (5.98)
London landmark proximity judgments (/10)	7.69 (1.34)	8.00 (1.12)	7.38 (1.43)
Little Man (/32)	30.76 (2.33)	31.75 (0.71)	30.67 (2.22)

WRMT=Warrington Recognition Memory Test; IR=immediate recall; DR=delayed recall (after 30 minutes); WMS-III=Wechsler Memory Scale.

Table 17 – Time 1 mean scores for the stress and anxiety tests

	Trainee London taxi drivers Mean (SD) (n=79)	Controls Mean (SD) (n=31)
Stress measures		
Perceived Stress Scale	13.77 (5.15)	13.87 (5.13)
State Trait Anxiety Inventory - T	29.55 (8.27)	30.16 (6.57)
State Trait Anxiety Inventory - Y	35.43 (7.24)	38.19 (9.84)
Life stress	4.92 (1.82)	4.51 (1.58)
Job stress	4.64 (2.10)	4.61 (2.33)

I next repeated the above analyses, this time considering the Time 1 stress measure scores for three groups: those trainees who ultimately qualified (n=39), those who did not (n=20), and the control participants (n=31) – see Table 18. There were no significant differences between the groups ($F(10,164)=0.77$; $p=0.64$).

Table 18 – Time 1 mean scores for the stress and anxiety tests

Group	Qualified trainee London taxi drivers Mean (SD) (n=39)	Non-qualified trainee London taxi drivers Mean (SD) (n=20)	Controls Mean (SD) (n=31)
Perceived Stress Scale	13.20 (5.23)	13.75 (4.60)	13.87 (5.13)
State Trait Anxiety Inventory - T	28.28 (6.03)	30.20 (11.30)	36.20 (7.05)
State Trait Anxiety Inventory - Y	35.10 (7.89)	30.16 (6.57)	38.19 (9.84)
Life stress	4.87 (1.70)	4.50 (1.93)	4.51 (1.58)
Job stress	4.71 (2.19)	4.05 (1.93)	4.61 (2.33)

These neuropsychological analyses clearly show that at Time 1, before training, the trainees and the controls did not differ. Most importantly, the trainees who went on to successfully qualify did not differ in terms of neuropsychological and memory profile from those trainees who failed to qualify. This shows that

qualification cannot be explained by a pre-existing pattern of cognitive performance.

6.2.5 Results: Time 2

The data acquired at Time 2 were analysed in the same way as Time 1 except that the analyses focused purely on three subject groups: those trainees who ultimately qualified (n=39), those who did not (n=20), and the control participants (n=31). In the first instance I compared the groups on a number of pertinent factors. I examined if there were any differences in total training time (months) or the average number of hours training, between the two trainee London taxi driver groups (Table 19). The two groups did not differ significantly in the length of time spent training. However the average number of hours training per week was significantly higher in the QTD group ($t(57)=6.62$ $p=0.001$). Unsurprisingly, the two taxi driver groups also differed significantly with regards to the average number of appearances, with the non-qualified group making hardly any ($t(57)=13.1$ $p=0.001$).

Table 19 – Training-related measures

Training related measures	Qualified Trainee taxi drivers Mean (SD) (N=39)	Non-qualified Trainee taxi drivers Mean (SD) (N=20)
Total time training (months)	38.84 (SD 7.02)	35.80 (SD 10.32)
Average hours training per week	34.56 (SD 12.40)	16.70 (SD 8.21)
Total number appearances	15.64 (3.66)	2.60 (3.45)

Next a one-way ANOVA was employed to ascertain if there were any differences between the groups in the time elapsed between scans (i.e. the time between Time 1 and Time 2) – see Table 20. There was no difference between the 3 groups ($F(2, 86) = 1.27$; $p= 0.28$).

Table 20 – Interval between testing

	Qualified Trainee taxi drivers Mean (SD) (n=39)	Non-qualified Trainee taxi drivers Mean (SD) (N=20)	Controls Mean (SD) (n=31)
Time between scans (months)	35.28 (8.19)	36.15 (8.53)	32.8 (7.37)

MRI data

As at Time 1, two types of analyses were undertaken (see also Chapter 2). In the first instance standard SPM8-VBM was used to analyse the data. In a subsequent analysis DARTEL was used.

No differences in grey matter volume were found between the groups using SPM8-VBM. To limit the chances of false-negatives, the above analysis was repeated with a much more liberal statistical threshold of $p < 0.005$ uncorrected for multiple comparisons. There were no significant differences in grey matter volume between the groups. Then the images were smoothed using a 4mm kernel and the analyses described above were repeated. Again, no significant differences were found between the groups. When DARTEL was employed, and the analyses repeated, no significant differences between the groups were apparent.

Neuropsychological data

Having established that at Time 2 there was a difference in hippocampal grey matter volume, I next investigated if there were neuropsychological differences between the groups post-training. I compared the groups on the 11 cognitive measures (see Table 21): London landmarks proximity judgements, Taylor complex figure copy and delayed recall, Warrington recognition memory for

faces, Warrington recognition memory for words, story recall at immediate and delayed recall, digits span, the flags test, object-place associations learning to criterion and object-place associations delayed recall (the latter had not been designed at the time of the start of this study, so could only be included at Time 2).

On this occasion there was an overall significant difference between the three groups ($F(22,152)=2.05$; $p=0.006$). This difference was investigated using the tests of between-subjects effects produced by MANOVA. There were three main effects.

Table 21 - Time 2 memory measures

Memory measures	Qualified taxi drivers Mean (SD) (n=39)	Non-qualified taxi drivers Mean (SD) (n=20)	Controls Mean (SD) (n=31)
Digit Span – Scaled score	11.12 (2.83)	11.10 (2.61)	12.29 (2.84)
WRMT – Faces scaled score	13.28 (3.16)	13.00 (2.82)	12.38(2.59)
WRMT – Words scaled score	12.43 (1.95)	13.70 (0.92)	12.41 (1.85)
AIMPB Form 2– IR (/60)	55.76 (30.66)	63.25 (33.05)	57.74(30.95)
AIMPB Form 2 – DR (/60)	55.74 (31.52)	59.75 (32.50)	59.83(32.41)
Taylor complex figure copy (/36)	35.92 (0.48)	35.75 (0.91)	35.93 (0.24)
Taylor complex figure – DR (/36)	21.93 (5.90)	22.60 (4.29)	25.77 (5.81)
London landmark proximity judgments (/10)	8.51 (0.94)	8.75 (1.25)	7.32 (1.97)
Flags (/126)	58.07 (23.53)	56.45 (20.04)	52.32(19.61)
Object-place learning to criterion	2.56 (1.35)	2.15 (1.18)	1.96 (0.91)
Object-place DR (/16)	12.79 (1.54)	13.75 (1.44)	13.74 (1.76)

WRMT=Warrington Recognition Memory Test; IR=immediate recall; DR=delayed recall (after 30 minutes).

London landmarks proximity judgements: A significant main effect was found between the groups on the London landmarks proximity judgements test ($F(2, 87)=8.09$ $p=0.001$). The pairwise post-hoc comparisons showed, qualified taxi

drivers were significantly better at judging relations between London landmarks than the control participants ($p=0.003$), and the non-qualified control taxi drivers were also significantly better than controls ($p=0.003$).

Taylor complex figure delayed recall: A significant main effect was found between the groups on the delayed recall of the Taylor complex figure ($F(2,87)=4.38$ $p=0.01$), where post-hoc comparisons showed qualified taxi drivers were significantly worse at recalling the complex figure after a delay than the control participants ($p=0.005$). The NQTD were also worse than controls ($p=0.05$).

Object-place associations test – delayed recall: A significant main effect was found between the groups on the delayed recall of the object-place associations test, ($F(2,87)=3.87$ $p=0.02$). The pairwise post-hoc comparisons showed qualified taxi drivers recalled fewer of the object-place associations after a delay than controls ($p=0.04$). Table 21 shows the mean performance for the two groups on this task.

Next, I examined the five stress measures – see Table 22. The three groups did not differ significantly ($F(10,164)=1.0$; $p=0.4$).

Table 22 – Time 2 stress and anxiety measures

Group	Qualified trainee London taxi drivers Mean (SD) (n=39)	Non-qualified trainee London taxi drivers Mean (SD) (n=20)	Controls Mean (SD) (n=31)
Perceived Stress Scale	12.28 (6.60)	13.20 (6.45)	12.48 (7.25)
State Trait Anxiety Inventory - T	27.02 (6.34)	25.30 (4.35)	27.64 (6.29)
State Trait Anxiety Inventory - Y	32.46 (8.75)	30.30 (6.42)	32.51 (8.63)
Life stress	4.61 (2.11)	5.55 (1.43)	4.83 (1.96)
Job stress	4.00 (2.93)	3.60 (2.83)	4.51 (2.04)

The first differences in neuropsychology have now emerged post-training. They bear a striking resemblance to the known differences documented in my earlier experiments and in Maguire et al. [151] namely that the now qualified trainees were significantly better on a London-based task, but worse on the delayed recall of two tests, the complex figure and the object-place associations test. These findings suggest that the acquisition of the Knowledge may be starting to affect memory. The real test, however, will be to compare the subjects with themselves, the direct comparison of Time 1 and Time 2.

6.2.6 Results: Time 1 vs. Time 2

MRI data

As explained in Chapter 2, standard VBM and DARTEL-VBM are used to examine grey matter volume differences in cross sectional studies, but should not be used for longitudinal assessment because of potential artefacts that could occur over time. Thus, in the comparison of Time 1 and Time 2 data in this experiment, high dimensional warping (HDW) was implemented SPM8. HDW safeguards against non-specific subtle differences that may arise between the first and second scans within subjects (for full details see [256]). The effects of the number of months between scans, total grey matter volume and participant age were excluded by modelling them as confounding variables.

Using SPM8 with HDW, I first compared the QTD group scans. Grey matter volume increased in the posterior hippocampi bilaterally (30, -42, 1, $z=5.44$, $p<0.02$; 20, -37, 10 $z=5.64$, $p=0.03$; 24, -39, 7, $z=4.02$, $p=0.04$; -29, -42, 3; $z=5.91$; $p<0.01$) at Time 2 relative to Time 1 (see Figure 36). The opposite comparison (T1>T2) did not show any significant differences. There were no other significant differences in grey matter anywhere else in the brain.

Similar analyses were performed for the NQTD group. No significant differences in grey matter volume were found anywhere in the brain, including the hippocampi, between Time 1 vs 2 and Time 2 vs Time 1. When the data from the two time points of the control participants were compared, again no significant differences in grey matter volume were found.

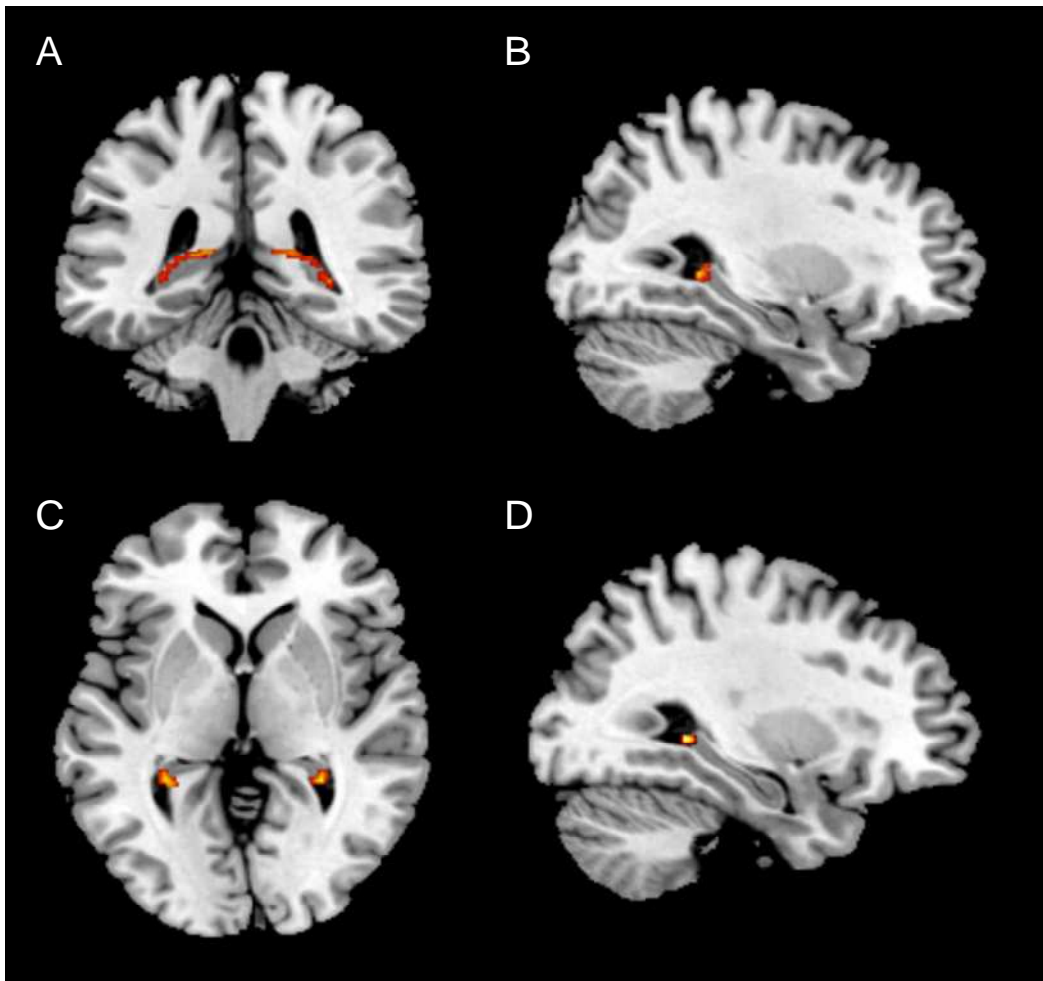


Figure 36. Grey matter differences between qualified London taxi drivers at start and end of training.

A: greater grey matter volume in the posterior hippocampi in qualified trainee taxi drivers at time of qualification compared with themselves as the start of training, shown on coronal section. B: montage shows the location of the peak voxel in the right hippocampus on shown on a sagittal section. C: greater grey matter volume in the posterior hippocampi in qualified trainee taxi drivers shown on an axial section. D: montage shows the location of the peak voxel in the left hippocampus on a coronal section. All images are shown on standard slices from the MNI template brain.

Thus using longitudinal SPM analysis techniques, the comparison of subjects with themselves before and after the successful acquisition of a detailed complex spatial representation of London, revealed structural changes in the posterior hippocampi. Failure to acquire the detailed spatial representation (NQTD) resulted in no such changes, and time *per se* – in the case of the control participants – also did not produce grey matter volume changes. The findings in the QTD group provide clear evidence for hippocampal structural plasticity in healthy adult humans.

Neuropsychological data

Having established these grey matter differences, I next examined whether there were any changes in neuropsychological test scores over time. A repeated measures MANOVA was performed on the cognitive measures from Time 1 and Time 2. Group (QTD, NQTD, controls) was the between factor variable, time (Time 1, Time 2) was the within factor variable, and the 8 cognitive measures served as the dependent variables, namely: digit span, story immediate and delayed Recall, Warrington face recognition memory, Warrington word recognition memory, complex figure delayed recall, London landmarks proximity judgements, mental rotation test. All scores were converted to Z-scores to make them comparable across time points. There was no main effect of group ($F(16,158)0.3$; $p=0.9$). There was no main effect of time ($F(8,80)0.1$; $p=0.9$). Nor was there a significant time x group interaction ($F(16,158)0.2$; $p=1.0$).

