HIPPOCAMPUS 18:981–984 (2008)

RAPID COMMUNICATION

Non-Spatial Expertise and Hippocampal Gray Matter Volume in Humans

Katherine Woollett, Janice Glensman, and Eleanor A. Maguire*

ABSTRACT: Previous work suggests that spatial expertise in licensed London taxi drivers is associated with differences in hippocampal gray matter volume relative to IQ-matched control subjects. Here we examined whether non-spatial expertise is associated with similar hippocampal gray matter effects. We compared medical doctors who, like taxi drivers, acquire a vast amount of knowledge over many years, with IQ-matched control subjects who had no tertiary education. Whole brain analysis of structural magnetic resonance imaging (MRI) scans using voxel-based morphometry (VBM) failed to identify any differences in gray matter volume between the groups, including in the hippocampus. Moreover, amount of medical experience that ranged from 0.5 to 22.5 yr did not correlate with gray matter volume in the hippocampus or elsewhere in the brain. We conclude that intensively acquiring a large amount of knowledge over many years is not invariably associated with hippocampal gray matter volume differences. Instead it would seem that hippocampal gray matter volume effects are more likely to be observed when the knowledge acquired concerns a complex and detailed large-scale spatial layout. © 2008 Wiley-Liss, Inc.

KEY WORDS: hippocampus; space; taxi drivers; VBM; memory

INTRODUCTION

In the UK, licensed London taxi drivers undergo extensive training over a period of 2–4 yr. This involves learning the layout of 25,000 streets, thousands of places of interest, leading to a stringent set of examinations in order to obtain an operating license. Gray matter volume differences in the hippocampus have been reported to accompany this expertise, with greater volume in posterior hippocampi and less volume in anterior hippocampi compared with control subjects. Furthermore, gray matter volume has been found to correlate with years of navigation experience in London taxi drivers, but not in London bus drivers, who experience a restricted set of routes though the city (Maguire et al., 2000, 2006).

Although gray matter volume differences have been described in cortical areas for other groups with expertise, such as musicians, bilinguals, and mathematicians (e.g., Gaser and Schlaug, 2003; Mechelli et al., 2004; Aydin et al., 2007), 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 practice 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 (Maguire et al., 2006). 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 gray matter volume patterns previously observed.

In this study we sought to examine one of these factors, namely whether the hippocampal gray 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 gray matter volume. Terrazas and McNaughton (2000) questioned whether other groups such as medical doctors or lawyers who also have high memory demands placed upon them might also show similar gray matter volume patterns as London taxi drivers. Their question is reasonable given the intimate link between memory and the hippocampus (Andersen et al., 2007), and long-standing debates about whether or not the hippocampus is preferentially engaged by spatial stimuli (O'Keefe and Nadel, 1978; Eichenbaum et al., 1999; Kumaran and Maguire, 2005; Moscovitch et al., 2006; Spiers and Maguire, 2007).

To examine this issue, we 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. We focused 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 it has been suggested that the knowledge they acquire is less spatial than that of taxi drivers (Terrazas and McNaughton, 2000). Gray matter volume has been examined previously in medical students where knowledge acquisition occurred over a 3-month period (Draganski et al., 2006). However, this short time

Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London, United Kingdom
Katherine Woollett and Janice Glensman contributed equally to this work.
Grant sponsor: Wellcome Trust.
*Correspondence to: Eleanor A. Maguire, Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London, UK. E-mail: e.maguire@fil.ion.ucl.ac.uk
Accepted for publication 14 May 2008
DOI 10.1002/hipo.20465
Published online 19 June 2008 in Wiley InterScience (www.interscience.wiley.com).

TABLE 1.

Background Details and Mean Test Scores

	Medical doctors mean (SD)	Control subjects mean (SD)
Participant characteristics		
Age	31.48 (5.42)	34.13 (8.01)
Matrix reasoning (scaled score—WASI)	14.04 (1.40)	13.65 (1.23)
Edinburgh Handedness Inventory	70.90 (34.25)	83.43 (16.76)
Years medical experience	12.07 (5.43)	–
Mean memory scores		
Digit span (scaled score—WAIS-III)	12.96 (3.32)	14.30 (4.06)
Rey Osterreith complex figure copy	35.52 (1.37)	35.43 (1.08)
Rey Osterreith complex figure delayed recall	24.54 (4.92)	23.02 (5.09)
Logical memory immediate recall (percentile score—WMS-R)	86.87 (13.31)	87.57 (15.54)
Logical memory delayed recall (percentile score—WMS-R)	81.04 (15.58)	87.00 (15.85)
Warrington Recognition Memory Test for Faces (scaled score)	12.35 (2.36)	11.09 (2.79)
Warrington Recognition Memory Test for Words (scaled score)	12.80 (0.38)	12.83 (1.15)
Mean stress scores		
Perceived Stress Scale	12.91 (6.09)	13.39 (7.05)
State Trait Anxiety Inventory (State)	31.43 (7.86)	29.96 (8.86)
State Trait Anxiety Inventory (Trait)	37.74 (9.13)	36.39 (11.66)
Rating of life stress (/10)*	5.04 (2.05)	4.96 (2.09)
Rating of job Stress (/10)*	5.00 (1.93)	4.65 (2.01)

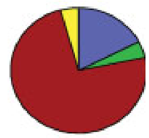
WASI, Wechsler Abbreviated Scale of Intelligence; WAIS-III, Wechsler Adult Intelligence Scale III; WMS-R, Wechsler Memory Scale Revised.
*Ratings 1 (low stress); 10 (high stress).

