

London Taxi Drivers and Bus Drivers: A Structural MRI and Neuropsychological Analysis

Eleanor A. Maguire,* Katherine Woollett, and Hugo J. Spiers

ABSTRACT: Licensed London taxi drivers show that humans have a remarkable capacity to acquire and use knowledge of a large complex city to navigate within it. Gray matter volume differences in the hippocampus relative to controls have been reported to accompany this expertise. While these gray matter differences could result from using and updating spatial representations, they might instead be influenced by factors such as self-motion, driving experience, and stress. We examined the contribution of these factors by comparing London taxi drivers with London bus drivers, who were matched for driving experience and levels of stress, but differed in that they follow a constrained set of routes. We found that compared with bus drivers, taxi drivers had greater gray matter volume in mid-posterior hippocampi and less volume in anterior hippocampi. Furthermore, years of navigation experience correlated with hippocampal gray matter volume only in taxi drivers, with right posterior gray matter volume increasing and anterior volume decreasing with more navigation experience. This suggests that spatial knowledge, and not stress, driving, or self-motion, is associated with the pattern of hippocampal gray matter volume in taxi drivers. We then tested for functional differences between the groups and found that the ability to acquire new visuo-spatial information was worse in taxi drivers than in bus drivers. We speculate that a complex spatial representation, which facilitates expert navigation and is associated with greater posterior hippocampal gray matter volume, might come at a cost to new spatial memories and gray matter volume in the anterior hippocampus. © 2006 Wiley-Liss, Inc.

KEY WORDS: hippocampus; memory; taxi; bus; MRI; VBM

INTRODUCTION

The volume of the hippocampus in nonhumans is known to vary as a function of the demands placed on spatial memory (Barnea and Nottebohm, 1994; Smulders et al., 1995; Volman et al., 1997; Lee et al., 1998; Biegler et al., 2001). Similar effects have been reported in licensed taxi drivers in London (UK) (Maguire et al., 2000). They undergo extensive training (in adulthood) over 2–4 yr known as acquiring “The Knowledge” that involves learning the layout of 25,000 streets in the city, thousands of places of interest, leading to stringent assessments at the Public Carriage Office in order to obtain an operating license. Examination of magnetic resonance imaging (MRI) scans using whole-brain voxel-based morphometry (VBM; Ashburner and Friston, 2005; Mechelli et al., 2005) showed greater gray

matter volume in the posterior hippocampi, and reduced anterior hippocampal gray matter volume in taxi drivers when compared with an age-matched control group. In addition, the longer taxi drivers navigated in London the greater the posterior hippocampal gray matter volume, and the more decreased the anterior gray matter volume. From this it was suggested that there may be a capacity for plastic change in the structure of the hippocampus in healthy adult humans that can accommodate the spatial representation of a very large and complex environment (Maguire et al., 2000).

If gray matter volume can be influenced by navigation experience, this has important implications for understanding the mechanisms of hippocampal operation, and for rehabilitation of memory-impaired patients. However, several other factors may have affected the findings. Navigation around London by taxi drivers necessarily involves self-motion. Since spatially selective hippocampal activity and theta oscillations require self-motion (Foster et al., 1989; O’Keefe and Recce, 1993; Terrazas et al., 2005), it could be argued that the structural changes observed are not the result of using and updating spatial representations but are instead the result of increased hippocampal activity caused by self-motion. Furthermore, movement around London is achieved by driving, which itself comprises multiple elements such as vigilance, attention, motor planning, and execution. In addition, driving all day in the poor air of a large city, dealing with customers, traffic, and fellow road users could be regarded as stressful, which might have impacted upon hippocampal gray matter (McEwen, 2001; McEwan and Magarinos, 2001).

In this study we explicitly addressed the effect of self-motion, driving experience, and stress on gray matter volume in licensed London taxi drivers. We did this by comparing taxi drivers with a control group who also spend all day driving about in the busy London traffic dealing with customers, namely London bus drivers. Unlike taxi drivers who navigate widely around the city, bus drivers operate along a constrained set of routes. If there was no difference in hippocampal gray matter volume between taxi and bus drivers, this would suggest that previous differences were explicable by nonspatial factors. If gray matter differences in taxi drivers were still apparent even accounting for self-motion, driving, and stress, then this would be further evidence that hippocampal gray matter volume may be responsive to demands placed on spatial memory by representing and navigating in a large scale space.

Wellcome Department of Imaging Neuroscience, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, United Kingdom

*Correspondence to: Eleanor A. Maguire, Wellcome Department of Imaging Neuroscience, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, United Kingdom.
E-mail: e.maguire@fil.ion.ucl.ac.uk

Accepted for publication 30 August 2006

DOI 10.1002/hipo.20233

Published online 5 October 2006 in Wiley InterScience (www.interscience.wiley.com).

If the pattern of greater posterior and reduced anterior hippocampal gray matter volume as reported by Maguire et al. (2000) was observed in this new cohort of taxi drivers, we sought to address another question. In addition to structural hippocampal differences, are there any functional differences evident in taxi drivers, in particular, might they acquire “The Knowledge” at the expense of something else?

MATERIALS AND METHODS

Participants

Thirty-five healthy male volunteers participated in the study. Of these, 18 were licensed London taxi drivers, and 17 were London bus drivers. The background details of the two groups are shown in Table 1. All taxi drivers had completed “The Knowledge” training, had passed the necessary Public Carriage Office examinations, and obtained a full (green badge) license. The London bus drivers had received formal training to drive a bus. Importantly, the two groups did not differ in the number of years they had been driving as taxi or bus drivers in London ($t(33) = 1.63$; $P = 0.11$). They all worked full-time, with no significant difference between the groups in the number of working hours per week ($t(33) = -0.34$; $P = 0.74$). Prior to commencing full-time driving, the taxi drivers also drove around London during their training, for between 2 and 4 yr. In contrast, bus driver basic training takes approximately 6 weeks. We therefore selected bus drivers, the majority of whom had a 2–4 yr history of driving (e.g., driving a delivery van) prior to becoming bus drivers, in order to control for overall amount of time driving in London. This driving tended to be along regular routes in restricted areas. None of the bus drivers had worked as licensed London taxi drivers or mini-cab drivers. None was training to be a licensed taxi driver or had ever been involved in such training. Of note, buses in London are one-person operated. That is, the driver also deals with the customers, and takes the fares, in a similar manner to taxi drivers. Bus drivers rotate around a repertoire of up to 12 routes. The taxi and bus drivers did not differ in terms of age ($t(33) = 1.97$; $P = 0.06$), handedness (laterality index) as measured using the Edinburgh Handedness Inventory (Oldfield, 1971) ($t(33) = 0.79$; $P = 0.43$), or age at which they left school ($t(33) = -0.09$; $P = 0.92$). As 8/17 London bus drivers were not born in the UK, and English was their second language, it was not appropriate to use vocabulary-based tests to estimate IQ. Instead, we assessed visual information processing and abstract reasoning skills using the Matrix Reasoning subtest of the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999). The mean scaled score for both groups was in the average range, and did not differ significantly ($t(33) = 0.03$; $P = 0.97$). The bus drivers who were not UK-born had resided in London on average for 14.23 yr (standard deviation 10.13; range 5–35 yr), and spoke English extremely well. All participants gave informed written consent to participation in the study in accordance with the local research ethics committee.