A repeated measures MANOVA was performed on the stress and anxiety measures from Time 1 and Time 2. Group (QTD, NQTD, controls) was the between factor variable, time (Time 1, Time 2) was the within factor variable, and the 5 stress measures served as the dependent variables, namely: Perceived Stress Scale, State-Trait Anxiety Inventory state and trait, job stress

rating, and life stress rating. There was no main effect of group ($F(10,164)=0.6$; $p=0.7$). There was a main effect of time ($F(5,83)=5.5$; $p=0.001$). There was no significant time x group interaction ($F(10,164)=1.06$; $p=0.3$). The source of the time difference was investigated using the post-hoc bonferroni corrected tests produced by MANOVA. The within subjects test showed an effect of time on the STAI-State ($F(1,87)=9.4$; $p=0.003$) and on the STAI-Trait ($F(1,87)=10.6$; $p=0.002$). To find out where the difference lay I compared the groups at Time 1 and Time 2 on these measures using paired t-tests. These showed that the control group reported higher scores on the STAI-State ($t(30)=2.2$; $p=0.029$) and on the STAI-Trait ($t(30)=2.0$; $p=0.048$) at Time 1 compared to Time 2. The QTD and NQTD groups showed no significant differences on these measures.

While there were no differences between the groups on the cognitive measures at Time 1, some differences emerged at Time 2. However, the direct comparison of the Time 1 and Time 2 tests scores within each group failed to show significant differences. This may be because the changes were in general subtle, given that these were healthy individuals.

6.2.7 Discussion

In this challenging study I recruited, tested, monitored over 3-4 years and re-tested a large group of healthy, average IQ'ed adult participants in order to examine whether intensive spatial learning resulted, within subjects, in changes to the structure of the hippocampus. I found that it did – in specifically those subjects who trained and successfully acquired a detailed spatial representation of London, This was associated with an increase in the volume of the posterior hippocampi. No such effects were seen for the group of trainees who did not acquire a mental map of London, despite training for a similar length of time.

Similarly, no grey matter volume changes were observed in control participants. Importantly the three groups did not differ in terms of grey matter volume at the outset of the study, nor in terms of background characteristics such as age and IQ. As far as I am aware, this is the first demonstration of hippocampal structural plasticity in healthy adult humans associated with higher cognition, i.e. memory, over a protracted timescale.

The grey matter volume changes I observed associated with acquiring the Knowledge were highly focal and restricted to the posterior hippocampi. No other changes were apparent anywhere else in the brain. The greater precision of the HDW permitted the within subjects effects to be observed in the context of SPM, underlining the usefulness of this technique for analysing longitudinal data. The grey matter volume increase associated with qualification was in the posterior hippocampus. This accords with the results of all previous cross-sectional studies involving licensed London taxi drivers, where their posterior hippocampus had greater grey matter volume relative to controls [147, 151].

In the cross-sectional studies, not only was posterior hippocampal grey matter volume increased relative to controls, but the anterior hippocampal volume was decreased. I found no evidence of anterior decreases in this longitudinal study. This is interesting because it may give a clue about the timeline and timescales of hippocampal plasticity. It may be that the timescale of this study – during and immediately after qualification – were too short for the anterior decreases to emerge; perhaps they take longer. A two-phase process would make sense, first the posterior volume increases and only then, under that pressure, does the anterior decrease. Of note, and in line with previous findings [151], the overall hippocampal volumes of the QTD and NQTD groups did not differ at any point, nor did they differ within subjects between Times 1 and 2. This underscores the

subtle nature of the hippocampal changes at the point they were studied here. I predict that if one were to examine the QTD in the future, the relative redistribution of hippocampal grey matter would be more obvious, with the decrease in the anterior volume in the context of increased posterior volume.

A hint that changes might be afoot in the anterior hippocampus come from the neuropsychological data. Whilst the group did not differ on any of the measures at the start of the study, at Time 2, the now established pattern [147], [151] of cognitive performance in taxi drivers noted in previous studies is evident. At Time 2, the now qualified trainees were better at judging the proximity of London landmarks, but were significantly worse at the delayed recall of a complex figure and on the delayed recall of object-place associations. The latter scores may reflect the changes in hippocampal processing that would accompany a redistribution of resources in the hippocampus [283]. Caution is required in making such a claim, however. Although differences between the groups emerged at Time 2, the direct comparisons between Time 1 and Time 2 within groups did not show any differences for the QTD. This may mean that the cognitive effects at this point and in these healthy adults of average IQ are too subtle to reliably detect, although the Time 2 data are suggestive that changes are occurring.

Overall, therefore, considering the hypotheses I outlined at the start of this chapter, I have found evidence of structural plasticity in the posterior hippocampus associated with successful acquisition of a complex spatial representation. However, it is also the case that the timescales involved may have been too short to observe the full effects, particularly in anterior hippocampus, and in terms of the impact on anterograde memory performance. Thus, aspects of several of the hypotheses have been supported. It seems that

what I have uncovered are genuine changes within subjects that are just on the cusp of being detectable. In Chapter 7, I will consider the issues raised by this experiment further.

6.3 Experiment 6

The results of Experiment 5 suggest that the development of a complex spatial representation of London results in structural changes to the hippocampus. Coupled with the results of previous cross-sectional studies, the effects of acquiring and using this mental map of the city results in increased posterior hippocampal grey matter volume, decreased anterior volume, expertise on tests of London-based knowledge, and impaired performance on the delayed recall of a complex figure and the object-place associations test. A question that arises from these findings is what happens if taxi drivers stop their navigation behaviour? Anecdotally it is suggested that skills can be lost if not practised – use it or lose it. If expertise and its expression are indeed plastic, then can the effects of expertise be reversed? I explored this question in Experiment 6, by comparing retired taxi drivers with those still working full-time, whilst matched for age and years experience taxi driving. If plasticity works ‘in reverse’ so to speak, then the retired taxi drivers should be worse than the full-time taxi drivers and so more like control participants on tests assessing London knowledge, but better at tests such as the delayed recall of a complex figure. In addition, if this hypothesis holds, then there should be a decrease in posterior hippocampal grey matter volume and an increase in anterior in the retired taxi drivers.

6.3.1 Material and methods

Participants

Recruiting retired licensed London taxi drivers proved to be surprisingly challenging. I found that the majority of London taxi drivers that are healthy continue to work, many full-time and others part-time, well into their late 70-80's. Those that are forced to retire for health reasons were typically not suitable for this study (e.g. pacemakers). I made strenuous efforts to track down suitable retired taxi drivers. Unfortunately the Public Carriage Office, so helpful for Experiment 5, could not for data protection reasons, release details of those who have retired. I asked taxi drivers taking part in my other studies to spread the word, messages about the study were put out over the in-cab computers, I placed advertisements in the main taxi driver publications and newspapers. A very high proportion of taxi drivers reside in the nearby county of Essex. I therefore placed advertisements in the most-widely circulated local Essex newspapers, in Essex retirements clubs, bowls and golf clubs. After more than two years of recruitment efforts, data for 30 subjects were obtained. Thus, Experiment 6 can be regarded as a preliminary cross-sectional study into the effect of taxi driver retirement.

There were 10 retired London taxi drivers, 10 still working London taxi drivers, and 10 retired control participants. The retired taxi drivers, full-time taxi drivers and control subjects did not differ in terms of age ($F(2,27)=2.84$; $p=0.08$) or handedness ($F(2,27)=1.09$; $p=0.3$). The control participants achieved significantly better scores on a measure of estimated verbal IQ (WTAR) ($F(2,25)=6.39$; $p=0.006$), although there were no differences between the two taxi driver groups (one retired taxi driver and one full-time taxi driver could not be tested using the WTAR). The retired taxi drivers and retired controls did not differ significantly in terms of the number of year retired ($F(1,18)=2.29$; $p=0.1$).

The retired and full-time taxi drivers were matched for the number of years experience of taxi driving, on average over 30 years of experience ($F(1,18)=0.1$; $p=0.6$). The background details of the three groups are shown on Table 23.

Table 23 - Volunteer characteristics

	Retired taxi drivers Mean (SD) (n= 10)	Full-time taxi drivers Mean (SD) (n=10)	Controls Mean(SD) (n=10)
Mean age (years)	69.50 (2.83)	66.10 (4.55)	66.00 (3.62)
Estimated verbal IQ (WTAR)*	100.77 (5.86)	99.33 (6.98)	108.60 (5.48)
Handedness - laterality index	72.10 (52.99)	90.50 (25.89)	93.30 (12.31)
Years retired	3.60 (6.32)	-	6.75 (5.43)
Years experience taxi driving	36.20 (11.43)	34.20 (11.32)	-

*Not all subjects did this test – see text

Structural MRI brain scan

A whole brain structural MRI scan was acquired for each participant as described in Chapter 2, using the MDEFT sequence without fat suppression.

Table 24 - Cognitive tests

Cognitive tests	Retired taxi drivers Mean (SD) (n=10)	Full-time taxi drivers Mean (SD) (n=10)	Retired Controls Mean (SD) (n=10)
Digit Span scaled score	10.77 (3.15)	12.75 (2.05)	13.80 (2.48)
WRMT – Faces scaled score	43.20 (2.52)	41.50 (4.14)	44.44 (3.87)
WRMT – Words scaled score	45.20 (4.46)	45.10 (3.84)	48.60 (0.84)
Logical memory (WMS) – IR	30.10 (7.20)	23.60 (5.18)	27.20(6.44)
Logical memory (WMS) – DR	25.30 (5.94)	19.00 (7.14)	23.80 (6.46)
Rey complex figure copy (/36)	34.90 (1.52)	35.50 (1.26)	36.0 (0.0)
Rey complex figure – DR (/36)	17.45 (3.31)	15.35 (4.57)	21.80 (5.97)
London landmarks recognition (/48)	39.70 (4.24)	40.20 (3.79)	39.83 (2.31)
London landmarks proximity judgments(/10)	8.10 (0.99)	8.90 (0.73)	7.83 (1.32)

WRMT=Warrington Recognition Memory Test; IR=immediate recall; DR=delayed recall (after 30 minutes); WMS=Wechsler Memory Scale. Standard deviations in parentheses.

Neuropsychological tests

A battery of neuropsychological tests was administered. The tests are listed on Table 24, and detailed in Section 3.2.1 of Chapter 3.

In addition to these tests, another interactive test was used to assess the in situ navigation skills of the taxi drivers. This test was designed and these data were collected and analysed by Hugo Spiers and Eleanor Maguire as part of another study [113].

The video game 'The Getaway' (© Sony Computer Entertainment Europe 2002) run on a Sony Playstation2 (© Sony Computer Games Inc) was used to present subjects with a ground-level first person perspective view of a simulation of central London. The game designers decided to truly recreate the city and a large team of photographers walked the streets of central London for two years recording many streets, shops, and other details. Over 110km (70 miles) of driveable roads have been accurately recreated from Ordnance Survey map data, covering fifty square kilometres (20 square miles) of the city centre. The one-way systems, working traffic lights, the busy London traffic, and an abundance of Londoners going about their business are all included. The area covered in the game stretches from Hyde Park in the west to Shoreditch and Bethnal Green in the east; from the Angel in the north to Lambeth Bridge in the south. There are no readable street signs in the game, so one has to rely on extant knowledge to navigate. Breaking all speed limits and ignoring all red traffic lights, it takes 15 minutes to travel between the furthest points east to west. Conveniently, one can simply drive freely around the city using the game console, with a normal ground-level first person perspective, in a car of one's choice, in our case a London taxi. See Figure 30 for still images from the environment (see also [331]), for additional details of the virtual reality London).

Crucially, the realistic nature of the environment, with one-way systems and traffic restrictions, embodies the need to be able to take detours and short-cuts. All of the taxi drivers confirmed that the game was very reminiscent of their experience of navigating in central London. One moves through the environment by controlling a virtual taxi cab using a game controller, consisting of two joysticks providing analogue control of acceleration, braking and steering left and right. They were instructed to navigate 'legally' as they would in actual London, observing all traffic restrictions and one-way systems. The 'Free Roaming' mode of the game was used, permitting free navigation with the normal game scenarios suspended. To avoid collisions with other vehicles in the environment, Action Replay Max software (© Datel Design and Development Ltd 2003) provided a 'cheat' modification to the game, permitting one to drive through other vehicles.

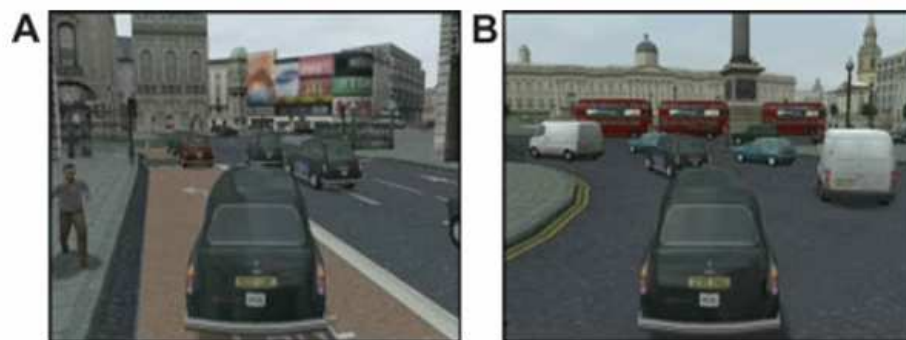


Figure 37. The virtual environment of London (UK).

Panels A) and B) show example views from within the video game "The Getaway" © 2002 Sony Computer Entertainment Europe. Panel (A) shows a view at Piccadilly Circus, panel (B) shows a view at Trafalgar Square. These images are reproduced with the kind permission of © Sony Computer Entertainment Europe.

Subjects' navigation was tested by picking up customers at start points in the city, and taking them to specified destinations. Our main criteria in designing the test was to have likely, legal routes that sampled widely across central London, with minimum overlap between routes, and minimal effects of the game (i.e. we avoided those areas where streets were not rendered in as much detail, or

where streets were missing in the game). There were thirteen routes – see Table 25. Navigation performance was measured in terms of the distance error. Based on a pilot study involving a different group of licensed London taxi drivers, consistently-chosen legal routes were established for each trial, and the ideal minimum length of these routes was computed. The distance each subject drove on each route and the ideal distance for each route were measured using Map24(UK) (<http://www.uk.map24.com>). The deviance from these likely routes was calculated as the percentage distance error (i.e. the amount of extra distance travelled compared with the ideal distance).

Table 25 - The routes in the virtual reality navigation task

Route

1. Piccadilly Circus to Big Ben
2. Big Ben to Horse Guards Parade
3. Broadwick Street to Berwick Street
4. Horse Guards Parade to Glasshouse Street
5. Kings Cross Station to the British Museum
6. Middlesex Hospital to Kings Cross Station
7. Holborn tube station to St Paul’s Cathedral
8. Holborn tube station to the River Thames
9. Berkeley Square to Berwick Street
10. Glasshouse Street to Berkeley Square
11. St Paul’s Cathedral to the Bank of England
12. Berwick Street to Golden Square
13. The British Museum to St Paul’s Cathedral

Note, the routes were not performed in the above order; order was randomised across subjects.

6.3.2 Procedure

Each participant was tested individually during one session of approximately 1.5 hours, and the taxi drivers in another separate session of approximately one hour for the virtual reality task. The order of neuropsychological testing and MRI scanning was random across subjects.