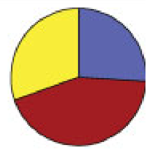
scale is not comparable with that of taxi drivers, and thus our interest was in testing medical doctors whose experience ranged over much longer timescales. In this context, the relationship between medical knowledge acquisition and gray matter volume is unknown.

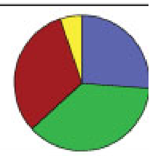
In group comparison studies, it is important that groups are matched on key variables such as gender, age, and IQ. IQ scores, for example, have been reported to correlate with hippocampal volume in healthy adults (Amat et al., 2008). Although previous work has shown taxi drivers and their control subjects have average IQs (Maguire et al., 2006), 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 socioeconomic background. Instead, we sought control participants 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 because of disparity on basic factors such as intellectual ability.

Methods and data analyses were identical to those described in a previous study (see Maguire et al., 2006 for full details). Forty-six healthy male volunteers participated in this 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 1. The medical doctors varied in degree of medical experience ranging from 0.5 to 22.5 yr, with a mean of 12.07 yr [standard deviation (SD) 5.43]. Interestingly, this is comparable with the previous study of licensed London taxi drivers (mean years experience 10.94 yr (SD 5.25), Maguire et al., 2006). 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, we 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. (2006), we assessed visual information processing and abstract reasoning skills using the Matrix Reasoning subtest of the Wechsler Abbreviated Scale of Intelligence (Weschler, 1999). The mean-scaled score for both groups was above average, and the groups did not differ significantly ($t(44){-}1.00$; $P = 0.31$). All participants gave informed written consent to participate in the study in accordance with the local research ethics committee.

Rating	Percentage
Is the medical knowledge you acquire mostly:	
Verbal	17.4
Visual images	4.3
A mixture	73.9
Other	4.3
Is the medical knowledge you acquire*:	
Mostly separate facts but with some degree of connectedness	26.1
The knowledge elements are well connected	43.5
The knowledge elements are all highly inter-connected	30.4
When would you say the amount of medical knowledge stored in your memory was greatest?	
Just after primary medical degree	26.3
After postgraduate qualification	36.8
It continues to accrue with increasing experience	31.6
Other	5.3







Percentage refers to the percentage of the group of doctors who chose this rating. Note that for the third rating (*When...greatest?*), medical students were excluded as it was not possible for them to answer. *No doctor selected the first option of: 'A large collection of separate bits of information'. Similarly 'Other' was not selected by any doctor for this rating and so is not included.

FIGURE 1. Knowledge in medical doctors. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Our goal in testing medical doctors was to assess a group of subjects whose acquired knowledge was less visual/spatial than that of taxi drivers. Ratings (see Fig. 1) 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 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, (see Maguire et al., 2006) 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 our expectation that medical doctors are an appropriate group in which to examine the relationship between hippocampal gray matter volume and expertise that is similar but less visual/spatial than that of London taxi drivers.

We first compared medical doctors and control subjects in terms of their neuropsychological profiles. A test battery was employed to assess a range of basic cognitive, memory, and affective functions. Full details of the tests are provided in Maguire et al. (2006) and mean scores are shown in Table 1. A multivariate analysis of variance (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, Rey Complex Figure copy, Rey Complex Figure delayed recall, logical memory immediate recall, logical memory delayed recall, word recognition, and face recognition. No significant differences were found between the groups ($F(7,38)1.64$; $P = 0.16$). An additional MANOVA was performed using five stress/anxiety measures: Perceived Stress Scale, State Trait Anxiety Inventory (state), State Trait Anxiety Inventory (trait), life stress rating, and job stress rating. No significant differences were apparent between the groups ($F(5,38)0.34$; $P = 0.89$).

Having found no differences between the medical doctors and control subjects on background characteristics, neuropsychological performance or stress measures, we then compared the magnetic resonance imaging (MRI) brain scans of the two groups to assess gray matter volume throughout the brain. The MRI scanner, scanning parameters, and data analysis were identical to those used by Maguire et al. (2006). To summarize, whole brain structural MRI scans were acquired on a 1.5 T Sonata scanner (Siemens Medical Systems, Erlangen, Germany), with a whole-body coil for RF transmission and an eight-element phased-array head coil for signal reception. Data analyses were performed using voxel-based morphometry (VBM) implemented in SPM5 using a smoothing kernel of 10 mm (Ashburner and Friston, 2005). 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. No significant differences in gray matter volume were found between the two groups. To limit the chances of false-negatives, a number of other analyses were undertaken. The above analysis was repeated with significance levels at $P < 0.005$ uncorrected for multiple comparisons. Even at this liberal threshold no significant difference in hippocampal gray matter volume was observed between the groups. Although a smoothing kernel of $\sim$10 mm is normally employed in SPM-VBM, it might be suggested that a smaller smoothing kernel is more appropriate for the hippocampus, such as the 4-mm kernel used by Maguire et al. (2000) and Gadian et al. (2000). Thus images were smoothed at both 8 mm and at 4 mm and the analyses described above were repeated. No significant differences were found between the groups.