TABLE 1.

Participant Characteristics

	London taxi drivers	London bus drivers
Mean age (yr)	39.33 (4.49)	35.88 (5.79)
Mean full-time driving experience in London (yr)	10.94 (5.25)	7.71 (6.47)
Mean age at leaving school (yr)	16.50 (0.90)	16.53 (1.00)
Mean laterality index (handedness)	69.72 (43.40)	57.12 (51.15)
Mean scaled score on matrix reasoning	11.61 (1.76)	11.59 (2.26)

SDs in parentheses. Group comparisons revealed no significant differences on any of these background characteristics (see text).

MRI Scan

Whole brain structural MRI scans were acquired on a 1.5 T Sonata whole body scanner (Siemens Medical Systems, Erlangen, Germany), with a whole-body coil for RF transmission and an 8-element phased-array head coil for signal reception, using a modified driven equilibrium Fourier transform (MDEFT) sequence (Ugurbil et al., 1993). Parameters were optimized as described in the literature (Deichmann et al., 2004): for each volunteer, 176 sagittal partitions were acquired with an image matrix of 256×224 (Read \times Phase). Twofold 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/4.15 ms/650 ms, BW = 85 Hz/Px, $\alpha = 20^\circ$. To increase the signal-to-noise ratio, an asymmetric position of the inversion pulse within the magnetization preparation experiment (duration TI) was chosen, and the delay between the initial saturation and the inversion amounted to 40% of TI (Deichmann et al., 2004). The total duration of the scan was 12 min. Special RF excitation pulses were used to compensate for B1 inhomogeneities of the transmit coil (Deichmann et al., 2002). 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.

Neuropsychological Test Battery

A test battery was employed to assess a range of basic cognitive, memory, and affective functions. Most of the tests were standardized tests with published normative data (see details later). We compared the two groups with each other on these tests, and also with the normative data where appropriate. A small number of additional nonstandard tests were also used.

To assess working memory and attention, the digit span subtest from the Wechsler Adult Intelligence Scale III (Wechsler, 1998) was used. Immediate and delayed recall (at 30 min) of verbal material was tested using the logical memory (stories) subtest from the Wechsler Memory Scale Revised (Wechsler, 1987). Perceptual organization and the free recall of visual material were tested

using the copy and delayed recall (at 30 min) of the Rey-Osterrieth Complex Figure (Osterrieth, 1944; see also normative data in Loring et al., 1990; Meyers and Meyers, 1995). Recognition memory was assessed using the Warrington Recognition Memory Test for words and faces (Warrington, 1984). Recognition of London landmarks was examined using a specially designed test. Subjects were shown 48 color photographs of landmarks one after another. Half of the photographs 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. In the case of all photographs, electronic alterations were made to exclude background cues to their whereabouts. Target and distractor landmarks were randomly intermixed. The test format was a yes/no recognition test where subjects were asked to state whether they recognized each landmark as a famous London landmark or not. The test was not formally timed; however, subjects on average took about 2–5 s per photograph. Subjects' knowledge of the spatial relationships between London landmarks was tested using a proximity judgments task. Stimuli were color photographs each depicting a famous London landmark. On each trial, subjects had to judge which of two London landmarks was closer (as the crow flies) to a third London landmark. There were 10 trials.

A number of widely used questionnaires were administered to assess stress and anxiety levels. The Perceived Stress Scale (Cohen et al., 1983) comprises 10 questions, and measures the degree to which situations in life are seen as stressful within the last month. The State-Trait Anxiety Inventory (Spielberger, 1983) is divided into two parts; the first assesses the subject's feelings of anxiety at the moment of testing (state), while the second part assesses the subject's anxiety generally in daily life (trait). Two other simple rating scales were constructed for this study, one measuring general life stress and the other general job stress. Ratings were made on a scale of 1–10, where 10 was very high stress.

Procedure

Each subject was tested individually during one session of approximately 2 h. The order of neuropsychological testing and MRI scanning was random across subjects. All subjects underwent neuropsychological assessment (taxi $n = 18$, bus $n = 17$). Two taxi drivers became claustrophobic shortly after commencement of the MRI scan, and scanning had to be terminated. The two groups who had MRI scans (taxi $n = 16$, bus $n = 17$) were still matched on participant characteristics with no significant differences between them for age ($t(31) = 1.93$; $P = 0.06$), handedness ($t(31) = 0.84$; $P = 0.41$), education ($t(31) = 0.1$; $P = 0.92$), driving experience in London ($t(31) = 2.01$; $P = 0.06$), and Matrix Reasoning score ($t(31) = 0.14$; $P = 0.89$).

Data Analysis

MRI data were analyzed using VBM implemented in the Statistical Parametric Mapping software (SPM5, Wellcome Department of Imaging Neuroscience, London, UK). This method, which permits automatic whole-brain analysis, has been described in detail elsewhere (Ashburner and Friston, 2005; Mechelli et al.,

2005). Briefly, this involves a number of fully automated preprocessing steps including extraction of brain, spatial normalization into stereotactic (MNI) space, segmentation into gray and white matter and CSF compartments, correction for volume changes induced by spatial normalization (modulation), and smoothing with a 10 mm full width at half maximum (FWHM) isotropic gaussian kernel. Analyses focused on gray matter. The two groups (taxi, bus) were compared using a two sample t -test to investigate differences in gray matter volume. In addition, parametric effects on gray matter volume of behavioral performance and participant characteristics were also examined. The effects of global gray matter volume and subject age were excluded by modeling them as confounding variables. Given that we used modulation, VBM in this context is comparing the absolute volume of gray matter structures. Our use of total gray matter volume as a covariate of no interest allows us to observe brain regions of absolute gray matter volume difference that cannot be explained by total gray matter volume differences. 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 behavioral data, basic group comparisons relating to participant characteristics were made using two-tailed t -tests. For the main analyses, data were screened for outliers, homogeneity of variance, and to ascertain whether 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 $P < 0.05$. Group (taxi, bus) was the independent variable, and the neuropsychological/stress measures were the dependent variables. Where the MANOVA indicated a significant effect, the between-subjects tests were employed to ascertain the source of the significance with a threshold of $P < 0.05$. Pearson correlations were used to assess relationships between variables identified as significant in the above analyses, with a significance threshold of $P < 0.05$ two tailed. A repeated measures analysis of variance (ANOVA) was used to analyze the Rey-Osterrieth Complex Figure error types.