6.3.3 Data analysis

MRI scans were analysed using VBM implemented in SPM5 as described in Chapter 2, with a smoothing kernel of 8mm. Given our *a priori* interest in the hippocampus, the significance level was set at $p < 0.05$ corrected for the volume of the hippocampus using a sphere of 4mm. The significance level for the rest of the brain was set at $p < 0.05$ corrected for multiple comparisons.

For the behavioural data, basic group comparisons relating to participant characteristics were made using two-tailed t-tests. For the main analyses MANOVA was used (see Chapter 2).

6.3.4 Results

MRI data

The three groups were compared to assess differences in grey matter volume. The MRI data for one retired taxi driver were not useable. Full-time taxi drivers had greater grey matter volume in the right posterior hippocampus than retired taxi drivers (36, -36 -4, $z=3.00$ $p=0.02$; see yellow peak on Figure 38) and also when compared to control participants in the right mid hippocampus (34 -12 -14, $z=2.78$, $p=0.04$). The retired taxi drivers had greater grey matter volume in the left posterior hippocampus than the retired control participants (-24, -36, 2, $z=2.76$, $p=0.04$; see red peak on Figure 38). By contrast, greater grey matter volume was found in the control group compared with the retired taxi drivers in the anterior left hippocampus (peak -22, -8, -24; $z=3.07$; $p=0.02$). No other significant differences were found anywhere else in the brain between the groups.

Thus, these preliminary data suggest that there may be changes in hippocampal structure associated with ceasing to use a previously acquired and well-used

spatial representation of London. A step-wise pattern of results, with the full-timers having the most posterior hippocampal grey matter volume, followed by the retired, and then the non-taxi driving controls, is exactly the pattern one would predict if 'reverse' plasticity were occurring.

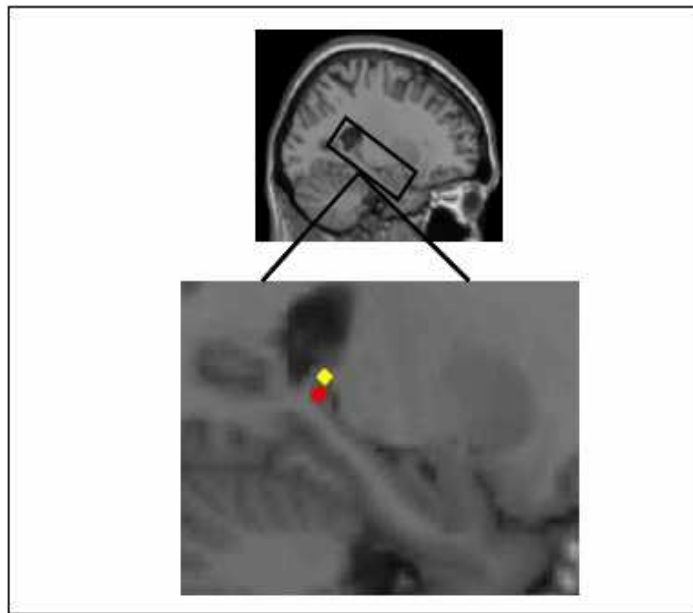


Figure 38. MRI findings in retired taxi drivers.

Full-time taxi drivers had greater grey matter volume in the posterior hippocampus than retired taxi drivers (voxel of peak difference shown in yellow), while the retired taxi drivers had greater grey matter volume in this region than the retired control participants (voxel of peak difference shown in red). Data are shown at a threshold of $p < 0.05$ corrected for the volume of the hippocampus.

Neuropsychological data

Next I compared the groups in terms of their neuropsychological profiles. Mean scores on the neuropsychological tests are shown on Table 24. An MANOVA was performed with group (retired taxi drivers, full-time taxi drivers and retired controls) as the independent variable. Five cognitive measures served as the dependent variables, namely: story immediate and delayed recall, word recognition, face recognition, and Rey complex figure delayed recall. Not all subjects completed the other tasks, and MANOVA as implemented in SPSS is not able to accommodate missing data. Thus ANOVA was used to analyse the

remaining data (digit span, London landmarks recognition memory, London landmarks proximity judgments). There was a significant difference between the groups ($F(10,44)=2.38$; $p=0.02$). The source of this difference was investigated using the tests of between-subjects effects produced by MANOVA. There was a main effect of the delayed recall of the Rey complex figure ($p=0.01$), and the Warrington word recognition memory test just reached significance ($p=0.05$). The multiple comparisons tests produced by MANOVA were used to ascertain the source of these differences. Retired controls were significantly better than full-time taxi drivers on delayed recall of the Rey complex figure ($p=0.01$) (see Figure 35). The comparison between groups for the word recognition memory test did not reach significance. There were no other significant differences. The ANOVA on the remaining three measures did not show any significant differences between the groups.

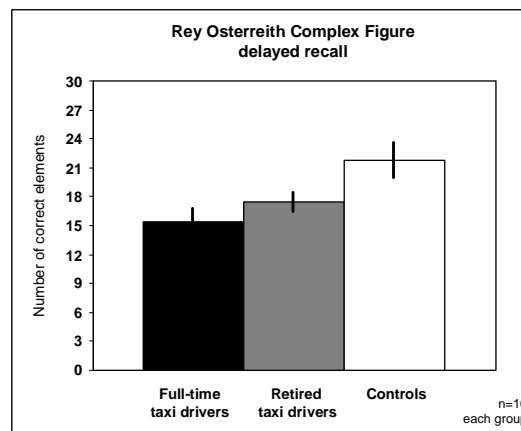


Figure 39 - Rey complex figure

Full-timers were significantly worse on the delayed recall of the Rey complex figure than retired taxi drives. Mean scores are shown +/- 1 standard error.

Eight retired and 7 full-time taxi drivers took part in the virtual reality navigation task (the control participants could not attempt this task). The average percent distance error for full-timers was 5.88% (SD 4.43) and for retired taxi drivers was 10.8% (SD 1.65), with the routes of the retired taxi drivers being significantly

longer and less ideal than those taxi drivers who were still working ($t(13)=2.18$, $p=0.05$ – see Figure 36).

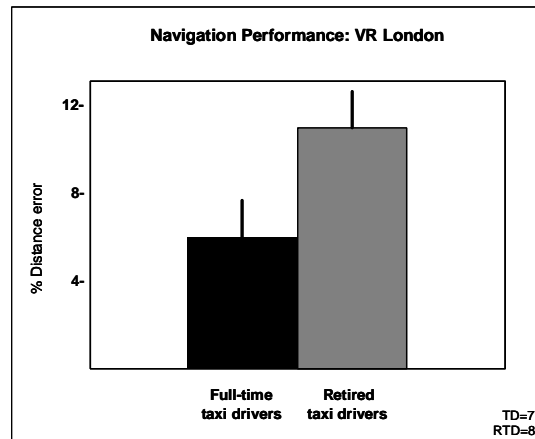


Figure 40 - Navigation performance

Full-timers were significantly better on tests of London knowledge. Mean scores are shown +/- 1 standard error.

As with the hippocampal grey matter volumes, the neuropsychological data show a step-wise pattern, with the control participants on average performing best on tests such as the delayed recall of a complex figure, followed by the retired taxi drivers, and with the full-time taxi drivers having the lowest scores. This is in clear contrast to the London navigation where the full-timers were significantly better than the retired taxi drivers. On the basis of this limited and preliminary evidence, it would seem that the intermediate position of the retired between full-time and control levels may indicate they are in transition and their anterograde memory is on the way to being 'normalised'.

6.3.5 Discussion

The preliminary data obtained here are by no means conclusive but nevertheless point towards the possibility that the structural brain changes observed in taxi drivers in previous studies may be reversible in retired taxi drivers. Full-time taxi drivers had significantly greater grey matter volume in the posterior hippocampus than retired taxi drivers, who had greater volume in this

region than the non-taxi driver retired control participants. This step-wise pattern was also apparent in the neuropsychological data, in particular for the delayed recall of the Rey complex figure, a test that is robust in indexing impairment in taxi drivers. Full-time taxi drivers performed worst, the retired taxi drivers a bit better and the controls best of all. The opposite was true for navigation around (virtual) London. On average the taxi drivers in this preliminary sample had only been retired for 3.6 years. It may be that with a longer period of retirement, performance on tests such as the Rey figure would normalise entirely. The improvement of performance on tests such as the Rey figure by the retired taxi drivers is also interesting in that it may indicate the 'elderly' hippocampus can support memory improvement, in contrast with the more traditional view of age-related memory decline.

Clearly increased numbers of participants are required to confirm these findings, and ideally a longitudinal study examining within subject effects the longer the taxi drivers are retired. Nevertheless, the possibility of plasticity effects in both directions, i.e. during acquisition and use of expertise, and then again when use of the skill ceases, serves to highlight that expertise and its behavioural consequences may be plastic and, in the case of the hippocampus at least, so are the related structural brain changes. Seasonal volume changes (increases and then subsequent decreases) in hippocampus have been reported in other species [119, 120] where increases in volume occur during the autumn/winter months when demands on spatial memory are also increased. Hence, plasticity may be maintained only when necessary for adaptational purposes arising from demands placed by interactions between the animal and the environment.

Chapter 7

General discussion

Précis

Using licensed London taxi drivers as a model system, in this thesis I reported the results of six experiments aimed at developing a body of knowledge in relation to the effect of experience (operationalised as navigational expertise/skill) on the structure and function of the healthy adult human hippocampus, with particular interest in whether experience induces structural plasticity in this brain structure. I found that successfully learning the spatial layout of a complex urban environment resulted in structural (volume) changes in the hippocampus within subjects, confirming plasticity in the healthy adult human brain in response to high level cognitive stimulation over a period of years. In a range of cross-sectional studies I was able to identify patterns of cognitive performance, both advantageous and disadvantageous, associated with navigational expertise, and the related pattern of hippocampal grey matter volume. I also established that the structural changes observed in the hippocampus associated with navigational expertise are not due to intensive cognitive stimulation *per se*, but seem to depend on the spatial nature of the stimulation. Finally, in a preliminary study, I found that hippocampal plasticity and the cognitive profile associated with navigational expertise is not unidirectional. In the case of retired taxi drivers, ceasing to use their navigational skills resulted in an apparent 'normalisation' of their hippocampal volume and cognitive performance. Having discussed the specific details of individual study findings in each experimental chapter, in this final chapter I will draw the data together to consider the results overall and how they inform the central themes of this thesis.

7.1 Plasticity in the human hippocampus

The central question in this thesis, and the one on which the other questions were implicitly predicated, was whether the intensive acquisition and use of a large-scale complex spatial representation of an environment drove structural changes in the healthy adult human hippocampus. Prior work in small mammals and birds showed that hippocampal volumes were larger in animals with a greater call on their spatial memory [118-120]. Moreover, in species where the demands on spatial memory varied seasonably, hippocampal volumes also varied similarly. Concordant with data from non-humans, previous cross-sectional studies of human London taxi drivers [147, 151], who had acquired and used a complex spatial representation of their environment, showed a different pattern of hippocampal grey matter volume compared with control participants. Taxi drivers had more grey matter in their posterior hippocampus and less in their anterior hippocampus. Moreover, hippocampal grey matter volume correlated with years spent navigating, positively in the case of the posterior hippocampal volumes and negatively for the anterior (Figure 8). In Experiment 1 I replicated these effects. Of note, the correlations suggested that the hippocampal grey matter volume effects were related to the amount of navigation experience, that continues to increase even years after qualification.

In order to definitively examine whether navigation experience drives changes in hippocampal structure in humans I conducted a longitudinal study examining trainee London taxi drivers before and after they acquired the Knowledge. Several possible outcomes could have emerged, including no effect within subjects of intense spatial knowledge acquisition on hippocampal structure because of no relationship between these factors (thus then rendering the three previous cross-sectional studies showing correlation very difficult to interpret).

Alternatively there might have been no differences because the time scale of the study was too short. Even though the study was conducted over a number of years, the previous correlation graphs suggest that navigation experience drives changes over much longer time scales. Thus measuring within subject effects just after qualification may be the least optimal time. Another possible outcome might have been that changes to hippocampal structure would be apparent when brain scans were compared within subjects.

My results show that hippocampal grey matter volume changes occur within adult humans in response to intense spatial stimulation. The whole brain (VBM) analysis technique that I employed meant that it was possible to detect grey matter differences anywhere in the brain. The grey matter changes I observed associated with acquiring the Knowledge were highly focal and restricted to the posterior hippocampi. The increase in grey matter volume in this region of the hippocampus accords with the results of all previous cross-sectional studies that showed greater grey matter volume in the posterior hippocampi of working taxi drivers compared to matched controls [147, 151], and my findings from Experiment 1. The changes in hippocampal structure were only evident in those who successfully qualified, and not in those who had undergone training for a similar length of time, but had not acquired a coherent representation of London. Moreover, it was not the case that those who successfully qualified had large posterior hippocampi to begin with, as there were no differences in hippocampal volume or cognitive profile between any of the groups at the start of the study. While differences were evident within the qualified subjects between time 1 and time 2, there were no differences between the groups at time 2. This suggests that the effects observed within subjects were subtle, and perhaps only just beginning to emerge after qualification. This fits with the cross-sectional data showing continued increases in posterior hippocampal volume for many years

after qualifying [146, 150], suggesting that the hippocampal changes are fully realised over a longer timescale than this longitudinal study.

On the same note, the previous cross-sectional studies (and Experiment 1) also reported a decrease in the volume in the anterior hippocampi of taxi drivers. A similar effect was not found here. Again, it may be that the timescale of the study was too short to observe these anterior changes. The absence of the anterior hippocampal effect may provide a clue about the sequence of events that underpin the cross-sectional findings. First the posterior hippocampal volume starts to increase and this then drives a later reduction in anterior hippocampal volume. The neuropsychological findings, discussed later in this chapter, also hint that this may well be the case. The groups did not differ at the start of training on any of the neuropsychological measures, but by time 2 the qualified trainees were significantly better on London knowledge, and significantly worse than controls on the delayed recall of a complex figure, and on the object-place associations task, thus mirroring the findings in the cross-sectional studies [151]. This latter finding may signal that changes are afoot in anterior hippocampus, changes that will manifest as an anterior volume decrease in due course. If it were practically possible, a follow up of participants in another 3-4 years would, I predict, show a further upwards trajectory in the posterior hippocampal volume, and a decrease in the anterior hippocampus as observed in all the cross sectional studies [147, 151].

The increase in posterior hippocampal grey matter volume in response to spatial learning suggests that the hippocampus is acting as a storage site for the information acquired throughout training or as a processing hub for navigation related information as proposed by the cognitive map theory. Alternatively the observed increases in posterior hippocampal grey matter volume may be the

result of increased usage of information. This latter explanation does not fit well with the findings reported in Chapter 5, if the hippocampus is not acting as a storage site or as a hub for navigational processing, one would expect to see similar changes in grey matter volume in other individuals that place high demands on memory during and after the acquisition of their knowledge. However, the above finding does not offer proof that environmental factors are invariably responsible for plasticity. There may be a genetic predisposition to plasticity in the context of environmental stimulation allowing for plasticity to express itself only in certain individuals. Genetic association studies have demonstrated influences of specific gene polymorphisms associated with memory on the volume of the hippocampus [40, 140, 142], and memory performance has been associated with the same polymorphisms [144]. Therefore the trainees that qualified may have a predisposition towards plasticity that the non-qualified individuals do not have.