We next examined if any of the neuropsychological or stress scores correlated with gray matter volume anywhere in the brain. We found no significant correlations for either group. Finally, we assessed whether gray 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 gray matter volume were apparent even at the liberal threshold of $P < 0.005$ uncorrected for multiple comparisons.

Naturally caution is required in interpreting the absence of positive findings. We 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. We therefore consider some implications of the latter possibility as suggested by the current findings.

The main difference between our two groups was that one group of subjects acquired a large amount of knowledge over many years, whereas the other group did not. Despite this difference, the groups were comparable in terms of gray matter volume in the hippocampi and indeed elsewhere in the brain. Furthermore, years of medical experience did not correlate with gray matter volume. The question, therefore, is why does spatial expertise and years navigating in London by taxi drivers correlate with gray matter volume (Maguire et al., 2000, 2006), whereas medical expertise and experience does not. It has been suggested that the hippocampus is particularly specialized for supporting allocentric spatial representations of the type that permits flexible navigation in large-scale space (O'Keefe and Dostrovsky, 1971; O'Keefe and Nadel, 1978). In non-humans the volume of the hippocampus is known to vary as a function of the demands placed on spatial memory (see Lee et al., 1998). The lack of positive findings in medical doctors lends further support to the view that the hippocampus may be preferentially engaged by spatial memory.

Although the non-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. Although medical doctors also acquire vast amounts of knowledge elements, and rated them as well or highly connected, it may be they are better characterized 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 characterized 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 us to conclude that intensively acquiring a large amount of knowledge over many years is not invariably associated with hippocampal gray matter volume differences relative to matched control subjects. 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.

Acknowledgments

We are grateful to all the volunteers for their participation. We thank David Bradbury, Peter Aston, and John Ashburner for their assistance and advice, and Peter Baimbridge and British Mensa for facilitating recruitment of control volunteers.

REFERENCES

Amat JA, Bansal R, Whiteman R, Haggerty R, Royal J, Peterson BS. 2008. Correlates of intellectual ability with morphology of the hippocampus and amygdala in healthy adults. Brain Cogn 66:105–114.

Andersen P, Morris RGM, Amaral DG, Bliss TVP, O'Keefe J, editors. 2007. The Hippocampus Book. New York: Oxford University Press.

Ashburner J, Friston KJ. 2005. Unified segmentation. Neuroimage 26:839–851.

Aydin K, Ucar A, Oguz K, Okur O, Agayev A, Unal Z, Yilmaz S, Ozturk C. 2007. Increased gray matter density in the parietal cortex of mathematicians: A voxel-based morphometry study. Am J Neuroradiol 28:1859–1864.

Draganski B, Gaser C, Kempermann G, Georg Kuhn H, Winkler J, Büchel C, May A. 2006. Temporal and spatial dynamics of brain structure changes during extensive learning. J Neurosci 26:6314–6317.

Eichenbaum H, Dudchenko P, Wood E, Shapiro M, Tanila H. 1999. The hippocampus, memory, and place cells: Is it spatial memory or a memory space? Neuron 23:209–226.

Gadian D, Aicardi J, Watkins K, Porter D, Mishkin M, Vargha-Khadem F. 2000. Developmental amnesia associated with early hypoxic-ischaemic injury. Brain 123:499–507.

Gaser C, Schlaug G. 2003. Gray matter differences between musicians and nonmusicians. Ann N Y Acad Sci 999:514–517.

Kumaran D, Maguire EA. 2005. The human hippocampus: Cognitive maps or relational memory? J Neurosci 25:7254–7259.

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.

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, Woollett K, Spiers HJ. 2006. London taxi drivers and bus drivers: A structural MRI, neuropsychological analysis. Hippocampus 16:1091–1101.

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.

Moscovitch M, Nadel L, Winocur G, Gilboa A, Rosenbaum RS. 2006. The cognitive neuroscience of remote episodic, semantic and spatial memory. Curr Opin Neurobiol 16:179–190.

O'Keefe J, Dostrovsky J. 1971. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res 34:171–175.

O'Keefe J, Nadel L. 1978. The Hippocampus as a Cognitive Map. Oxford: Oxford University Press.

Spiers HJ, Maguire EA. 2007. The neuroscience of remote spatial memory: A tale of two cities. Neuroscience 149:7–27.

Terrazas A, McNaughton BL. 2000. Brain growth and the cognitive map. Proc Natl Acad Sci USA 97:4414–4416.

Weschler D. 1999. Weschler Abbreviated Scale of Intelligence. San Antonio, USA: The Psychological Corporation.