RESULTS

Group Differences

In the first instance the two groups were compared to assess differences in gray matter volume anywhere in the brain. Greater gray matter volume was found in the taxi drivers when compared with the bus drivers in the right mid-posterior hippocampus (peak (x, y, z) 32, -22 , -14 ; $z = 4.01$; extent in $y -18$ to -33 mm) and the left mid-posterior hippocampus (peak -32 , -16 , -16 ; $z = 4.2$; extent in $y -14$ to -30 mm) (see Fig. 1A). By contrast, less gray matter volume was found in the taxi drivers when compared with the bus drivers in the very anterior head of the right (peak 24, -4 , -24 ; $z = 4.05$; extent in $y -4$ to -12 mm) and left hippocampus (peak -26 , -6 , -22 ; $z = 4.22$; extent in $y -4$ to -10 mm) (see Fig. 1B). No other gray matter differences between the groups

were apparent elsewhere in the brain. Of note, there was no difference in gray matter volume in the caudate nucleus (even at an uncorrected threshold) between taxi and bus drivers. It has been

shown using functional MRI (fMRI) that repeatedly following a single short route can lead to activation of the caudate nucleus rather than the hippocampus (Hartley et al., 2003; Iaria et al.,

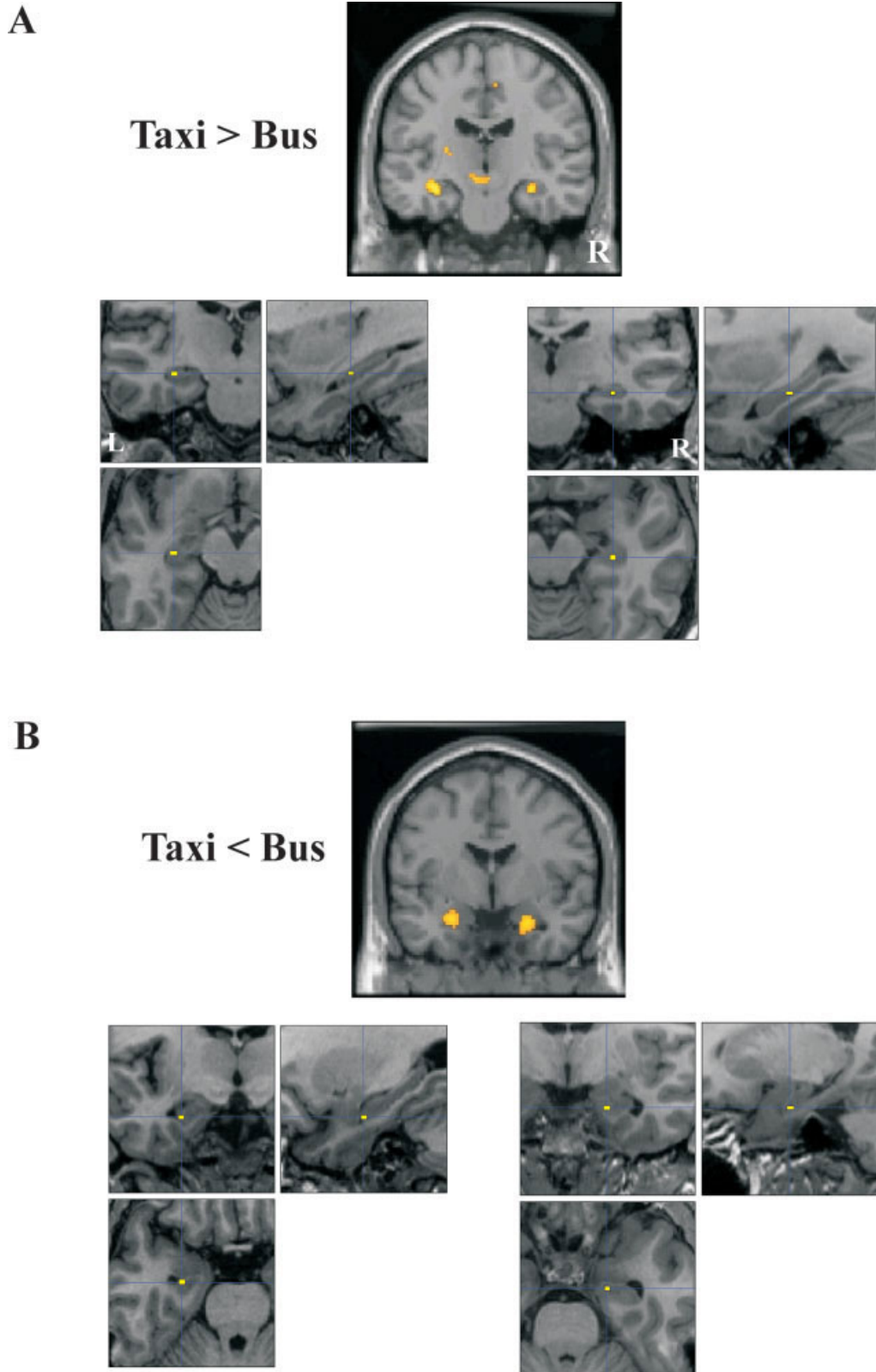


FIGURE 1

TABLE 2.

Performance on the Neuropsychological Tests

Measure	London taxi drivers		London bus drivers		Group comparisons
	Mean score	Mean percentile	Mean score	Mean percentile	
Digit span (scaled score)	11.17 (2.40)	66	11.24 (3.07)	66	ns
Story recall (immediate)	27.22 (4.45)	56	29.00 (4.96)	64	ns
Story recall (delayed)	23.33 (4.56)	56	25.47 (6.19)	62	ns
Rey-Osterrieth complex figure (copy)	33.89 (2.47)	80	34.88 (1.76)	90	ns
Rey-Osterrieth complex figure (delayed recall)	17.33 (4.49)	14	21.32 (5.07)	42	$P = 0.02$
Recognition memory test words	47.17 (2.72)	50	48.76 (1.75)	75	$P = 0.05$
Recognition memory test faces	44.00 (3.89)	50	43.18 (3.89)	50	ns
London landmark recognition memory test	40.46 (2.77)	n/a	35.41 (4.70)	n/a	$P = 0.0001$
London landmark proximity judgements	8.33 (1.09)	n/a	6.88 (1.41)	n/a	$P = 0.009$

SDs in parentheses. ns = not significantly different. n/a = not applicable.

2003; Voermans et al., 2004; and in nonhumans Packard and McGaugh, 1996; White and McDonald, 2002). However, London bus drivers do not follow one single route for years. Instead, the routes are long (across London) and their route repertoire is more extensive (see Materials and Methods).

Having observed differences in hippocampal gray matter volume, we next wanted to know how the two groups compared in terms of their neuropsychological profiles. Mean scores on the neuropsychological tests are shown in Table 2. A MANOVA was performed with group (taxi drivers, bus drivers) as the independent variable. The nine cognitive measures served as the dependent variables, namely: digit span, story immediate recall, story delayed recall, Rey-Osterrieth Complex Figure Copy, Rey-Osterrieth Complex Figure delayed recall, word recognition, face recognition, London landmark recognition, London proximity judgments. There was an overall significant difference between the two groups ($F(9,25) = 4.621$; $P = 0.001$). The source of this difference was investigated using the tests of between-subjects effects produced by MANOVA. There were four main effects. Taxi drivers were significantly better than bus drivers on the London landmark recognition memory test ($F(1,33) = 15.75$; $P = 0.0001$) and London proximity judgments test ($F(1,33) = 11.73$; $P = 0.009$) (see Figs. 2A,B). By contrast, taxi drivers were significantly poorer than bus drivers on the delayed recall of the Rey-Osterrieth Complex Figure ($F(1,33) = 6.08$; $P = 0.02$) (see Fig. 2C). Bus drivers also scored better than taxi drivers on word recogni-

tion ($F(1,33) = 4.19$; $P = 0.05$). However, this difference only just reached statistical significance, and both groups scored very highly and well within the normal range on this test.