I found that 49% of trainees went on to qualify. Why did 51% fail to qualify, and could this be related to them having less of a predisposition to hippocampal plasticity? Attempting to pinpoint the reasons for failure in this sample is difficult, and the non-qualified group is heterogeneous. A number of different reasons were given when asked why they ceased training, such as financial reasons, not having enough time to study, work and family commitments (most trainees were in non-taxi related full-time employment and had young families). Of note, none of the trainees that ceased reported that they found the spatial learning and memory demands too difficult to cope with, although it is possible, even likely, that the reasons mentioned above may have masked such difficulties in some individuals. The non-qualified group as a whole, whilst training for a similar timescale actually spent fewer hours training per week than their qualified counterparts. If neurogenesis or dendritogenesis are the candidate underlying

mechanisms involved in the structural changes, then neither of these mechanisms would have been tapped at the same rate/level in the non-qualified group than they would in proficient learners. It is possible that if trainees had dedicated similar amounts of time to training they all would have qualified. Likewise, it is possible that trainees that did not qualify spent less time training because they found it too difficult, but failed to report this. Thus, it is difficult to make a clear conclusion about why some trainees qualified and some did not. In the future, incorporating analysis of genetic polymorphisms may provide more definitive answers.

As well as examining the effect of intensive spatial knowledge acquisition on hippocampal volumes and memory, I also wanted to explore if plasticity is bidirectional, in other words, can the changes in hippocampal volume and memory-related advantages and disadvantages be reversed once taxi drivers stop using their expertise? The preliminary data I obtained from a small cross-sectional study of retired taxi drivers (Experiment 6) offer further support for the hypothesis that plasticity occurs as a response to environmental demands. Navigation in London in retired taxi drivers had worsened in comparison to age and experienced matched still working taxi drivers, while the retired taxi drivers were better at tests such as the delayed recall of a complex figure. Moreover, the greater grey matter volume in posterior and less in anterior hippocampus in still-working taxi drivers, and less in retired taxi drivers is completely concordant with a reversal of the effects noted in the previous cross-sectional studies. A longitudinal study assessing retired taxi drivers within subjects is required to definitively test this 'reverse plasticity' hypothesis, however, these preliminary findings in the retired cohort suggest that plasticity in response to environmental factors is continuous, reversible and related to the acquisition and continued use of expertise. It is also notable that the plasticity occurred in the elderly brain. The

improvement of performance on tests such as the complex figure by the retired taxi drivers is interesting in that it may indicate the 'elderly' hippocampus can support memory improvement, in contrast with the more traditional view of age-related memory decline. The pattern observed in both the trainee and retired taxi drivers concords with animal findings of environmentally-driven plasticity. In food-caching animals the volume of the hippocampus increases at the peak of the hoarding season when demands on spatial memory are at the highest, while the volume returns to a lower 'normal' volume during non-hording periods [119, 120, [338]. If the brain has limited resources to carry out a task, then maintaining additional neurons and/or connections that are no longer in use would be energetically expensive, especially if they can be created when needed.

What are the mechanisms that may underpin the hippocampal plasticity I observed in Experiment 5? Using in vivo techniques such as the structural MRI scanning employed here it is not possible to address this question directly. Nevertheless, one might consider a number of microscopic processes that may affect regional brain volumes by drawing on animal and electrophysiological work. Recent studies in rodents have demonstrated that when learning requires cognitive effort and where learning actually takes place (i.e. is remembered after a delay) there is an effect on the rate of hippocampal neurogenesis [202]. Moreover, the animals that learn best have more new neurons after training than those that do not learn, or do not learn efficiently [192, 204-206]. If neurogenesis is what underlies the hippocampal volume change in the qualified taxi drivers it would suggest that the neuron is the unit of learning, meaning that there is continued need for new neurons in long-term memory maintenance throughout an individual's lifetime [204, 339]. The greater grey matter volume in the posterior hippocampus of taxi drivers may be related to recruitment of new

neurons following neurogenesis [303, 304] that are pressed into the service of spatial processing. Alternatively, the development of greater communication between neurons in the form of increased synaptogenesis [305] and increases in dendritic arborisation will increase connectivity between neurons, which in turn can increase the potential memory capacity and also lead to volumetric changes [340]. Applying this logic to the pool of trainee taxi drivers in the longitudinal study one could speculate that the trainees that did not qualify, did not show grey matter changes because they did not actually learn, or did not do so efficiently. Whilst the qualified group did learn and thereby may have recruited more new neurons into the network, or increased the connectivity between neurons needed to support spatial memory.

7.2 Neuropsychological consequences of navigational expertise

The licensed London taxi drivers in my (Experiments 1, 2) and previous cross-sectional experiments [151], and the qualified trainees in the longitudinal study of this thesis showed a robust and consistent neuropsychological profile. They were comparable to matched controls in terms of learning of and recognition memory for individual stimuli, for retrograde memory for autobiographical and semantic information, executive and perceptual functions, as well as working memory. However, they were impaired on the delayed recall of a complex figure and the learning and delayed recall of object in place associations, two tests that involve object-place associations. An initial explanation for this pattern of performance was put forward in Experiment 1 whereby findings could be attributed to a visual associative memory problem. However, it could also be characterised as difficulty with between-domain associations [104]. I

investigated this further (Experiment 2) and found that taxi drivers performance was comparable to controls on all other associative memory tests, whether these were from within domain, comprising items from different modalities (between domains), presented in context or without a context. Thus the difficulty taxi drivers have on these tests cannot be attributed to a memory problem per se, nor to a general associative problem, nor to a visual memory/processing problem. The extreme demands placed on taxi drivers through both training and everyday navigation clearly involves object-place associations, and this may reduce the resources available specifically for the processing of new object-place associations.

The right hippocampus is thought to have an important role in object-place memory. Complex figures of the type I employed in this thesis are reported to be particularly sensitive to right-sided temporal lobe damage [265, 341] as is the object-place memory test [239, 240, 342, 343]. Negative correlations between the volume of the right hippocampus and delayed recall of complex figures after brain injury have been reported . Right hippocampal damage can also affect object-place learning [240], and recall [241], with the amount of tissue removal negatively related with recall performance [239, 240]. Functional MRI studies of healthy participants also show activation in the right hippocampus when engaging in navigation [344, 345], and object-place memory tasks [346, 347].

The taxi drivers included in the current experiments are all relatively young and healthy, however, the pattern of neuropsychological results suggest that the structural changes in the hippocampus may have altered its functioning. Studies of aged rats have shown that changes in information processing can lead to place cell rigidity that results in impaired spatial memory performance [348]. The interplay between rapid forms of synaptic plasticity and homeostatic

plasticity allows a network to remain sufficiently flexible so that it can accommodate experience-dependent changes, while preserving network stability. Modifying this delicate balance can reduce the storage capacity of the neuronal network [283]. Hence, in taxi drivers the change in memory load may affect the hippocampal network in a way that mimics the aged hippocampus, rendering the entire circuitry less flexible.

In addition to performing worse than controls on the tests that involved object-place associations, taxi drivers also had difficulty incorporating new information into their spatial representation of the layout of London. Trying to learn the layout of new parts of the city would involve learning new object-place associations and adding them to the existing layout of the city. Thus, if the hippocampi of taxi drivers have altered the balance of information processing in a way similar to that of the aged hippocampi, taxi drivers may have difficulty incorporating information to the map of London that is very similar to what they already know. In aged rodents, alterations to area CA3 have been associated with behavioural impairments of spatial learning. Hyperactivity in this region results in a disinhibition of the network, where rapid encoding of environmental specific features is absent and inability to inhibit access to existing memory representations (pattern completion) is increased becoming the dominant pattern of activation [348].

Although this explanation is plausible, the findings from the performance of taxi drivers on the New Town test suggest that this explanation may not be sufficient. In contrast to learning new parts of London, when learning a completely new environment, which obviously involves learning object-place associations, taxi drivers had an advantage over control participants. I propose that taxi drivers through their training have perfected their strategy for learning,

making them better than controls at this type of naturalistic task. The numerous additional cues provided during the learning of New Town, coupled with the taxi drivers' perfected navigational learning strategies may have helped them learn to a better standard than controls and drive them to form an allocentric representation of the environment from the outset. However, when the additional cues are absent, performance suffers and they cannot learn simple object-place associations as effectively due to the limited availability of neural resources.

While this would account for the difficulties processing object-place associations, it does not account for the difficulties taxi drivers encountered in New London. I suggest that in addition to possible limited storage capacity taxi drivers also had to deal with a mismatch between the information held in memory about the city and the newly 'modified' information received during learning. In this case taxi drivers did not just have to learn new information they had to learn by modifying the stored information they held about the layout of the city. As mentioned earlier, the changes in the structure of the hippocampus of taxi drivers may adversely affect the balance of the network and the connections between the subfield. Alternatively, the deficits observed in the taxi drivers can be interpreted as a very specific type of memory interference. Colgin et al. [349] proposed that if a process known as 'remapping' (remapping is the formation of distinct representations in populations of place cells after slight changes to inputs to the hippocampus) is disrupted, spatial memories that are similar to each other may not be accurately separated. It has been suggested that the dentate gyrus is where remapping starts in the hippocampus, if so, the increase in grey matter volume in taxi drivers may reflect alterations to

the neuronal circuit of the dentate gyrus in the posterior hippocampus, that interferes with pattern separation.

On the other hand, the inability to incorporate new information that is similar to existing information held in memory may be associated with the decrease in the anterior volume of the hippocampus. The hippocampus has been proposed as a match-mismatch processor given its role in associative novelty. Mismatch signalling depends on the recall of information held in memory and is thought to occur only when novel inputs overlap significantly with stored representations [350]. In turn, a number of functional neuroimaging studies have found increased activity in anterior hippocampus in response to novel associations [233-235, 351]; see also [350]. One explanation is that the anterior hippocampus may simply be unable to encode new information effectively, however taxi drivers do not exhibit problems encoding material other than object-place associations. A recent study reported that rats with hippocampal damage did not show impairments on encoding objects, places or contexts but did show impairment of object-place recall [352]. It seems that a more likely explanation may be that because the hippocampus is already storing a large amount of visuo-spatial information (the representation of London) there is much less capacity to consolidate the new spatial information into memory.

I propose therefore, that there are two different negative cognitive effects of navigational expertise in London taxi drivers: (a) reduced availability of resources making it harder to learn object-place associations, arising from limited of storage capacity; (b) there is interference from information held in memory when there is similarity between incoming and existing information. This is all quite speculative, and future work is clearly necessary in order to more

directly determine the origins of the behavioural effects in taxi drivers, and how they might be mediated by the anterior or the posterior hippocampi or a combination of both.

7.3 Spatial and non-spatial expertise

Clearly in the case of taxi drivers, we are concerned particularly with spatial representations in memory. Hippocampal volume changes in animals too are primarily associated with spatial memory demands, but of course this is may be merely because testing spatial rather than other types of memory is easiest in non-humans. In this thesis I also considered whether the plasticity of hippocampal grey matter was specifically associated with intensive and extensive knowledge acquisition in the spatial domain, or whether knowledge that is less or non-spatial would be associated with similar patterns of hippocampal grey matter volume. I addressed this question in Experiment 4 by employing medical doctors as a model of knowledge acquisition that is less spatial. I failed to find grey matter volume changes in the hippocampus or anywhere else in brain when they were compared with age, gender and IQ matched control participants.

This finding suggests that intensively learning a body of knowledge over a long time scale is not invariably associated with hippocampal volume changes, and that space may be the key to inducing structural plasticity in the hippocampus. Draganski et al.'s [53] finding of hippocampal volume changes in medical students cramming for the examinations seems at odds with both my findings and a spatial conclusion. However, careful examination of the areas of volume change in the Draganski paper reveals that the changes do not seem to be the

hippocampus proper. Whilst smoothing means that one cannot be sure of the exact location, the peaks nevertheless seem to be adjacent to the hippocampus. One could also argue that the effect in taxi drivers and not in medical doctors is because the nature of the material they are learning is very different. The knowledge elements for taxi drivers (e.g. roads, landmarks, road junctions, etc) are multiply connected with each other, and the resultant overall mental representation is a large, highly integrated coherent whole. Whilst medical doctors also acquire vast amounts of knowledge elements, and rated them as well or highly connected, it may be they are better characterised as sets of associations, with less multiple connections. It is also arguable whether their medical knowledge is represented as a single coherent entity, perhaps being better characterised as a reservoir of multiple sets of knowledge elements. One might argue, therefore, that one should test another group of subjects whose knowledge is non-spatial like medical doctors, but that is more like taxi drivers in its nature. However, it is difficult to think of another instance of expertise, besides London taxi drivers, where knowledge is so vast and multiply connected, and forms one coherent single representation.

Although the findings from mine and previous experiments implicated the right hippocampus, it is also the case that changes in the left hippocampus were also apparent, for example, in the longitudinal study within subjects changes were observed in the qualified group in both left and right posterior hippocampi. Whilst some neuropsychological studies have indicated a pre-eminent role for the right hippocampus in visuo-spatial memory [276, 353, 354], other studies have found navigation and spatial memory decrements in those with left hippocampal damage also [296, 350, 352]. In fMRI studies of navigation the left hippocampus is often activated (e.g. [72, 116, 244]). Thus both hippocampi seem to be involved in spatial representations of the environment, but in

different ways, with recent [244] work suggesting a role for the right hippocampus in allocentric place memory and the left in egocentric spatial sequencing.

7.4 Implications for the hippocampus

As well as having relevance for understanding plasticity in the human brain in general, and the effects of expertise, my findings permit some conclusions that have relevance for theories of hippocampal function. The hippocampus seems to be particularly responsive to spatial processing, given that intensive spatial learning drives structural changes. The involvement of the hippocampus in spatial memory is not short-term, the changes in hippocampal structure continue for many years after qualification. The continued increase in posterior hippocampal volume seems to be associated with the continued development and coherence of the complex spatial representation of London, which may suggest the ‘mental map’ of London, or key parts of it, is stored in the hippocampus. If one considers the main theories outlined in Chapter 1, then the findings seem most compatible with the cognitive map theory [97], that posits that a fundamental function of the hippocampus is the construction and maintenance of spatial maps of the environment. The theory suggests that the hippocampus stores an allocentric representation of the environment that allows for flexible wayfinding.

The pattern of both structural and functional findings across different stages of the working life of taxi drivers, points to the role of the hippocampus as ongoing in the maintenance and use of the allocentric representation of the environment. This is contrary to the premises of the consolidation theory [107] that holds the

hippocampus is no longer required once memory has been consolidated. According to this view, taxi drivers should have highly semanticised spatial information about the city. Although this explanation could seem to fit for well known routes, i.e. from a main train station to the financial centre of the city, it does not adequately explain the volume differences. If this were the case one would expect to see the hippocampus returning to its normal size after training, or in any case a few years after training at most (e.g. once sufficient time has lapsed for routes to be semanticised). In real life taxi drivers are required to navigate to novel destinations that are less well known or new for which the hippocampus is needed, and hence may be the reason why the changes in hippocampal volume persist. However, even taking this argument into account this would not support the positive correlation observed between the posterior grey matter volume and number of years taxi driving in several of the cross-sectional studies. It also does not offer an explanation for the 'reverse' plasticity hinted at by the retired data. Similarly, my findings are not wholly compatible with the relational theory [102, 103], given the apparently spatial preference of the hippocampus, and its role long beyond what might be expected for consolidation processes.

The findings do not fit well with the multiple trace theory either. Although the increase in grey matter could be ascribed to many overlapping memories of a route, it seems highly implausible that taxi drivers would find it effective to use the recall of a memory of a route to be able to navigate – they do not report this as a strategy. In the same vein, accessing the same route over and over again could lead to multiple traces of the same route which could be related to volume increases, but given the tens of thousands of routes that they travel, such a system would involve so many traces that the system would become

unmanageable and would inevitably lead to taxi drivers being less proficient when taking less used or new routes.

My findings also provide further evidence for functional segregation down the long axis of the hippocampus which has been observed in previous studies (see Chapter 1), and may be important for understanding the mechanisms of hippocampal operation. There are known afferent and efferent differences along the septotemporal axis of the hippocampus [355]; see also [219], as well as differences in gene expression [222]. Overall, the results in this thesis are in line with those from previous neuropsychological and MRI studies involving navigation (e.g.[147, 151, 242, 291]) where the effects diverge in anterior and posterior hippocampus. The results of Experiment 5 suggest that structural changes emerge first in the posterior hippocampus, and these may then have knock-on effects in anterior hippocampus. Whilst adding to the accruing body of evidence of distinct functions for anterior and posterior hippocampus, and now providing some information about the timescale of plasticity in the two regions, it is still not clear what exactly the anterior hippocampus does, nor how the anterior-posterior distinction maps onto the functions of individual hippocampal subfields.

7.5 Conclusions

In this thesis I presented novel data that offer evidence for experience-driven plasticity in the adult human hippocampus, the timescales and bidirectional nature of these changes, and the neuropsychological consequences of hippocampal plasticity, both positive and negative. Whilst providing new insights, my findings raise numerous additional issues.

For instance, future structure-genetic-cognitive studies are needed to directly investigate whether hippocampal plasticity is possible in any individual or more likely in those with a certain genetic predisposition. Establishing this will allow us to know if it is possible through training, to overcome possible genetic disadvantages and reach desirable levels of navigational expertise. This may benefit not only candidates to become London taxi drivers, but have implications for learning and acquiring expertise in general. Further work is also required to ascertain exactly how the coarse anterior-posterior distinction relates to specific sub-regions of the hippocampus, and the exact nature of laterality. These questions have already started to be addressed using innovative in vivo imaging techniques that employ multivariate high spatial resolution functional MRI [72] that can be mapped onto structural subfields of the hippocampus [72, 356]. Future work using methods complementary to MRI and VBM is also needed to ascertain the microscopic properties of grey matter volume changes i.e. are they due to neuronal size, neurogenesis, or perhaps dendritic or axonal arborisation. This will be a major challenge for the field as non-invasive microscopic methods are not readily available in the study of humans.

Developing a clear understanding of structural brain plasticity would have a major impact on both rehabilitation of neurological patients and education. It could influence the way in which adults and aging individuals are taught later in life, career changes, and the constant updating of skills and knowledge which are a requirement for most working individuals. Likewise, the prospect of using intensive training as part of rehabilitation programs needs to be carefully weighed against the effort required to achieve change. My data suggest that intensive training might not confer an unequivocal cognitive advantage, as might generally be assumed. There may be a naturally occurring amount of plasticity

needed for normal functioning, perhaps mediated by neurogenesis, synaptogenesis, cell apoptosis, aggregation and elimination of synapses. But the brain is constantly performing a balancing act. When any one of these processes is altered, there are consequences, which may certainly advantage an individual in some circumstances, but perhaps at a cost.

References

1. Kalbfleisch, M.L., *Functional neural anatomy of talent*. *Anat Rec B New Anat*, 2004. **277**(1): p. 21-36.
2. Ericsson, K.C., N.; Feltovich, PJ; Hoffman, RR, *The Cambridge Handbook of Expertise and Expert Performance*. 2006, New York: Cambridge University Press.
3. Grafton, S., *Developing talent: expertise and the brain*. In *Your brain on cubs: inside the heads of players and fans* ed. D. Gordon. 2008, Washington, DC: Dana Press.
4. Howe, M.J., J.W. Davidson, and J.A. Sloboda, *Innate talents: reality or myth?* *Behav Brain Sci*, 1998. **21**(3): p. 399-407; discussion 407-42.
5. Simonton, D.K., *Scientific talent, training, and performance: Intellect, personality, and genetic endowment*. *Review of General Psychology*, 2008(12): p. 28-46.
6. Owen, A.M., et al., *Putting brain training to the test*. *Nature*, 2010. **465**(7299): p. 775-8.
7. Purcell, E.M., H.C. Torrey, and R.V. Pound, *Resonance Absorption by Nuclear Magnetic Moments in a Solid*. *Physical Review*, 1946. **69**.
8. Behrens, T.E., et al., *Characterization and propagation of uncertainty in diffusion-weighted MR imaging*. *Magn Reson Med*, 2003. **50**(5): p. 1077-88.
9. Pruessner, J.C., et al., *Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories*. *Cereb Cortex*, 2000. **10**(4): p. 433-42.
10. Fischl, B., et al., *Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain*. *Neuron*, 2002. **33**(3): p. 341-55.
11. Ashburner, J. and K.J. Friston, *Voxel-based morphometry--the methods*. *Neuroimage*, 2000. **11**(6 Pt 1): p. 805-21.
12. Ashburner, J. and K.J. Friston, *Unified segmentation*. *Neuroimage*, 2005. **26**(3): p. 839-51.
13. Friston, K.J., *Testing for anatomically specified regional effects*. *Hum Brain Mapp*, 1997. **5**(2): p. 133-6.
14. Frackowiak, R.S.J., et al., *Human Brain Function 1997*: Academic Press USA.
15. Tsivilis, D., et al., *A disproportionate role for the fornix and mammillary bodies in recall versus recognition memory*. *Nat Neurosci*, 2008. **11**(7): p. 834-42.
16. Nenadic, I., H. Sauer, and C. Gaser, *Distinct pattern of brain structural deficits in subsyndromes of schizophrenia delineated by psychopathology*. *Neuroimage*, 2010. **49**(2): p. 1153-60.
17. Koolschijn, P.C., et al., *Brain volume abnormalities in major depressive disorder: a meta-analysis of magnetic resonance imaging studies*. *Hum Brain Mapp*, 2009. **30**(11): p. 3719-35.
18. Lepore, N., et al., *Pattern of hippocampal shape and volume differences in blind subjects*. *Neuroimage*, 2009. **46**(4): p. 949-57.

19. Chebat, D.R., et al., *Alterations in right posterior hippocampus in early blind individuals*. Neuroreport, 2007. **18**(4): p. 329-33.
20. Fortin, M., et al., *Wayfinding in the blind: larger hippocampal volume and supranormal spatial navigation*. Brain, 2008. **131**(Pt 11): p. 2995-3005.
21. Bonilha, L., et al., *Gray and white matter imbalance--typical structural abnormality underlying classic autism?* Brain Dev, 2008. **30**(6): p. 396-401.
22. Vythilingam, M., et al., *Smaller head of the hippocampus in Gulf War-related posttraumatic stress disorder*. Psychiatry Res, 2005. **139**(2): p. 89-99.
23. Sowell, E.R., et al., *Mapping cortical change across the human life span*. Nat Neurosci, 2003. **6**(3): p. 309-15.
24. Shaw, P., et al., *Intellectual ability and cortical development in children and adolescents*. Nature, 2006. **440**(7084): p. 676-9.
25. Giedd, J.N., et al., *Quantitative magnetic resonance imaging of human brain development: ages 4-18*. Cereb Cortex, 1996. **6**(4): p. 551-60.
26. Rapoport, J.L. and N. Gogtay, *Brain neuroplasticity in healthy, hyperactive and psychotic children: insights from neuroimaging*. Neuropsychopharmacology, 2008. **33**(1): p. 181-97.
27. Paus, T., et al., *Sexual dimorphism in the adolescent brain: Role of testosterone and androgen receptor in global and local volumes of grey and white matter*. Horm Behav, 2010. **57**(1): p. 63-75.
28. Pantelis, C., et al., *Structural brain imaging evidence for multiple pathological processes at different stages of brain development in schizophrenia*. Schizophr Bull, 2005. **31**(3): p. 672-96.
29. Shaw, P., et al., *Longitudinal mapping of cortical thickness and clinical outcome in children and adolescents with attention-deficit/hyperactivity disorder*. Arch Gen Psychiatry, 2006. **63**(5): p. 540-9.
30. Aylward, E.H., et al., *MRI volumes of amygdala and hippocampus in non-mentally retarded autistic adolescents and adults*. Neurology, 1999. **53**(9): p. 2145-50.
31. Andreasen, N.C., et al., *Intelligence and brain structure in normal individuals*. Am J Psychiatry, 1993. **150**(1): p. 130-4.
32. Narr, K.L., et al., *Relationships between IQ and regional cortical gray matter thickness in healthy adults*. Cereb Cortex, 2007. **17**(9): p. 2163-71.
33. Paus, T., et al., *In vivo morphometry of the intrasulcal gray matter in the human cingulate, paracingulate, and superior-rostral sulci: hemispheric asymmetries, gender differences and probability maps*. J Comp Neurol, 1996. **376**(4): p. 664-73.
34. Good, C.D., et al., *Cerebral asymmetry and the effects of sex and handedness on brain structure: a voxel-based morphometric analysis of 465 normal adult human brains*. Neuroimage, 2001. **14**(3): p. 685-700.
35. Kloppel, S., et al., *Nurture versus nature: long-term impact of forced right-handedness on structure of pericentral cortex and basal ganglia*. J Neurosci, 2010. **30**(9): p. 3271-5.
36. Giorgio, A., et al., *Age-related changes in grey and white matter structure throughout adulthood*. Neuroimage, 2010. **51**(3): p. 943-51.

37. Tisserand, D.J., et al., *A voxel-based morphometric study to determine individual differences in gray matter density associated with age and cognitive change over time*. Cereb Cortex, 2004. **14**(9): p. 966-73.
38. Fjell, A.M. and K.B. Walhovd, *Structural brain changes in aging: courses, causes and cognitive consequences*. Rev Neurosci, 2010. **21**(3): p. 187-221.
39. Pezawas, L., et al., *The brain-derived neurotrophic factor val66met polymorphism and variation in human cortical morphology*. J Neurosci, 2004. **24**(45): p. 10099-102.
40. Erickson, K.I., et al., *Brain-derived neurotrophic factor is associated with age-related decline in hippocampal volume*. J Neurosci, 2010. **30**(15): p. 5368-75.
41. Munte, T.F., E. Altenmuller, and L. Jancke, *The musician's brain as a model of neuroplasticity*. Nat Rev Neurosci, 2002. **3**(6): p. 473-8.
42. Gaser, C. and G. Schlaug, *Gray matter differences between musicians and nonmusicians*. Ann N Y Acad Sci, 2003. **999**: p. 514-7.
43. Sluming, V., et al., *Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians*. Neuroimage, 2002. **17**(3): p. 1613-22.
44. Bermudez, P. and R.J. Zatorre, *Differences in gray matter between musicians and nonmusicians*. Ann N Y Acad Sci, 2005. **1060**: p. 395-9.
45. Bengtsson, S.L., et al., *Extensive piano practicing has regionally specific effects on white matter development*. Nat Neurosci, 2005. **8**(9): p. 1148-50.
46. Cannonieri, G.C., et al., *Practice and perfect: length of training and structural brain changes in experienced typists*. Neuroreport, 2007. **18**(10): p. 1063-6.
47. Jancke, L., et al., *The architecture of the golfer's brain*. PLoS One, 2009. **4**(3): p. e4785.
48. Aydin, K., et al., *Increased gray matter density in the parietal cortex of mathematicians: a voxel-based morphometry study*. AJNR Am J Neuroradiol, 2007. **28**(10): p. 1859-64.
49. Mechelli, A., et al., *Structural covariance in the human cortex*. J Neurosci, 2005. **25**(36): p. 8303-10.
50. Hufner, K., et al., *Structural and functional plasticity of the hippocampal formation in professional dancers and slackliners*. Hippocampus, 2010.
51. Martin, S.B., et al., *Human experience seeking correlates with hippocampus volume: convergent evidence from manual tracing and voxel-based morphometry*. Neuropsychologia, 2007. **45**(12): p. 2874-81.
52. Draganski, B., et al., *Neuroplasticity: changes in grey matter induced by training*. Nature, 2004. **427**(6972): p. 311-2.
53. Draganski, B., et al., *Temporal and spatial dynamics of brain structure changes during extensive learning*. J Neurosci, 2006. **26**(23): p. 6314-7.
54. Ceccarelli, A., et al., *Cognitive learning is associated with gray matter changes in healthy human individuals: a tensor-based morphometry study*. Neuroimage, 2009. **48**(3): p. 585-9.
55. Boyke, J., et al., *Training-induced brain structure changes in the elderly*. J Neurosci, 2008. **28**(28): p. 7031-5.

56. Hyde, K.L., et al., *The effects of musical training on structural brain development: a longitudinal study*. Ann N Y Acad Sci, 2009. **1169**: p. 182-6.
57. Filippi, M., et al., *Motor learning in healthy humans is associated to gray matter changes: a tensor-based morphometry study*. PLoS One, 2010. **5**(4): p. e10198.
58. Casey, B.J., et al., *Imaging the developing brain: what have we learned about cognitive development?* Trends Cogn Sci, 2005. **9**(3): p. 104-10.
59. Chklovskii, D.B., B.W. Mel, and K. Svoboda, *Cortical rewiring and information storage*. Nature, 2004. **431**(7010): p. 782-8.
60. Quallo, M.M., et al., *Gray and white matter changes associated with tool-use learning in macaque monkeys*. Proc Natl Acad Sci U S A, 2009. **106**(43): p. 18379-84.
61. Nottebohm, F., *Neuronal replacement in adult brain*. Brain Res Bull, 2002. **57**(6): p. 737-49.
62. Barnea, A., *Interactions between environmental changes and brain plasticity in birds*. Gen Comp Endocrinol, 2009. **163**(1-2): p. 128-34.
63. Bottjer, S.W., E.A. Miesner, and A.P. Arnold, *Changes in neuronal number, density and size account for increases in volume of song-control nuclei during song development in zebra finches*. Neurosci Lett, 1986. **67**(3): p. 263-8.
64. Amaral, D.G., *Introduction: what is where in the medial temporal lobe?* Hippocampus, 1999. **9**(1): p. 1-6.
65. Barnes, C.A., et al., *Multistability of cognitive maps in the hippocampus of old rats*. Nature, 1997. **388**(6639): p. 272-5.
66. Lever, C., et al., *Long-term plasticity in hippocampal place-cell representation of environmental geometry*. Nature, 2002. **416**(6876): p. 90-4.
67. Muller, R.U. and J.L. Kubie, *The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells*. J Neurosci, 1987. **7**(7): p. 1951-68.
68. O'Keefe, J. and A. Speakman, *Single unit activity in the rat hippocampus during a spatial memory task*. Exp Brain Res, 1987. **68**(1): p. 1-27.
69. Jung, M.W., S.I. Wiener, and B.L. McNaughton, *Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat*. J Neurosci, 1994. **14**(12): p. 7347-56.
70. Kjelstrup, K.B., et al., *Finite scale of spatial representation in the hippocampus*. Science, 2008. **321**(5885): p. 140-3.
71. Redish, A.D., et al., *Independence of firing correlates of anatomically proximate hippocampal pyramidal cells*. J Neurosci, 2001. **21**(5): p. RC134.
72. Hassabis, D., et al., *Decoding neuronal ensembles in the human hippocampus*. Curr Biol, 2009. **19**(7): p. 546-54.
73. Andersen, P., et al., *The Hippocampus Book*. 1 ed. 2006: Oxford University Press.
74. O'Keefe, J. and N. Burgess, *Dual phase and rate coding in hippocampal place cells: theoretical significance and relationship to entorhinal grid cells*. Hippocampus, 2005. **15**(7): p. 853-66.