If one examines the published normative data for the standardized tests, bus drivers are typically scoring around or above the 50th percentile (see Table 2). This was also the case for taxi drivers, including on the word recognition memory test, but with one exception. Taxi drivers scored on average between 10th and 20th percentile on the delayed recall of the Rey-Osterrieth Complex Figure. This is below what one would have expected for healthy males of this age. An assessment of the types of error made on the Rey-Osterrieth Complex Figure delayed recall was carried out by classifying errors into one of four types. To correct for overall error numbers, errors were expressed as a percentage of the total errors made: pure omission (taxi = 57%; bus = 61%); elements correctly placed but incorrectly drawn (taxi = 11%; bus = 16%); elements correctly drawn but incorrectly placed (taxi = 20%; bus = 19%); elements incorrectly drawn and incorrectly placed (taxi = 12%; bus = 4%). There was no significant difference in error pattern between the two groups ($F(3,31) = 0$; $P = 1.0$).

Thus, we have found that London taxi drivers have greater gray matter volume in the mid-posterior hippocampi and are better at identifying London landmarks and knowing their proximal relations than London bus drivers. By contrast, taxi drivers have less gray matter volume in their anterior hippocampi and are worse at acquiring or retrieving new visuo-spatial information.

FIGURE 1. Gray matter differences between London taxi and bus drivers. **A, Top panel:** greater gray matter volume in the mid-posterior hippocampi in London taxi drivers when compared with London bus drivers, shown on coronal section from the standard SPM canonical MRI. Another area in the vicinity of the red nucleus (0, -20, -10; $z = 4.47$) also had greater gray matter volume in taxi drivers, although this did not survive whole brain correction. Montage below left shows the location of the peak voxel in the left hippocampus on high resolution (medial temporal lobe magnified) MRI scan sections from a taxi driver chosen at random: top left coronal, top right sagittal, bottom left axial. Montage below right shows the

location of the peak voxel in the right hippocampus on the same types of section. **B, Top panel:** less gray matter volume in the very anterior head of the hippocampi in taxi drivers when compared with bus drivers, shown on coronal section from the standard SPM canonical MRI. Montage below left shows the location of the peak voxel in the left hippocampus on high resolution (medial temporal lobe magnified) MRI scan sections from another taxi driver chosen at random: top left coronal, top right sagittal, bottom left axial. Montage below right shows the location of the peak voxel in the right hippocampus on the same types of section. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

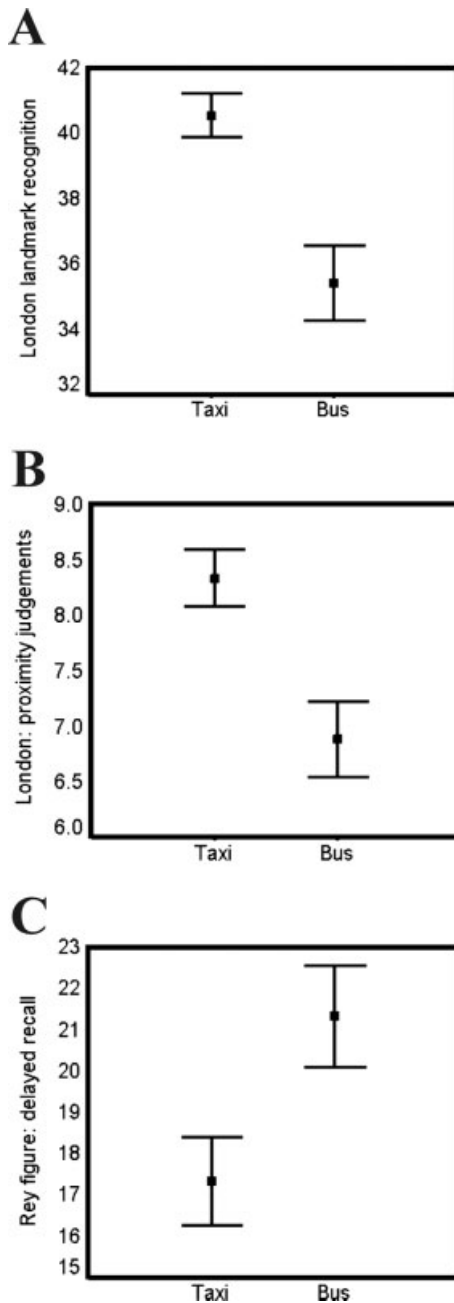


FIGURE 2. Significant main effects for the neuropsychological tests. **A:** London taxi drivers were significantly better than London bus drivers at identifying London landmarks from among visually similar distractors and **B,** making judgments about proximal relations between London landmarks. **C:** London taxi drivers were significantly poorer at recalling newly-learned visual information (Rey-Osterrieth Complex Figure) after a delay than London bus drivers. Bars represent SE.

Driving constantly in a large and busy city such as London, with traffic, polluted air, and dealing with members of the public could be regarded as stressful. High levels of stress in taxi drivers might have influenced the VBM and neuropsychological findings. To assess this, we examined the taxi and bus drivers on a range of stress and anxiety measures. Mean scores are shown on

TABLE 3.
Mean Scores on Stress and Anxiety Measures

	London taxi drivers	London bus drivers
Perceived stress scale	16.39 (7.96)	15.59 (5.68)
Life stress rating	5.50 (2.72)	4.29 (1.75)
Job stress rating	4.94 (2.55)	4.29 (1.96)
State trait anxiety inventory state	29.50 (7.99)	26.41 (9.46)
State trait anxiety inventory trait	38.11 (12.31)	34.35 (9.91)

SDs in parentheses. Group comparisons revealed no significant differences on any of these measures (see text).

Table 3. A MANOVA was performed with group (taxi drivers, bus drivers) as the independent variable. The five stress/anxiety measures served as the dependent variables, namely: Perceived Stress Scale, Life Stress rating, Job Stress rating, State-Trait Anxiety Inventory (state), State-Trait Anxiety Inventory (trait). The taxi drivers and bus drivers did not differ significantly on this range of measures ($F(5,29) = 0.64$; $P = 0.66$). Our simple Life and Job Stress ratings showed both groups to be in the middle of the scale, with “average” life and job stress. The original normative data for the Perceived Stress Scale and State-Trait Anxiety Inventory are not appropriate for our samples. Our subjects were not from the same culture, education system, or era where and when the normative data were gathered. This does not undermine the value of these tests; however, using these norms for comparison with our data serves no useful purpose. The important issue with regard to these measures is that the taxi and bus drivers did not differ, thus allowing us to answer our question as to whether group differences in perceived stress and anxiety explained hippocampal volume differences, which they did not.

Effect of Experience

Taxi drivers did not differ from bus drivers in terms of the amount of experience driving in London. As the two groups only differed in the extent of their knowledge of London’s layout, this would seem to support the idea that the structural differences in taxi drivers may relate to spatial memory. Nevertheless, we also examined directly the effect of experience on gray matter volume by entering number of years engaged in professional driving in London as a covariate of interest in the VBM analysis (controlling for subject age, see Materials and Methods). Nowhere in the brain of bus drivers was gray matter volume found to covary as a function of amount of time driving. However, among taxi drivers, time taxi driving did covary with gray matter volume. Specifically, gray matter volume the right posterior hippocampus (peak 32, -34, -2; $z = 3.86$; extent in $y -30$ to -36 mm) increased in taxi drivers the longer they had been taxi driving (see Fig. 3A). The correlation between gray matter volume and the number of years taxi driving was significant ($r = 0.72$, $P = 0.002$). By contrast, gray matter volume in the anterior hippocampi decreased in taxi drivers the longer they

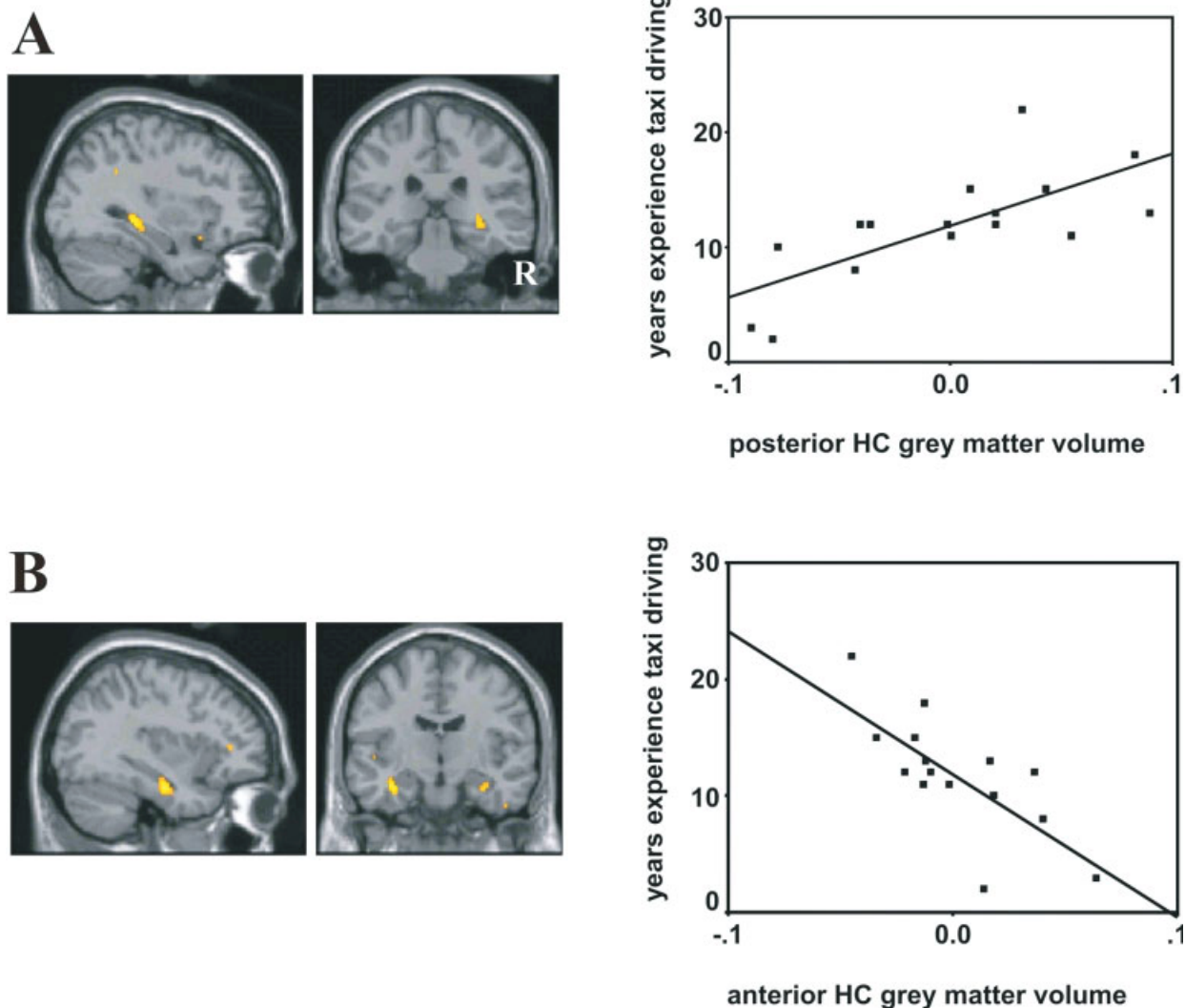


FIGURE 3. The effect of taxi driving experience on gray matter volume. **A:** Sagittal and coronal sections (left and middle panels) showing right posterior hippocampal gray matter volume varies positively with time taxi driving. Right panel shows a plot of the correlation, with increasing gray matter volume the more years of taxi driving. **B:** Sagittal and coronal sections (left and middle panels) showing anterior hippocampal gray matter volume varies nega-

tively with time taxi driving. Right panel shows a plot of the correlation (at the peak voxel in the right anterior hippocampus) with decreasing gray matter volume the more years of taxi driving. Note the plot points for two subjects overlap. HC = hippocampal. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

had been taxi driving; left anterior hippocampus (peak 34, -10, -24; $z = 3.48$; extent in $y -6$ to -12 mm) and right anterior hippocampus (peak -34, -12, -22; $z = 3.54$; extent in $y -8$ to -14 mm) (see Fig. 3B). The correlations between gray matter volume and the number of years taxi driving were significant (left anterior hippocampus $r = -0.73$, $P = 0.001$; right anterior hippocampus $r = -0.69$, $P = 0.002$). This confirms that driving experience or self-motion per se does not appear to influence gray matter volume, but rather the key factor seems to be utilizing a complex spatial representation over years of navigation.

There was a trend for performance on delayed recall of the Rey-Osterrieth Complex Figure to correlate negatively with years experience taxi driving ($r = -0.44$, $P = 0.04$ one tailed, $P = 0.07$ two tailed). Thus, the longer taxi driving, the lower the score on

Rey-Osterrieth Complex Figure delayed recall. No such effect was apparent in bus drivers ($r = 0.07$, $P = 0.78$).

No other neuropsychological or stress variables correlated significantly with years experience or gray matter volume in either group.

DISCUSSION

In this study we compared licensed London taxi drivers with London bus drivers using VBM and neuropsychological tests. The groups were matched for gender, age, education, intellectual level, and amount of driving experience in London. They also rated their feelings of stress and anxiety equivalently. We show

that compared with bus drivers, taxi drivers had greater gray matter volume in the posterior hippocampi and less gray matter volume in the anterior hippocampi. Furthermore, years of navigation experience correlated with hippocampal gray matter volume only in taxi drivers, with right posterior gray matter volume increasing and anterior gray matter volume decreasing with more navigation experience. In addition to structural brain differences, our neuropsychological testing revealed that taxi drivers, while performing better than bus drivers on tests relating to knowledge of London, were significantly poorer at acquiring or retrieving novel visuospatial information.

Our findings are consistent with the previous study that examined gray matter in licensed London taxi drivers when compared with a group of age and gender-matched control subjects (Maguire et al., 2000). The coordinates of the voxel of peak gray matter difference in the right mid-posterior hippocampus in this study and those observed by Maguire et al. (2000) are almost identical. The peak coordinate for the left hippocampus is more anterior than that observed by Maguire et al. (2000), although the area of difference extends posteriorly and overlaps with that previously reported. The coordinates of the voxels of peak gray matter difference in the anterior hippocampi are also similar to those observed by Maguire et al. (2000). As before, these differences are at the most anterior tip of the hippocampus. This region can be difficult to distinguish from the neighboring amygdala. However, the more inferior location of the gray matter differences (see Fig. 1, and Duvernoy et al., 2005), and the similarity with the previous results which were independently verified by a traditional pixel counting region of interest analysis (ROI) focused on the hippocampus (Maguire et al., 2000) are consistent with the effects being hippocampal.