75. Scoville, W.B. and B. Milner, *Loss of recent memory after bilateral hippocampal lesions*. J Neurol Neurosurg Psychiatry, 1957. **20**(1): p. 11-21.
76. Corkin, S., *What's new with the amnesic patient H.M.?* Nat Rev Neurosci, 2002. **3**(2): p. 153-60.
77. Steinworth, S., B. Levine, and S. Corkin, *Medial temporal lobe structures are needed to re-experience remote autobiographical memories: evidence from H.M. and W.R.* Neuropsychologia, 2005. **43**(4): p. 479-96.
78. Bohbot, V.D. and S. Corkin, *Posterior parahippocampal place learning in H.M.* Hippocampus, 2007. **17**(9): p. 863-72.
79. Squire, L.R., *Memory systems of the brain: a brief history and current perspective*. Neurobiol Learn Mem, 2004. **82**(3): p. 171-7.
80. Cohen, N.J. and L.R. Squire, *Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that*. Science, 1980. **210**(4466): p. 207-10.
81. Bayley, P.J., et al., *The neuroanatomy of remote memory*. Neuron, 2005. **46**(5): p. 799-810.
82. Manns, J.R., et al., *Recognition memory and the human hippocampus*. Neuron, 2003. **37**(1): p. 171-80.
83. Cipolotti, L., et al., *Long-term retrograde amnesia. The crucial role of the hippocampus*. Neuropsychologia, 2001. **39**(2): p. 151-72.
84. Spiers, H.J., E.A. Maguire, and N. Burgess, *Hippocampal amnesia*. Neurocase, 2001. **7**(5): p. 357-82.
85. Nadel, L. and M. Moscovitch, *Memory consolidation, retrograde amnesia and the hippocampal complex*. Curr Opin Neurobiol, 1997. **7**(2): p. 217-27.
86. Cohen, N.J. and H. Eichenbaum, *Memory, Amnesia and the Hippocampal System* 1993, Cambridge, MA MIT Press.
87. O'Keefe, J. and J. Dostrovsky, *The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat*. Brain Res, 1971. **34**(1): p. 171-5.
88. Mayes, A.R., et al., *Associative recognition in a patient with selective hippocampal lesions and relatively normal item recognition*. Hippocampus, 2004. **14**(6): p. 763-84.
89. Holdstock, J.S., et al., *Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans?* Hippocampus, 2002. **12**(3): p. 341-51.
90. Eichenbaum, H., A.P. Yonelinas, and C. Ranganath, *The medial temporal lobe and recognition memory*. Annu Rev Neurosci, 2007. **30**: p. 123-52.
91. Kirwan, C.B. and C.E. Stark, *Overcoming interference: an fMRI investigation of pattern separation in the medial temporal lobe*. Learn Mem, 2007. **14**(9): p. 625-33.
92. Bakker, A., et al., *Pattern separation in the human hippocampal CA3 and dentate gyrus*. Science, 2008. **319**(5870): p. 1640-2.
93. Kumaran, D. and E.A. Maguire, *Novelty signals: a window into hippocampal information processing*. Trends Cogn Sci, 2009. **13**(2): p. 47-54.
94. Hassabis, D. and E.A. Maguire, *Deconstructing episodic memory with construction*. Trends Cogn Sci, 2007. **11**(7): p. 299-306.

95. Schacter, D.L. and D.R. Addis, *The cognitive neuroscience of constructive memory: remembering the past and imagining the future*. Philos Trans R Soc Lond B Biol Sci, 2007. **362**(1481): p. 773-86.
96. Buckner, R.L. and D.C. Carroll, *Self-projection and the brain*. Trends Cogn Sci, 2007. **11**(2): p. 49-57.
97. O'Keefe, J. and L. Nadel, *The Hippocampus as a Cognitive Map*. 1978: Oxford University Press.
98. Moscovitch, M. and L. Nadel, *Consolidation and the hippocampal complex revisited: in defense of the multiple-trace model*. Curr Opin Neurobiol, 1998. **8**(2): p. 297-300.
99. Moser, M.B. and E.I. Moser, *Functional differentiation in the hippocampus*. Hippocampus, 1998. **8**(6): p. 608-19.
100. Morris, R.G., et al., *Place navigation impaired in rats with hippocampal lesions*. Nature, 1982. **297**(5868): p. 681-3.
101. Tulving, E., *Episodic memory: from mind to brain*. Annu Rev Psychol, 2002. **53**: p. 1-25.
102. Eichenbaum, H., *The hippocampus and mechanisms of declarative memory*. Behav Brain Res, 1999. **103**(2): p. 123-33.
103. Eichenbaum, H., *Hippocampus: cognitive processes and neural representations that underlie declarative memory*. Neuron, 2004. **44**(1): p. 109-20.
104. Mayes, A., D. Montaldi, and E. Migo, *Associative memory and the medial temporal lobes*. Trends Cogn Sci, 2007. **11**(3): p. 126-35.
105. Staresina, B.P. and L. Davachi, *Differential encoding mechanisms for subsequent associative recognition and free recall*. J Neurosci, 2006. **26**(36): p. 9162-72.
106. Kumaran, D. and E.A. Maguire, *The human hippocampus: cognitive maps or relational memory?* J Neurosci, 2005. **25**(31): p. 7254-9.
107. Squire, L.R., *Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans*. Psychol Rev, 1992. **99**(2): p. 195-231.
108. Reed, J.M. and L.R. Squire, *Retrograde amnesia for facts and events: findings from four new cases*. J Neurosci, 1998. **18**(10): p. 3943-54.
109. Bayley, P.J., R.O. Hopkins, and L.R. Squire, *Successful recollection of remote autobiographical memories by amnesic patients with medial temporal lobe lesions*. Neuron, 2003. **38**(1): p. 135-44.
110. Teng, E. and L.R. Squire, *Memory for places learned long ago is intact after hippocampal damage*. Nature, 1999. **400**(6745): p. 675-7.
111. Rosenbaum, R.S., et al., *Remote spatial memory in an amnesic person with extensive bilateral hippocampal lesions*. Nat Neurosci, 2000. **3**(10): p. 1044-8.
112. Manns, J.R., R.O. Hopkins, and L.R. Squire, *Semantic memory and the human hippocampus*. Neuron, 2003. **38**(1): p. 127-33.
113. Maguire, E.A., R. Nannery, and H.J. Spiers, *Navigation around London by a taxi driver with bilateral hippocampal lesions*. Brain, 2006. **129**(Pt 11): p. 2894-907.
114. Nadel, L. and V. Bohbot, *Consolidation of memory*. Hippocampus, 2001. **11**(1): p. 56-60.
115. Gilboa, A., et al., *Remembering our past: functional neuroanatomy of recollection of recent and very remote personal events*. Cereb Cortex, 2004. **14**(11): p. 1214-25.

116. Chadwick, M.J., et al., *Decoding individual episodic memory traces in the human hippocampus*. *Curr Biol*, 2010. **20**(6): p. 544-7.
117. Lee, D.W., L.E. Miyasato, and N.S. Clayton, *Neurobiological bases of spatial learning in the natural environment: neurogenesis and growth in the avian and mammalian hippocampus*. *Neuroreport*, 1998. **9**(7): p. R15-27.
118. Volman, S.F., T.C. Grubb, Jr., and K.C. Schuett, *Relative hippocampal volume in relation to food-storing behavior in four species of woodpeckers*. *Brain Behav Evol*, 1997. **49**(2): p. 110-20.
119. Barnea, A. and F. Nottebohm, *Seasonal recruitment of hippocampal neurons in adult free-ranging black-capped chickadees*. *Proc Natl Acad Sci U S A*, 1994. **91**(23): p. 11217-21.
120. Smulders, T.V., A.D. Sasson, and T.J. DeVoogd, *Seasonal variation in hippocampal volume in a food-storing bird, the black-capped chickadee*. *J Neurobiol*, 1995. **27**(1): p. 15-25.
121. Pravosudov, V.V., A.S. Kitaysky, and A. Omanska, *The relationship between migratory behaviour, memory and the hippocampus: an intraspecific comparison*. *Proc Biol Sci*, 2006. **273**(1601): p. 2641-9.
122. Cnotka, J., M. Mohle, and G. Rehkemper, *Navigational experience affects hippocampus size in homing pigeons*. *Brain Behav Evol*, 2008. **72**(3): p. 233-8.
123. Galea, L.A. and B.S. McEwen, *Sex and seasonal differences in the rate of cell proliferation in the dentate gyrus of adult wild meadow voles*. *Neuroscience*, 1999. **89**(3): p. 955-64.
124. Jacobs, L.F. and W.D. Spencer, *Natural space-use patterns and hippocampal size in kangaroo rats*. *Brain Behav Evol*, 1994. **44**(3): p. 125-32.
125. Preston, S.D. and L.F. Jacobs, *Cache decision making: the effects of competition on cache decisions in Merriam's kangaroo rat (*Dipodomys merriami*)*. *J Comp Psychol*, 2005. **119**(2): p. 187-96.
126. Sherry, D.F., B.G. Galef, Jr., and M.M. Clark, *Sex and intrauterine position influence the size of the gerbil hippocampus*. *Physiol Behav*, 1996. **60**(6): p. 1491-4.
127. Cook, M.J., et al., *Hippocampal volumetric and morphometric studies in frontal and temporal lobe epilepsy*. *Brain*, 1992. **115** (Pt 4): p. 1001-15.
128. Jutila, L., et al., *MR volumetry of the entorhinal, perirhinal, and temporopolar cortices in drug-refractory temporal lobe epilepsy*. *AJNR Am J Neuroradiol*, 2001. **22**(8): p. 1490-501.
129. Jack, C.R., Jr., et al., *Rate of medial temporal lobe atrophy in typical aging and Alzheimer's disease*. *Neurology*, 1998. **51**(4): p. 993-9.
130. Teipel, S.J., et al., *Regional pattern of hippocampus and corpus callosum atrophy in Alzheimer's disease in relation to dementia severity: evidence for early neocortical degeneration*. *Neurobiol Aging*, 2003. **24**(1): p. 85-94.
131. Velakoulis, D., et al., *Selective bilateral hippocampal volume loss in chronic schizophrenia*. *Biol Psychiatry*, 2001. **50**(7): p. 531-9.
132. Flaum, M., et al., *Symptom dimensions and brain morphology in schizophrenia and related psychotic disorders*. *J Psychiatr Res*, 1995. **29**(4): p. 261-76.

133. Herbert, M.R., et al., *Dissociations of cerebral cortex, subcortical and cerebral white matter volumes in autistic boys*. *Brain*, 2003. **126**(Pt 5): p. 1182-92.
134. Frodl, T., et al., *Hippocampal changes in patients with a first episode of major depression*. *Am J Psychiatry*, 2002. **159**(7): p. 1112-8.
135. MacQueen, G.M., et al., *Course of illness, hippocampal function, and hippocampal volume in major depression*. *Proc Natl Acad Sci U S A*, 2003. **100**(3): p. 1387-92.
136. Raz, N., et al., *Trajectories of brain aging in middle-aged and older adults: regional and individual differences*. *Neuroimage*, 2010. **51**(2): p. 501-11.
137. Golomb, J., et al., *Hippocampal formation size in normal human aging: a correlate of delayed secondary memory performance*. *Learn Mem*, 1994. **1**(1): p. 45-54.
138. Raz, N., et al., *Differential effects of age and sex on the cerebellar hemispheres and the vermis: a prospective MR study*. *AJNR Am J Neuroradiol*, 1998. **19**(1): p. 65-71.
139. Gur, R.C., et al., *Brain region and sex differences in age association with brain volume: a quantitative MRI study of healthy young adults*. *Am J Geriatr Psychiatry*, 2002. **10**(1): p. 72-80.
140. Bohbot, V.D., et al., *Gray matter differences correlate with spontaneous strategies in a human virtual navigation task*. *J Neurosci*, 2007. **27**(38): p. 10078-83.
141. Bueller, J.A., et al., *BDNF Val66Met allele is associated with reduced hippocampal volume in healthy subjects*. *Biol Psychiatry*, 2006. **59**(9): p. 812-5.
142. Szeszko, P.R., et al., *Brain-derived neurotrophic factor val66met polymorphism and volume of the hippocampal formation*. *Mol Psychiatry*, 2005. **10**(7): p. 631-6.
143. Toro, R., et al., *Brain volumes and Val66Met polymorphism of the BDNF gene: local or global effects?* *Brain Struct Funct*, 2009. **213**(6): p. 501-9.
144. Egan, M.F., et al., *The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function*. *Cell*, 2003. **112**(2): p. 257-69.
145. Hariri, A.R., et al., *Brain-derived neurotrophic factor val66met polymorphism affects human memory-related hippocampal activity and predicts memory performance*. *J Neurosci*, 2003. **23**(17): p. 6690-4.
146. Sullivan, E.V., et al., *Heritability of hippocampal size in elderly twin men: equivalent influence from genes and environment*. *Hippocampus*, 2001. **11**(6): p. 754-62.
147. Maguire, E.A., et al., *Navigation-related structural change in the hippocampi of taxi drivers*. *Proc Natl Acad Sci U S A*, 2000. **97**(8): p. 4398-403.
148. Maguire, E.A., et al., *Navigation expertise and the human hippocampus: a structural brain imaging analysis*. *Hippocampus*, 2003. **13**(2): p. 250-9.
149. Maguire, E.A., et al., *Routes to remembering: the brains behind superior memory*. *Nat Neurosci*, 2003. **6**(1): p. 90-5.
150. Yates, F.A., *The Art of Memory*. 1966, London: Routledge and Kegan Paul.