It could be argued that using bus drivers as the comparison group might have obscured potentially interesting effects associated with driving per se or with following routes repeatedly. However, given that the results of the VBM analyses from this study are so concordant with the previous study (Maguire et al., 2000) where control subjects were not in a driving profession suggests that the bus drivers were an appropriate control group. While it would have been preferable to have matched the taxi and bus driver groups in terms of their country of origin, it was not possible, given the high proportion of non-UK born people in the London bus driving profession. We ensured, however, that non-UK born bus drivers had resided in the UK for some time, and were very proficient in English. That the VBM results were very similar to those obtained in the original study (Maguire et al., 2000) where control subjects were UK born, speaks against a cohort effect explaining our findings in the present study.

While confirming the robustness of the previously reported structural results, our study extends them in several ways. In the original study, the anterior gray matter differences were identified as significant using the ROI method. We now show that these anterior effects replicate and can be measured effectively using VBM. The current findings also more clearly reveal the nature of the association between hippocampal gray matter and years navigating as a taxi driver. The plots in Figure 3 show that this relationship is best characterized as linear. A more varied distribution

of experience, coupled with better quality MRI scans, and developments in the VBM technique are likely reasons for the improved clarity of the current findings.

The two subject groups are better matched than those tested by Maguire et al. (2000) in terms of basic neuropsychological profiles. In the previous study, the control group comprised nontaxi drivers. This meant that the taxi and control groups differed considerably not only in terms of spatial representation and navigation experience, but in the self-motion, proprioceptive, vestibular, and optic cue processing that this encompasses. The influence of self-motion signals on hippocampal place cells has been suggested to be substantial (Foster et al., 1989; O'Keefe and Recce, 1993; Terrazas et al., 2005), and could have been a powerful influence on gray matter. The current findings, however, allow us to rule this out. Taxi and bus drivers had similar driving experience in London, they all drove full time, and for a comparable number of hours each day. Both flexible navigation by the taxi drivers and the more spatially constrained route following of the bus drivers relied on processing self-motion signals, yet hippocampal gray matter differences between the two groups persisted. Likewise, by comparing two groups matched for driving a vehicle on a continuous basis, with concomitant attentional, vigilance, and motoric demands, we can rule out driving as having affected the gray matter findings.

Driving a taxi in London, constantly dealing with traffic, road works, and customers, could be regarded as stressful. Glucocorticoids (GCs) are produced in response to stress. GCs can regulate neuronal metabolism, gene expression, and hippocampal synaptic plasticity (McEwan et al., 1999; McEwen, 2001; McEwan and Magarinos, 2001). Increases in plasma or salivary cortisol are often used as a biochemical marker of stress. Elevated cortisol levels have been associated with poor performance on memory tasks (Newcomer et al., 1999; Cho et al., 2000) and reduced hippocampal volume (Lupien et al., 1998). Thus, stress may have induced the anterior hippocampal volume reduction in the taxi drivers. It was not possible to measure cortisol levels in this study, although this would be interesting to consider in future studies. However, the bus drivers operated in the same environment as the taxi drivers, coping with the same stressors. On a range of stress and anxiety measures, the groups' self-ratings did not differ significantly, suggesting that stress is not the basis of the reduced anterior hippocampal gray matter volume in taxi drivers.

Thus, the results from the present study continue to permit the view that learning, representing, and using a spatial representation of a highly complex and large-scale environment is a primary function of the hippocampus in humans such that this brain region might adapt structurally to accommodate its elaboration. It might seem somewhat paradoxical that the gray matter in posterior hippocampus increases with more navigation experience, since one might have predicted instead that taxi drivers would have most knowledge when they are newly qualified (Terrazas and McNaughton, 2000). Taxi drivers report that their "mental map" of London becomes more integrated and coheres over a long period as it gets more fine-tuned through experience of the relationships among roads and between places (Maguire EA, unpublished observations). Similar findings of apparently environmentally driven plasticity have been reported in several

other groups including musicians, jugglers, and bilinguals (Munte et al., 2002; Gaser and Schlaug, 2003; Draganski et al., 2004; Mechelli et al., 2004). In these groups, positive correlations between gray matter and the time spent learning and practicing their specialization have also been found. Professional musicians show a greater increase in gray matter volume in motor and auditory areas (Gaser and Schlaug, 2003) and a frontal region (Sluming et al., 2002) the longer the time spent practicing and playing, as do early bilinguals in parietal cortex (Mechelli et al., 2004).

The correlations between time taxi driving and hippocampal gray matter suggest that experience drives the changes observed. However, it should be acknowledged that in order to categorically conclude this, a within-subjects longitudinal study over many years facilitating comparison before and after acquiring "The Knowledge" is required (and is ongoing). Such a study will also permit examination of the characteristics of those who succeed at taxi driver training, and whether innate pretraining cognitive factors and/or hippocampal volume are predictive of successful qualification. That said, in a previous study, the correlation between navigation ability in a newly learned (virtual) environment in nontaxi drivers and its relation to hippocampal volume was examined (Maguire et al., 2003). No significant relationship was found, i.e., better navigators did not have larger posterior hippocampal volume. However, this was a cross-sectional study, and does not preclude the possibility that taxi driver potential is influenced by innate factors. It should also be noted that our study focused on spatial knowledge and its relationship with the hippocampus. It could be that intense acquisition and prolonged use of nonspatial knowledge might have similar effects.

While the correlation between time taxi driving and posterior hippocampal gray matter volume is right lateralized, and may indicate a preferential role for this area in representing space (O'Keefe and Nadel, 1978; Smith and Milner, 1981; Burgess et al., 2002), by and large the results are bilateral. Previous neuropsychological studies have documented deficits in topographical memory following left medial temporal lobe damage (Maguire et al., 1996; Lambrey et al., 2003), and the left, as well as the right, hippocampus has been observed to activate in some neuroimaging studies of navigation (Ghaem et al., 1997; Maguire et al., 1998; Gron et al., 2000; Spiers and Maguire, 2006). The precise contribution of the left hippocampus to navigation, and whether it somehow relates to episodic memory, remains to be determined.

Given that licensed London taxi drivers have a particular pattern of hippocampal gray matter volume, we next asked the question what might the functional consequences of such an arrangement be on memory? Clearly, taxi drivers are expert at navigation in London, and their significantly better scores when compared with bus drivers on the London landmarks and proximity tests confirmed this. We suggest this enhanced performance might be associated with the greater posterior hippocampal gray matter volume in taxi drivers. However, acquiring or retrieving new visuo-spatial information was below average in taxi drivers, while bus drivers performed in an average fashion on the delayed recall of the Rey-Osterrieth Complex Figure. This difference cannot be explained

by difficulties perceiving, organizing, or drawing the figure as both groups scored equivalently on the copy trial. The Rey-Osterrieth Complex Figure is one of the most widely used tests of spatial memory in neuropsychology, and is reported to be particularly sensitive to right-sided damage (Meyers and Meyers, 1995; Lezak, 2004). For instance Bohbot et al. (1998) found deficits in patients with surgical removal of the right hippocampus, and more recently, a negative correlation between the volume of the right hippocampus and delayed recall of the Rey-Osterrieth Complex Figure after brain injury has been reported (Tomaiuolo et al., 2004).

While previous neuropsychological studies have not specifically linked the anterior hippocampus with performance on the delayed recall of the Rey-Osterrieth Complex Figure, we suggest the below average score observed in the taxi drivers on this test could be related to their reduced right anterior hippocampal gray matter volume. This accords with findings across species that there may be functional differentiation along the anterior-posterior axis of the hippocampus (e.g., Moser et al., 1993; Jung et al., 1994; Colombo et al., 1998; Hock and Bunsey, 1998; Maguire et al., 2000; Bannerman et al., 2004; Gogtay et al., 2006). In humans, functional neuroimaging studies have linked anterior hippocampal activity with the detection of stimulus novelty and to encoding, while the posterior hippocampus has been associated with retrieval (Lepage et al., 1998; Strange and Dolan, 2001; but see Small et al., 2001). In line with this view, and of particular relevance here, successful navigation in familiar environments has often resulted in activation of the posterior hippocampus (Hartley et al., 2003; Kumaran and Maguire, 2005; Spiers and Maguire, 2006). By contrast, learning the layout of a new virtual environment has been associated with activity in the anterior hippocampus (Wolbers and Buchel, 2005). We note, however, that scores on the delayed recall of the Rey-Osterrieth Complex Figure did not correlate with gray matter volume directly in taxi drivers. The limited variance in performance among taxi drivers might reflect an early and adverse impact of taxi training on anterograde visuo-spatial memory, and so a categorical difference rather than a correlation may result. Thus, while the present Rey-Osterrieth Complex Figure finding is intriguing and may reflect a cognitive trade-off in acquiring The Knowledge, clearly, further studies of taxi drivers are required to explore in more depth their anterograde visuo-spatial memory, and also its relationship with the anterior hippocampus.

In conclusion, licensed London taxi drivers show that humans have an amazing capacity to acquire and use for navigation a highly complex spatial representation of a large city. Our findings suggest that this is accompanied by greater gray matter volume in the posterior hippocampus. We speculate that there may be a price to pay for this expertise, however, as reflected in less gray matter volume elsewhere in the hippocampus, and a compromised ability to acquire or retrieve new visuo-spatial information. Longitudinal within-subjects studies are required to examine direct relationships between navigation expertise, changes in hippocampal volume, and their cognitive sequelae. In addition, future work using methods complementary to MRI and VBM may ascertain if gray matter volume differences are due to neuronal size, neurogenesis, or perhaps dendritic or axonal arborisation.

Acknowledgments

We thank the participants for their time and interest in this work. We are grateful to Stagecoach London and Arriva London for facilitating bus driver recruitment, and all the licensed London taxi companies, particularly Dial-a-Cab, for facilitating taxi driver recruitment. Thanks also to J. Ashburner for helpful discussions, and R. Deichmann and B. de Martino for MRI technical support. This work was funded by a Wellcome Trust senior research fellowship in basic biomedical science to EAM.

REFERENCES

- Ashburner J, Friston KJ. 2005. Unified segmentation. *Neuroimage* 26: 839–851.
- Bannerman DM, Rawlins JN, McHugh SB, Deacon RM, Yee BK, Bast T, Zhang WN, Pothuizen HH, Feldon J. 2004. Regional dissociations within the hippocampus—Memory and anxiety. *Neurosci Biobehav Rev* 28:273–283.
- Barnea A, Nottebohm F. 1994. Seasonal recruitment of hippocampal neurons in adult free-ranging black-capped chickadees. *Proc Natl Acad Sci USA* 91:11217–11221.
- Biegler R, McGregor A, Krebs JR, Healy SD. 2001. A larger hippocampus is associated with longer-lasting spatial memory. *Proc Natl Acad Sci USA* 98:6941–6944.
- Bohbot VD, Kalina M, Stepankova K, Spackova N, Petrides M, Nadel L. 1998. Spatial memory deficits in patients with lesions to the right hippocampus and to the right parahippocampal cortex. *Neuropsychologia* 36:1217–1238.
- Burgess N, Maguire EA, O'Keefe J. 2002. The human hippocampus and spatial and episodic memory. *Neuron* 35:625–641.
- Cho K, Ennaceur A, Cole JC, Suh CK. 2000. Chronic jet lag produces cognitive deficits. *J Neurosci* 20:RC66.
- Cohen S, Kamarck T, Mermelstein R. 1983. A global measure of perceived stress. *J Health Soc Behav* 24:385–396.
- Colombo M, Fernandez T, Nakamura K, Gross CG. 1998. Functional differentiation along the anterior-posterior axis of the hippocampus in monkeys. *J Neurophysiol* 80:1002–1005.
- Deichmann R, Good CD, Turner R. 2002. RF inhomogeneity compensation in structural brain imaging. *Magn Reson Med* 47:398–402.
- Deichmann R, Schwarzbauer C, Turner R. 2004. Optimisation of the 3D MDEFT sequence for anatomical brain imaging: Technical implications at 1.5 and 3 T. *Neuroimage* 21:757–767.
- Draganski B, Gaser C, Busch V, Schuierer G, Bogdahn U, May A. 2004. Neuroplasticity: Changes in grey matter induced by training. *Nature* 427:311–312.
- Duvernoy HM. 2005. *The Human Hippocampus: Functional Anatomy, Vascularisation and Serial Sections with MRI*. Berlin: Springer.
- Foster TC, Castro CA, McNaughton BL. 1989. Spatial selectivity of rat hippocampal neurons: Dependence on preparedness for movement. *Science* 244:1580–1582.
- Gaser C, Schlaug G. 2003. Gray matter differences between musicians and nonmusicians. *Ann NY Acad Sci* 999:514–517.
- Ghaem O, Mellet E, Crivello F, Tzourio N, Mazoyer B, Berthoz A, Denis M. 1997. Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. *Neuroreport* 8: 739–744.
- Gogtay N, Nugent TF III, Herman DH, Ordóñez A, Greenstein D, Hayashi KM, Clasen L, Toga AW, Giedd JN, Rapoport JL, Thompson PM. 2006. Dynamic mapping of normal human hippocampal development. *Hippocampus* 16:664–672.
- Gron G, Wunderlich AP, Spitzer M, Tomczak R, Riepe MW. 2000. Brain activation during human navigation: Gender-different neural networks as substrate of performance. *Nat Neurosci* 3:404–408.
- Hartley T, Maguire EA, Spiers HJ, Burgess N. 2003. The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron* 37:877–888.
- Hock BJ Jr, Bunsey MD. 1998. Differential effects of dorsal and ventral hippocampal lesions. *J Neurosci* 18:7027–7032.
- Iaria G, Petrides M, Dagher A, Pike B, Bohbot VD. 2003. Cognitive strategies dependent on the hippocampus and caudate nucleus in human navigation: Variability and change with practice. *J Neurosci* 23:5945–5952.
- Jung MW, Wiener SI, McNaughton BL. 1994. Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. *J Neurosci* 14:7347–7356.
- Kumaran D, Maguire EA. 2005. The human hippocampus: Cognitive maps or relational memory? *J Neurosci* 25:7254–7259.
- Lambrey S, Samson S, Dupont S, Baulac M, Berthoz A. 2003. Reference frames and cognitive strategies during navigation: Is the left hippocampal formation involved in the sequential aspects of route memory? *Int Congr Ser* 1250:261–274.
- Lee DW, Miyasato LE, Clayton NS. 1998. Neurobiological bases of spatial learning in the natural environment: Neurogenesis and growth in the avian and mammalian hippocampus. *Neuroreport* 9: R15–R27.
- Lepage M, Habib R, Tulving E. 1998. Hippocampal PET activations of memory encoding and retrieval: The HIPER model. *Hippocampus* 8:313–322.
- Lezak MD. 2004. *Neuropsychological assessment*. Oxford: Oxford University Press.
- Loring DW, Martin RC, Meador KJ, Lee GP. 1990. Psychometric construction of the Rey-Osterrieth complex figure: Methodological considerations and interrater reliability. *Arch Clin Neuropsychol* 5: 1–14.
- Lupien SJ, de LM, de SS, Convit A, Tarshish C, Nair NP, Thakur M, McEwen BS, Hauger RL, Meaney MJ. 1998. Cortisol levels during human aging predict hippocampal atrophy and memory deficits. *Nat Neurosci* 1:69–73.
- Maguire EA, Burke T, Phillips J, Staunton H. 1996. Topographical disorientation following unilateral temporal lobe lesions in humans. *Neuropsychologia* 34:993–1001.
- Maguire EA, Burgess N, Donnett JG, Frackowiak RS, Frith CD, O'Keefe J. 1998. Knowing where and getting there: A human navigation network. *Science* 280:921–924.
- Maguire EA, Gadian DG, Johnsrude IS, Good CD, Ashburner J, Frackowiak RSJ, Frith CD. 2000. Navigation-related structural change in the hippocampi of taxi drivers. *Proc Natl Acad Sci USA* 97: 4398–4403.
- Maguire EA, Spiers HJ, Good CD, Hartley T, Frackowiak RSJ, Burgess N. 2003. Navigation expertise and the human hippocampus: A structural brain imaging analysis. *Hippocampus* 13:208–217.
- McEwen BS. 2001. Plasticity of the hippocampus: Adaptation to chronic stress and allostatic load. *Ann NY Acad Sci* 933:265–277.
- McEwen BS, Magarinos AM. 2001. Stress and hippocampal plasticity: Implications for the pathophysiology of affective disorders. *Hum Psychopharmacol* 16(Suppl 1):S7–S19.
- McEwen BS, de Leon MJ, Lupien SJ, Meaney MJ. 1999. Corticosteroids, the aging brain and cognition. *Trends Endocrinol Metab* 10: 92–96.
- Mechelli A, Crinion JT, Noppeney U, O'Doherty J, Ashburner J, Frackowiak RS, Price CJ. 2004. Neurolinguistics: Structural plasticity in the bilingual brain. *Nature* 431:757.
- Mechelli A, Price CJ, Friston KJ, Ashburner J. 2005. Voxel-based morphometry of the human brain: Methods and applications. *Curr Med Imaging Rev* 1:105–113.
- Meyers JE, Meyers KR. 1995. *Rey Complex Figure Tests and Recognition Trial*. Odessa, FL: Psychological Assessment Resources.
- Moser E, Moser MB, Andersen P. 1993. Spatial learning impairment parallels the magnitude of dorsal hippocampal lesions, but is hardly present following ventral lesions. *J Neurosci* 13:3916–3925.