151. Maguire, E.A., K. Woollett, and H.J. Spiers, *London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis*. Hippocampus, 2006. **16**(12): p. 1091-101.
152. Musacchia, G., et al., *Musicians have enhanced subcortical auditory and audiovisual processing of speech and music*. Proc Natl Acad Sci U S A, 2007. **104**(40): p. 15894-8.
153. Strait, D.L., et al., *Musical experience promotes subcortical efficiency in processing emotional vocal sounds*. Ann N Y Acad Sci, 2009. **1169**: p. 209-13.
154. Herdener, M., et al., *Musical training induces functional plasticity in human hippocampus*. J Neurosci, 2010. **30**(4): p. 1377-84.
155. Mohamed Zied, K., et al., *Bilingualism and adult differences in inhibitory mechanisms: evidence from a bilingual stroop task*. Brain Cogn, 2004. **54**(3): p. 254-6.
156. Luria, A.R., *The Mind of a Mnemonist*. 1987, Cambridge, MA: Harvard University Press.
157. Parker, E.S., L. Cahill, and J.L. McGaugh, *A case of unusual autobiographical remembering*. Neurocase, 2006. **12**(1): p. 35-49.
158. Rakic, P., *Neurogenesis in adult primate neocortex: an evaluation of the evidence*. Nat Rev Neurosci, 2002. **3**(1): p. 65-71.
159. Anderson, J.C., et al., *Map of the synapses formed with the dendrites of spiny stellate neurons of cat visual cortex*. J Comp Neurol, 1994. **341**(1): p. 25-38.
160. Kolb, B., et al., *Environmental complexity has different effects on the structure of neurons in the prefrontal cortex versus the parietal cortex or nucleus accumbens*. Synapse, 2003. **48**(3): p. 149-53.
161. Ming, G.L. and H. Song, *Adult neurogenesis in the mammalian central nervous system*. Annu Rev Neurosci, 2005. **28**: p. 223-50.
162. Gould, E., et al., *Hippocampal neurogenesis in adult Old World primates*. Proc Natl Acad Sci U S A, 1999. **96**(9): p. 5263-7.
163. Rakic, P., *Neurogenesis in adult primates*. Prog Brain Res, 2002. **138**: p. 3-14.
164. Sanai, N., et al., *Unique astrocyte ribbon in adult human brain contains neural stem cells but lacks chain migration*. Nature, 2004. **427**(6976): p. 740-4.
165. Bhardwaj, R.D., et al., *Neocortical neurogenesis in humans is restricted to development*. Proc Natl Acad Sci U S A, 2006. **103**(33): p. 12564-8.
166. Curtis, M.A., et al., *Human neuroblasts migrate to the olfactory bulb via a lateral ventricular extension*. Science, 2007. **315**(5816): p. 1243-9.
167. Sanai, N., et al., *Comment on "Human neuroblasts migrate to the olfactory bulb via a lateral ventricular extension"*. Science, 2007. **318**(5849): p. 393; author reply 393.
168. Greenough, W.T., H.M. Hwang, and C. Gorman, *Evidence for active synapse formation or altered postsynaptic metabolism in visual cortex of rats reared in complex environments*. Proc Natl Acad Sci U S A, 1985. **82**(13): p. 4549-52.
169. Eadie, B.D., V.A. Redila, and B.R. Christie, *Voluntary exercise alters the cytoarchitecture of the adult dentate gyrus by increasing cellular proliferation, dendritic complexity, and spine density*. J Comp Neurol, 2005. **486**(1): p. 39-47.

170. Moser, M.B., M. Trommald, and P. Andersen, *An increase in dendritic spine density on hippocampal CA1 pyramidal cells following spatial learning in adult rats suggests the formation of new synapses*. Proc Natl Acad Sci U S A, 1994. **91**(26): p. 12673-5.
171. Faherty, C.J., D. Kerley, and R.J. Smeyne, *A Golgi-Cox morphological analysis of neuronal changes induced by environmental enrichment*. Brain Res Dev Brain Res, 2003. **141**(1-2): p. 55-61.
172. Briones, T.L., et al., *Changes in number of synapses and mitochondria in presynaptic terminals in the dentate gyrus following cerebral ischemia and rehabilitation training*. Brain Res, 2005. **1033**(1): p. 51-7.
173. Toni, N., et al., *LTP promotes formation of multiple spine synapses between a single axon terminal and a dendrite*. Nature, 1999. **402**(6760): p. 421-5.
174. Darmopil, S., et al., *Environmental enrichment alters dentate granule cell morphology in oldest-old rat*. J Cell Mol Med, 2009. **13**(8B): p. 1845-56.
175. Leuner, B., J. Falduto, and T.J. Shors, *Associative memory formation increases the observation of dendritic spines in the hippocampus*. J Neurosci, 2003. **23**(2): p. 659-65.
176. Geinisman, Y., *Structural synaptic modifications associated with hippocampal LTP and behavioral learning*. Cereb Cortex, 2000. **10**(10): p. 952-62.
177. Geinisman, Y., et al., *Associative learning elicits the formation of multiple-synapse boutons*. J Neurosci, 2001. **21**(15): p. 5568-73.
178. Bruel-Jungerman, E., C. Rampon, and S. Laroche, *Adult hippocampal neurogenesis, synaptic plasticity and memory: facts and hypotheses*. Rev Neurosci, 2007. **18**(2): p. 93-114.
179. Marrone, D.F., *Ultrastructural plasticity associated with hippocampal-dependent learning: a meta-analysis*. Neurobiol Learn Mem, 2007. **87**(3): p. 361-71.
180. Hongpaisan, J. and D.L. Alkon, *A structural basis for enhancement of long-term associative memory in single dendritic spines regulated by PKC*. Proc Natl Acad Sci U S A, 2007. **104**(49): p. 19571-6.
181. Eriksson, P.S., et al., *Neurogenesis in the adult human hippocampus*. Nat Med, 1998. **4**(11): p. 1313-7.
182. Knoth, R., et al., *Murine features of neurogenesis in the human hippocampus across the lifespan from 0 to 100 years*. PLoS One, 2010. **5**(1): p. e8809.
183. Kempermann, G., et al., *Early determination and long-term persistence of adult-generated new neurons in the hippocampus of mice*. Development, 2003. **130**(2): p. 391-9.
184. van Praag, H., et al., *Functional neurogenesis in the adult hippocampus*. Nature, 2002. **415**(6875): p. 1030-4.
185. Hastings, N.B. and E. Gould, *Rapid extension of axons into the CA3 region by adult-generated granule cells*. J Comp Neurol, 1999. **413**(1): p. 146-54.
186. Kempermann, G., H.G. Kuhn, and F.H. Gage, *Experience-induced neurogenesis in the senescent dentate gyrus*. J Neurosci, 1998. **18**(9): p. 3206-12.

187. van Praag, H., G. Kempermann, and F.H. Gage, *Running increases cell proliferation and neurogenesis in the adult mouse dentate gyrus*. Nat Neurosci, 1999. **2**(3): p. 266-70.
188. Gould, E., et al., *Learning enhances adult neurogenesis in the hippocampal formation*. Nat Neurosci, 1999. **2**(3): p. 260-5.
189. Malberg, J.E., et al., *Chronic antidepressant treatment increases neurogenesis in adult rat hippocampus*. J Neurosci, 2000. **20**(24): p. 9104-10.
190. Tanapat, P., L.A. Galea, and E. Gould, *Stress inhibits the proliferation of granule cell precursors in the developing dentate gyrus*. Int J Dev Neurosci, 1998. **16**(3-4): p. 235-9.
191. Monje, M.L., H. Toda, and T.D. Palmer, *Inflammatory blockade restores adult hippocampal neurogenesis*. Science, 2003. **302**(5651): p. 1760-5.
192. Dayer, A.G., et al., *Short-term and long-term survival of new neurons in the rat dentate gyrus*. J Comp Neurol, 2003. **460**(4): p. 563-72.
193. Biebl, M., et al., *Analysis of neurogenesis and programmed cell death reveals a self-renewing capacity in the adult rat brain*. Neurosci Lett, 2000. **291**(1): p. 17-20.
194. Shors, T.J., et al., *Neurogenesis in the adult is involved in the formation of trace memories*. Nature, 2001. **410**(6826): p. 372-6.
195. Shors, T.J., et al., *Neurogenesis may relate to some but not all types of hippocampal-dependent learning*. Hippocampus, 2002. **12**(5): p. 578-84.
196. Shors, T.J., *Memory traces of trace memories: neurogenesis, synaptogenesis and awareness*. Trends Neurosci, 2004. **27**(5): p. 250-6.
197. Gross, C.G., *Neurogenesis in the adult brain: death of a dogma*. Nat Rev Neurosci, 2000. **1**(1): p. 67-73.
198. Waddell, J. and T.J. Shors, *Neurogenesis, learning and associative strength*. Eur J Neurosci, 2008. **27**(11): p. 3020-8.
199. Kempermann, G., *Why new neurons? Possible functions for adult hippocampal neurogenesis*. J Neurosci, 2002. **22**(3): p. 635-8.
200. Snyder, J.S., et al., *A role for adult neurogenesis in spatial long-term memory*. Neuroscience, 2005. **130**(4): p. 843-52.
201. Olariu, A., et al., *A natural form of learning can increase and decrease the survival of new neurons in the dentate gyrus*. Hippocampus, 2005. **15**(6): p. 750-62.
202. Shors, T.J., *Saving new brain cells*. Sci Am, 2009. **300**(3): p. 46-52, 54.
203. Creer, D.J., et al., *Running enhances spatial pattern separation in mice*. Proc Natl Acad Sci U S A, 2010. **107**(5): p. 2367-72.
204. Barnea, A. and F. Nottebohm, *Recruitment and replacement of hippocampal neurons in young and adult chickadees: an addition to the theory of hippocampal learning*. Proc Natl Acad Sci U S A, 1996. **93**(2): p. 714-8.
205. Becker, S., *A computational principle for hippocampal learning and neurogenesis*. Hippocampus, 2005. **15**(6): p. 722-38.
206. Becker, S., G. Macqueen, and J.M. Wojtowicz, *Computational modeling and empirical studies of hippocampal neurogenesis-dependent memory: Effects of interference, stress and depression*. Brain Res, 2009. **1299**: p. 45-54.
207. Marr, D., *Simple memory: a theory for archicortex*. Philosophical Transactions of The Royal Society B Biological Sciences

- 1971(262): p. 23-81.
208. Treves, A. and E.T. Rolls, *What determines the capacity of autoassociative memories in the brain?* . Network, 1991. **2**: p. 371-397.
 209. Imayoshi, I., et al., *Roles of continuous neurogenesis in the structural and functional integrity of the adult forebrain.* Nat Neurosci, 2008. **11**(10): p. 1153-61.
 210. Saxe, M.D., et al., *Ablation of hippocampal neurogenesis impairs contextual fear conditioning and synaptic plasticity in the dentate gyrus.* Proc Natl Acad Sci U S A, 2006. **103**(46): p. 17501-6.
 211. Madsen, T.M., et al., *Arrested neuronal proliferation and impaired hippocampal function following fractionated brain irradiation in the adult rat.* Neuroscience, 2003. **119**(3): p. 635-42.
 212. Winocur, G., et al., *Inhibition of neurogenesis interferes with hippocampus-dependent memory function.* Hippocampus, 2006. **16**(3): p. 296-304.
 213. Clelland, C.D., et al., *A functional role for adult hippocampal neurogenesis in spatial pattern separation.* Science, 2009. **325**(5937): p. 210-3.
 214. Saxe, M.D., et al., *Paradoxical influence of hippocampal neurogenesis on working memory.* Proc Natl Acad Sci U S A, 2007. **104**(11): p. 4642-6.
 215. Kempermann, G., *The neurogenic reserve hypothesis: what is adult hippocampal neurogenesis good for?* Trends Neurosci, 2008. **31**(4): p. 163-9.
 216. Kempermann, G., et al., *Natural variation and genetic covariance in adult hippocampal neurogenesis.* Proc Natl Acad Sci U S A, 2006. **103**(3): p. 780-5.
 217. Kempermann, G., et al., *Milestones of neuronal development in the adult hippocampus.* Trends Neurosci, 2004. **27**(8): p. 447-52.
 218. Bayer, S.A., J.W. Yackel, and P.S. Puri, *Neurons in the rat dentate gyrus granular layer substantially increase during juvenile and adult life.* Science, 1982. **216**(4548): p. 890-2.
 219. Gogtay, N., et al., *Dynamic mapping of normal human hippocampal development.* Hippocampus, 2006. **16**(8): p. 664-72.
 220. Deng, W., et al., *Adult-born hippocampal dentate granule cells undergoing maturation modulate learning and memory in the brain.* J Neurosci, 2009. **29**(43): p. 13532-42.
 221. Thompson, C.L., et al., *Genomic anatomy of the hippocampus.* Neuron, 2008. **60**(6): p. 1010-21.
 222. Leonardo, E.D., et al., *Molecular heterogeneity along the dorsal-ventral axis of the murine hippocampal CA1 field: a microarray analysis of gene expression.* Neuroscience, 2006. **137**(1): p. 177-86.
 223. Moser, E., M.B. Moser, and P. Andersen, *Spatial learning impairment parallels the magnitude of dorsal hippocampal lesions, but is hardly present following ventral lesions.* J Neurosci, 1993. **13**(9): p. 3916-25.
 224. Hock, B.J., Jr. and M.D. Bunsey, *Differential effects of dorsal and ventral hippocampal lesions.* J Neurosci, 1998. **18**(17): p. 7027-32.
 225. Royer, S., et al., *Distinct representations and theta dynamics in dorsal and ventral hippocampus.* J Neurosci, 2010. **30**(5): p. 1777-87.

226. Colombo, M., et al., *Functional differentiation along the anterior-posterior axis of the hippocampus in monkeys*. J Neurophysiol, 1998. **80**(2): p. 1002-5.
227. Dong, H.W., et al., *Genomic-anatomic evidence for distinct functional domains in hippocampal field CA1*. Proc Natl Acad Sci U S A, 2009. **106**(28): p. 11794-9.
228. Fanselow, M.S. and H.W. Dong, *Are the dorsal and ventral hippocampus functionally distinct structures?* Neuron, 2010. **65**(1): p. 7-19.
229. Achim, A.M. and M. Lepage, *Neural correlates of memory for items and for associations: an event-related functional magnetic resonance imaging study*. J Cogn Neurosci, 2005. **17**(4): p. 652-67.
230. Lepage, M., et al., *The interfering effect of related events on recognition memory discriminability: a functional magnetic resonance imaging study*. Brain Res Cogn Brain Res, 2005. **22**(3): p. 429-37.
231. Pihlajamaki, M., et al., *Encoding of novel picture pairs activates the perirhinal cortex: an fMRI study*. Hippocampus, 2003. **13**(1): p. 67-80.
232. Zeineh, M.M., et al., *Dynamics of the hippocampus during encoding and retrieval of face-name pairs*. Science, 2003. **299**(5606): p. 577-80.
233. Strange, B.A., et al., *Segregating the functions of human hippocampus*. Proc Natl Acad Sci U S A, 1999. **96**(7): p. 4034-9.
234. Hunkin, N.M., et al., *Novelty-related activation within the medial temporal lobes*. Neuropsychologia, 2002. **40**(8): p. 1456-64.
235. Kohler, S., et al., *Novelty responses to relational and non-relational information in the hippocampus and the parahippocampal region: a comparison based on event-related fMRI*. Hippocampus, 2005. **15**(6): p. 763-74.
236. Rains, G.D. and B. Milner, *Verbal recall and recognition as a function of depth of encoding in patients with unilateral temporal lobectomy*. Neuropsychologia, 1994. **32**(10): p. 1243-56.
237. Frisk, V. and B. Milner, *Does left temporal lobectomy adversely affect the rate at which verbal material can be processed?* Neuropsychologia, 1991. **29**(2): p. 113-23.
238. Jones-Gotman, M., et al., *Learning and retention of words and designs following excision from medial or lateral temporal-lobe structures*. Neuropsychologia, 1997. **35**(7): p. 963-73.
239. Nunn, J.A., et al., *Differential spatial memory impairment after right temporal lobectomy demonstrated using temporal titration*. Brain, 1999. **122 (Pt 1)**: p. 47-59.
240. Crane, J. and B. Milner, *What went where? Impaired object-location learning in patients with right hippocampal lesions*. Hippocampus, 2005. **15**(2): p. 216-31.
241. Smith, M.L. and B. Milner, *Right hippocampal impairment in the recall of spatial location: encoding deficit or rapid forgetting?* Neuropsychologia, 1989. **27**(1): p. 71-81.
242. Maguire, E.A., et al., *Knowing where and getting there: a human navigation network*. Science, 1998. **280**(5365): p. 921-4.
243. Maguire, E.A., R.S. Frackowiak, and C.D. Frith, *Recalling routes around london: activation of the right hippocampus in taxi drivers*. J Neurosci, 1997. **17**(18): p. 7103-10.