- Munte TF, Altenmuller E, Jancke L. 2002. The musician's brain as a model of neuroplasticity. *Nat Rev Neurosci* 3:473–478.
- Newcomer JW, Selke G, Melson AK, Hershey T, Craft S, Richards K, Alderson AL. 1999. Decreased memory performance in healthy humans induced by stress-level cortisol treatment. *Arch Gen Psychiatry* 56:527–533.
- O'Keefe J, Nadel L. 1978. *The Hippocampus as a Cognitive Map*. Oxford: Oxford University Press.
- O'Keefe J, Recce ML. 1993. Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3:317–330.
- Oldfield RC. 1971. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9:97–113.
- Osterrieth PA. 1944. Le test de copie d'une figure complexe. *Archives de Psychologie* 30:206–356, Translated by J. Corwin and FW Bylsma.
- Packard MG, McGaugh JL. 1996. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol Learn Mem* 65:65–72.
- Sluming V, Barrick T, Howard M, Cezayirli E, Mayes A, Roberts N. 2002. Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *Neuroimage* 17:1613–22.
- Small SA, Nava AS, Perera GM, DeLaPaz R, Mayeux R, Stern Y. 2001. Circuit mechanisms underlying memory encoding and retrieval in the long axis of the hippocampal formation. *Nat Neurosci* 4:442–449.
- Smith ML, Milner B. 1981. The role of the right hippocampus in the recall of spatial location. *Neuropsychologia* 19:781–793.
- Smulders TV, Sasson AD, DeVoogd TJ. 1995. Seasonal variation in hippocampal volume in a food-storing bird, the black-capped chickadee. *J Neurobiol* 27:15–25.
- Spielberger CD. 1983. *State-Trait Anxiety Inventory for Adults (Form Y)*. California: Mind Garden.
- Spiers HJ, Maguire EA. 2006. Thoughts, behavior, and brain dynamics during navigation in the real world. *Neuroimage* 31:1826–1840.
- Strange BA, Dolan RJ. 2001. Adaptive anterior hippocampal responses to oddball stimuli. *Hippocampus* 11:690–698.
- Terrazas A, McNaughton BL. 2000. Brain growth and the cognitive map. *Proc Natl Acad Sci USA* 97:4414–4416.
- Terrazas A, Krause M, Lipa P, Gothard KM, Barnes CA, McNaughton BL. 2005. Self-motion and the hippocampal spatial metric. *J Neurosci* 25:8085–8096.
- Tomaiuolo F, Carlesimo GA, Di PM, Petrides M, Fera F, Bonanni R, Formisano R, Pasqualetti P, Caltagirone C. 2004. Gross morphology and morphometric sequelae in the hippocampus, fornix, and corpus callosum of patients with severe non-missile traumatic brain injury without macroscopically detectable lesions: A T1 weighted MRI study. *J Neurol Neurosurg Psychiatry* 75:1314–1322.
- Ugurbil K, Garwood M, Ellermann J, Hendrich K, Hinke R, Hu X, Kim S-G, Menon R, Merkle H, Ogawa S, Salmi R. 1993. Imaging at high magnetic fields: Initial experiences at 4 T. *Magn Reson Q* 9:259–277.
- Voermans NC, Petersson KM, Daudey L, Weber B, Van Spaendonck KP, Kremer HP, Fernandez G. 2004. Interaction between the human hippocampus and the caudate nucleus during route recognition. *Neuron* 43:427–435.
- Volman SE, Grubb TC Jr, Schuett KC. 1997. Relative hippocampal volume in relation to food-storing behavior in four species of woodpeckers. *Brain Behav Evol* 49:110–120.
- Warrington EK. 1984. *Recognition Memory Test*. England, UK: Nfer-Nelson.
- Weschler D. 1987. *Weschler Memory Scale-Revised*. San Antonio: The Psychological Corporation.
- Weschler D. 1998. *Weschler Adult Intelligence Scale-III UK*. London, UK: The Psychological Corporation.
- Weschler D. 1999. *Weschler Abbreviated Scale of Intelligence*. San Antonio: The Psychological Corporation.
- White NM, McDonald RJ. 2002. Multiple parallel memory systems in the brain of the rat. *Neurobiol Learn Mem* 77:125–184.
- Wolbers T, Buchel C. 2005. Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *J Neurosci* 25:3333–3340.