244. Igloi, K., et al., *Lateralized human hippocampal activity predicts navigation based on sequence or place memory*. Proc Natl Acad Sci U S A, 2010. **107**(32): p. 14466-71.
245. Svoboda, E., M.C. McKinnon, and B. Levine, *The functional neuroanatomy of autobiographical memory: a meta-analysis*. Neuropsychologia, 2006. **44**(12): p. 2189-208.
246. Bloch, F., Hansen, W. W. & Packard, M. , *Nuclear Induction*. . Physical Review, 1946. **70**.
247. Purcell, E.M., Torrey, H. C. & Pound, R. V. , *Resonance Absorption by Nuclear Magnetic Moments in a Solid*. Physical Review, 1946 **69**.
248. Ugurbil, K., et al., *Imaging at high magnetic fields: initial experiences at 4 T*. Magn Reson Q, 1993. **9**(4): p. 259-77.
249. Deichmann, R., C. Schwarzbauer, and R. Turner, *Optimisation of the 3D MDEFT sequence for anatomical brain imaging: technical implications at 1.5 and 3 T*. Neuroimage, 2004. **21**(2): p. 757-67.
250. Deichmann, R., C.D. Good, and R. Turner, *RF inhomogeneity compensation in structural brain imaging*. Magn Reson Med, 2002. **47**(2): p. 398-402.
251. Howarth, C., C. Hutton, and R. Deichmann, *Improvement of the image quality of T1-weighted anatomical brain scans*. Neuroimage, 2006. **29**(3): p. 930-7.
252. Ashburner, J. and K.J. Friston, *Nonlinear spatial normalization using basis functions*. Hum Brain Mapp, 1999. **7**(4): p. 254-66.
253. Good, C.D., et al., *A voxel-based morphometric study of ageing in 465 normal adult human brains*. Neuroimage, 2001. **14**(1 Pt 1): p. 21-36.
254. Salmond, C.H., et al., *Distributional assumptions in voxel-based morphometry*. Neuroimage, 2002. **17**(2): p. 1027-30.
255. Ashburner, J., *A fast diffeomorphic image registration algorithm*. Neuroimage, 2007. **38**(1): p. 95-113.
256. Kipps, C.M., et al., *Progression of structural neuropathology in preclinical Huntington's disease: a tensor based morphometry study*. J Neurol Neurosurg Psychiatry, 2005. **76**(5): p. 650-5.
257. Ashburner, J., J.L. Andersson, and K.J. Friston, *High-dimensional image registration using symmetric priors*. Neuroimage, 1999. **9**(6 Pt 1): p. 619-28.
258. Wechsler, *Wechsler Test of Adult Reading*. , ed. U. The Psychological Corporation: San Antonio. 2001.
259. Wechsler, *Wechsler Abbreviated Scale of Intelligence*. , ed. U. The Psychological Corporation: San Antonio. 1999.
260. Oldfield, R.C., *The assessment and analysis of handedness: The Edinburgh inventory*. Neuropsychologia, 1971. **9**(1): p. 97.
261. Sluming, V., et al., *Broca's area supports enhanced visuospatial cognition in orchestral musicians*. J Neurosci, 2007. **27**(14): p. 3799-806.
262. Biegler, R., et al., *A larger hippocampus is associated with longer-lasting spatial memory*. Proc Natl Acad Sci U S A, 2001. **98**(12): p. 6941-4.
263. Wechsler, D., *Wechsler Adult Intelligence Scale III 1997*, San Antonio, TX The Psychological Corporation.
264. Wechsler, D., *Wechsler Memory Scale-Third Edition*. 1997, San Antonio, TX: The Psychological Corporation.

265. Lezak, M.D., *Neuropsychological assessment*. 4 ed. 2004: Oxford University Press.
266. Burgess, P. and T. Shallice, *The Hayling and Brixton Tests. Test manual*. 1997, Bury St Edmunds: UK: Thames Valley Test Company.
267. Warrington, E.K. and M. James, *Visual Object and Space Perception Battery*. 1991, Bury St. Edmunds, UK: Thames Valley Test.
268. Baddeley, A.D., H. Emslie, and I. Nimmo-Smith, *Doors and People: A Test of Visual and Verbal Recall and Recognition* 1994, Bury St. Edmunds, UK Thames Valley Test Company.
269. Warrington, E.K., *Manual for Recognition Memory Test*. 1984, Windsor, England: NFER-Nelson.
270. Kopelman, M.D., *The autobiographical memory interview*. 1990, Bury St-Edmunds, UK.: Thames Valley Test Company,.
271. Osterrieth, P.A., *Le test de copie d'une figure complexe*. Archives de Psychologie, 1944. **30**: p. 206–356
272. Muter, V., S. Taylor, and F. Vargha-Khadem, *A longitudinal study of early intellectual development in hemiplegic children*. Neuropsychologia, 1997. **35**(3): p. 289-98.
273. Cohen, S., T. Kamarck, and R.J. Mermelstein, *A global measure of perceived stress*. . Journal of Health and Social Behavior, 1983. **24**: p. 385–396.
274. Spielberger, C.D., *State-Trait Anxiety Inventory for Adults (Form Y)*. 1983, California: Mind Garden.
275. Cipolotti, L. and E.A. Maguire, *A combined neuropsychological and neuroimaging study of topographical and non-verbal memory in semantic dementia*. Neuropsychologia, 2003. **41**(9): p. 1148-59.
276. Smith, M.L. and B. Milner, *The role of the right hippocampus in the recall of spatial location*. Neuropsychologia, 1981. **19**(6): p. 781-93.
277. Hubley, A.M. and D. Tremblay, *Comparability of total score performance on the Rey-Osterrieth Complex Figure and a modified Taylor Complex Figure*. J Clin Exp Neuropsychol, 2002. **24**(3): p. 370-82.
278. Hubley, A.M. and S. Jassal, *Comparability of the Rey-Osterrieth and Modified Taylor Complex Figures using total scores, completion times, and construct validation*. J Clin Exp Neuropsychol, 2006. **28**(8): p. 1482-97.
279. Casey, M.B., et al., *Does processing style affect recall of the Rey-Osterrieth or Taylor complex figures?* J Clin Exp Neuropsychol, 1991. **13**(4): p. 600-6.
280. Van Petten, C., *Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: review and meta-analysis*. Neuropsychologia, 2004. **42**: p. 1394-1413.
281. Di Paola, M., et al., *Hippocampal atrophy is the critical brain change in patients with hypoxic amnesia*. Hippocampus, 2008. **18**(7): p. 719-28.
282. Wilson, I.A., et al., *Neurocognitive aging: prior memories hinder new hippocampal encoding*. Trends Neurosci, 2006. **29**(12): p. 662-70.
283. Roth-Alpermann, C., et al., *Homeostatic shutdown of long-term potentiation in the adult hippocampus*. Proc Natl Acad Sci U S A, 2006. **103**(29): p. 11039-44.

284. Quamme, J.R., A.P. Yonelinas, and K.A. Norman, *Effect of unitization on associative recognition in amnesia*. *Hippocampus*, 2007. **17**(3): p. 192-200.
285. Vargha-Khadem, F., et al., *Differential effects of early hippocampal pathology on episodic and semantic memory*. *Science*, 1997. **277**(5324): p. 376-80.
286. Holdstock, J.S., et al., *Item recognition is less impaired than recall and associative recognition in a patient with selective hippocampal damage*. *Hippocampus*, 2005. **15**(2): p. 203-15.
287. Holdstock, J.S., et al., *Two case studies illustrating how relatively selective hippocampal lesions in humans can have quite different effects on memory*. *Hippocampus*, 2008. **18**(7): p. 679-91.
288. Aggleton, J.P., et al., *Sparing of the familiarity component of recognition memory in a patient with hippocampal pathology*. *Neuropsychologia*, 2005. **43**(12): p. 1810-23.
289. Coltheart, M., *The MRC Psycholinguistic Database*. *Quarterly Journal of Experimental Psychology*, 1981. **33A**: p. 497-505. .
290. Lynch, K., *The Image of the City* 1960: MIT Press.
291. Woollett, K. and E.A. Maguire, *Navigational expertise may compromise anterograde associative memory*. *Neuropsychologia*, 2009. **47**(4): p. 1088-95.
292. Rosenbaum, R.S., et al., *Memory for familiar environments learned in the remote past: fMRI studies of healthy people and an amnesic person with extensive bilateral hippocampal lesions*. *Hippocampus*, 2007. **17**(12): p. 1241-51.
293. Rovine, M.J. and G.D. Weisman, *Sketch-map variables as predictors of way-finding performance*. *Journal of Environmental Psychology*, 1989. **9**: p. 217-232.
294. Gale, N., et al., *The acquisition and integration of route knowledge in an unfamiliar neighborhood*. *Journal of Environmental Psychology*, 1990. **10**: p. 3-25.
295. Blades, M., *The reliability of data collected from sketch maps*. *Journal of Environmental Psychology*, 1990. **10**: p. 327-339.
296. Maguire, E.A., et al., *Topographical disorientation following unilateral temporal lobe lesions in humans*. *Neuropsychologia*, 1996. **34**(10): p. 993-1001.
297. Golledge, R.G., *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*. 1999, Baltimore, MD: Johns Hopkins Press.
298. Beck, R.J. and D. Wood, *Cognitive transformation of information from urban geographic fields to mental maps*. *Environment and Behavior*, 1976. **8**: p. 199-238.
299. Montello, D.R., *A new framework for understanding the acquisition of spatial knowledge in large-scale environments*. . *Spatial and Temporal Reasoning in Geographic Information Systems* (pp. 143-154). , ed. R.G.G. In M. J. Egenhofer. 1998: Oxford University Press.
300. Daniel, T.C., & Boster, R. S., *Measuring landscape esthetics: the scenic beauty estimation method*. Research paper RM-167. Ft. Collins, CO: US Forest Service, Rocky Mountain Forest and Range Experiment Station., 1976.
301. Brush, R.O., *The attractiveness of woodlands: perceptions of forest landowners in Massachusetts*. *Forest Science*, 1979. **25**: p. 495-506.

302. Craik, K.H., & Feimer, N. R., *Environmental assessment.* , in *Handbook of Environmental Psychology*, I.A.E. In D. Stokols, Editor. 1987, New York, Wiley.
303. Doetsch, F. and R. Hen, *Young and excitable: the function of new neurons in the adult mammalian brain.* *Curr Opin Neurobiol*, 2005. **15**(1): p. 121-8.
304. Shors, T.J., *From stem cells to grandmother cells: how neurogenesis relates to learning and memory.* *Cell Stem Cell*, 2008. **3**(3): p. 253-8.
305. Toni, N., et al., *Synapse formation on neurons born in the adult hippocampus.* *Nat Neurosci*, 2007. **10**(6): p. 727-34.
306. Bootsma, R.J. and P.C. Van Wieringen, *Timing and Attacking Forehand Drive in Table Tennis.* *Journal of Experimental Psychology. Human Perception and Performance*, 1990. **16**(1): p. 21-29.
307. Chase, W. and H. Simon, *Perception in chess.* . *Cognitive Psychology.*, 1973. **1**: p. 33-81.
308. Ericsson, K.A., R.T. Krampe, and S. Heizmann, *Can we create gifted people?* *Ciba Found Symp*, 1993. **178**: p. 222-31; discussion 232-49.
309. Reif, F. and S. Allen, *Cognition for interpreting scientific concepts: a study of acceleration.* *Cognition and Instruction*, 1992. **9**(1): p. 1-44.
310. Norman, D.A., *Categorization of action slips.* *Psychological Review Vol 88*(1) Jan 1981, 1-15, 1981.
311. Hecht, H. and D.R. Proffitt, *The price of expertise - effects of experience on the water-level task.* *Psychological Science*, 1995. **6**(2): p. 90-95.
312. Wiley, J., *Expertise as mental set: the effects of domain knowledge in creative problem solving.* *Mem Cognit*, 1998. **26**(4): p. 716-30.
313. Saariluoma, P., *Error in chess: the apperception-restructuring view.* *Psychol Res*, 1992. **54**(1): p. 17-26.
314. Luchins, A.S., *Mechanization in problem solving. The effect of Einstellung.*
 . *Psychological Monographs.* Vol. 54. 1942: American Psychological Association. 95.
315. Woltz, D.J., M.K. Gardner, and B.G. Bell, *Negative transfer errors in sequential cognitive skills: strong-but-wrong sequence application.* *J Exp Psychol Learn Mem Cogn*, 2000. **26**(3): p. 601-25.
316. Frensch, P. and R. Sternberg, *Complex Problem Solving.* 1991: Lawrence Erlbaum Associates, Inc.
317. Sternberg, R. and E. Grigorenko, *The Psychology of abilities, competencies and expertise.* 2003: Cambridge University Press.
318. Henkel, L.A., M.K. Johnson, and D.M. De Leonardis, *Aging and source monitoring: cognitive processes and neuropsychological correlates.* *J Exp Psychol Gen*, 1998. **127**(3): p. 251-68.
319. Parkin, A.J., et al., *Data-driven recognition memory: a new technique and some data on age differences.* *Psychon Bull Rev*, 2001. **8**(4): p. 812-9.
320. Wilson, I.A., et al., *Place cell rigidity correlates with impaired spatial learning in aged rats.* *Neurobiol Aging*, 2003. **24**(2): p. 297-305.
321. Hart, R.A. and G.T. Moore, eds. *The development of spatial cognition: A review.* *Image and Environment.*, ed. R. Downs and D. Stea. 1973, Chicago, Aldine.

322. Siegel, A.W., *The externalisation of cognitive maps by children and adults: in search of better ways to ask questions.* , in *Spatial Representation and Behaviour Across the Life Span: Theory and Application.*, L. Liben, A. Paterson, and N. Newcombe, Editors. 1981, New York, Academic Press. p. 1.
323. McNamara, T.P., *Mental representations of spatial relations.* . *Cognitive Psychology*, 1986. **18**: p. 87-121.
324. Evans, G.W., D.G. Marrero, and P.A. Butler, *Environmental learning and cognitive mapping.* *Environment and Behavior*, 1981. **13**: p. 83-104.
325. Hoshoooley, J.S. and D.F. Sherry, *Greater hippocampal neuronal recruitment in food-storing than in non-food-storing birds.* *Dev Neurobiol*, 2007. **67**(4): p. 406-14.
326. Sherry, D.F. and J.S. Hoshoooley, *Seasonal hippocampal plasticity in food-storing birds.* *Philos Trans R Soc Lond B Biol Sci*, 2010. **365**(1542): p. 933-43.
327. Scholz, J., et al., *Training induces changes in white-matter architecture.* *Nature Neuroscience*, 2009. **12**(11): p. 1370-71.
328. Coughlan, A.K. and S.E. Hollows, *The Adult Memory and Information Processing Battery Test Manual.* 1985, Leeds, UK: Psychology Department, University of Leeds.
329. Ratcliff, G., *Spatial thought, mental rotation and the right cerebral hemisphere.* *Neuropsychologia*, 1979. **17**(1): p. 49-54.
330. Thurstone, L.L. and T.E. Jefferey, *Space Thinking (Flags).* 1984, Chicago, IL: London House Pearson Reid.
331. Spiers, H.J. and E.A. Maguire, *Spontaneous mentalizing during an interactive real world task: an fMRI study.* *Neuropsychologia*, 2006. **44**(10): p. 1674-82.