Three-dimensional space representation in the human brain

Misun Kim

Submitted for a PhD in Cognitive Neuroscience

May 2018

Supervisor: Eleanor A. Maguire

I, Misun Kim, 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

Brain structures that support spatial cognition by encoding one's position and direction have been extensively studied for decades. In the majority of studies, neural substrates have been investigated on a horizontal two-dimensional plane, whereas humans and other animals also move vertically in a three-dimensional (3D) world.

In this thesis, I investigated how 3D spatial information is represented in the human brain using functional MRI experiments and custom-built 3D virtual environments. In the first experiment, participants moved on flat, tilted-up or tilted-down pathways in a 3D lattice structure. Multivoxel pattern analysis revealed that the anterior hippocampus expressed 3D location information that was similarly sensitive to the vertical and horizontal axes. The retrosplenial cortex and posterior hippocampus represented direction information that was only sensitive to the vertical axis. In the second experiment, participants moved in a virtual building with multiple levels and rooms. Using an fMRI repetition suppression analysis, I observed a hierarchical representation of this 3D space, with anterior hippocampus representing local information within a room, while retrosplenial cortex, parahippocampal cortex and posterior hippocampus represented room information within the wider building. As in the first experiment, vertical and horizontal location information was similarly encoded. In the last experiments, participants were placed into a virtual zero-gravity environment where they could move freely along all 3 axes. The thalamus and subiculum expressed horizontal heading information, whereas retrosplenial cortex showed dominant encoding of vertical heading. Using novel fMRI analyses, I also found preliminary evidence of a 3D grid code in the entorhinal cortex.

Overall, these experiments demonstrate the capacity of the human brain to implement a flexible and efficient representation of 3D space. The work in this thesis will, I hope,

serve as a stepping-stone in our understanding of how we navigate in the real – 3D – world.

Impact statement

Our world is three-dimensional (3D). We navigate in multi-level buildings, on undulating terrain and in the open air. However, the neural circuitry underlying spatial cognition (our “internal GPS system”) has mainly been studied on horizontal two-dimensional (2D) surfaces. The experiments I report here extend our understanding of spatial cognition by incorporating the uncharted third dimension. I used custom-built virtual reality navigation paradigms and a non-invasive brain scanning technique (functional magnetic resonance imaging - fMRI) to test where, and how, 3D location and direction information is encoded in the human brain. The main findings were: (1) the hippocampus contained information about where participants were in 3D environments and it was similarly sensitive to vertical and horizontal axes; (2) 3D heading was collectively encoded by the thalamus, subiculum and retrosplenial cortex, and these brain areas were sensitive to either horizontal or vertical direction; (3) the brain seemed to have flexible representations of 3D space, adapted for the shape of environments and behavioural demands.

My work will hopefully promote further research into 3D spatial cognition in both humans and other animals, as we still know so little about this crucial aspect of cognition. For example, the retrosplenial cortex, which represented vertical direction information in my fMRI studies, is a candidate brain structure to contain head direction cells tuned to vertical pitch, and this could be tested in animal electrophysiological studies. Taking the lead from my work, which revealed the basic neural circuitry for 3D spatial representations, future research should seek to elucidate how we actively navigate and find goals in a realistic 3D world. This might help us to develop better navigation strategies. I also showed the advantage of using virtual reality techniques and a quasi-naturalistic setup, and I believe similar approaches should be deployed more widely in the field of spatial research.

The impact of my research also extends beyond the domain of spatial navigation. Recent findings suggest that the brain uses a map-like code to solve more general cognitive problems, such as social interactions, perception and decision making. The analysis methods I developed – including probing 3D grid codes using fMRI – could be applied to broader cognitive problems which are inherently high-dimensional.

In the longer term, studying 3D spatial representation might bring clinical benefits. The hippocampus, entorhinal cortex and retrosplenial cortex – the main brain structures investigated in my work – are early casualties in Alzheimer's disease, and spatial disorientation is a common initial symptom. The prognostic value of 2D grid signals in the brain has been previously suggested, but spatial navigation tasks in 3D might be more sensitive in aiding early detection of impending cognitive decline.

It may sound far-fetched, but understanding the neural basis of our sense of direction in 3D could be pertinent when humans are freed from Earth's gravity. Space exploration is an ever-increasing endeavour (e.g. missions to Mars by NASA and commercial companies). An understanding of 3D space representations in the brain will be crucial to facilitating the ergonomic design of microgravity environments during space travel.

Acknowledgements

I first want to give my enormous thanks to my primary supervisor, Eleanor Maguire. She was extremely generous with her time, teaching me how to do science, how to write, how to talk and, most importantly, how to keep my chin up when science occasionally tested me. I have always been supported and trusted by her, and this is by far the best thing any supervisor can do. I will miss our regular Tuesday meetings.

I am grateful to my funders, the Wellcome Trust ("for the incurably curious") and Samsung Scholarship. I also want to thank my ~120 participants who bore my experiments without VR sickness or claustrophobia (my apologies to the few people who were afflicted).

I was so lucky to get my PhD training at the Wellcome Centre for Human Neuroimaging (aka the "FIL"). All of the staff (Kamlyn, Peter, David, IT, the radiographers, etc.) provided awesome support. The scientific advice was excellent (Karl, the guru, who patiently reminded me of Neyman Pearson lemma, Guillaume who helped me to excavate SPM, Pete Z, Gareth, Martina, Tim Behrens). I can't imagine my PhD without the Maguire gang – Anna, Conny, Daniel, Dina, Gloria, Ian, Marshall + honorary member Swan – who have always lifted me up scientifically and in many other ways. Thanks to my amazing office mates who rescued me from my 2D desk space to the outside 3D world like Nitzan, Thomas, Philipp, Claudia, Mona, Sofie, Alexa, Pradeep, Michael, etc.

I also want to say thanks to Alasdair Gibb and David Attwell – the Wellcome 4 year PhD committee – who brought me to UCL and kindly looked after me in my first year and onwards. Special thanks to Fabio and Elina in my PhD program. Thanks also to my secondary supervisor, Kate Jeffery, who sowed the seed of 3D space research.

Last, but not least, I am eternally grateful to my beloved family in Korea.

항상 믿어주고 기도해주는 엄마, 아빠, 형주, 민서야 고마워.

Table of contents

List of figures

List of tables

List of abbreviations

2D: two-dimensional

3D: three-dimensional

BOLD: blood oxygenation level dependent

fMRI: functional magnetic resonance imaging

GLM: general linear model

MEG: magnetoencephalography

MRI: magnetic resonance imaging

PET: positron emission tomography

ROI: region of interest

SD: standard deviation

SEM: standard error of mean

SPM: statistical parametric mapping

VR: virtual reality

Chapter 1 General introduction

1.1 Overview

Spatial navigation is an essential skill for all animals (e.g. bees finding honey, rats exploring a myriad of underground sewers, humans commuting to work, travellers in a busy city or on exotic islands). During navigation, we calculate an efficient route to goal destinations and constantly monitor our location and direction within the environment. Nowadays, humans often rely on a mobile phone map and GPS technology, but our brains are equipped with a biological compass and map system. Decades of extensive research in animal electrophysiology, lesion studies, human neuroimaging and computational neuroscience have revealed the basic neural circuitry underlying spatial cognition. The hippocampus, entorhinal cortex, thalamus and retrosplenial cortex are some of the key neural structures for building an internal representation of space by encoding position, head direction and distance information. Considering the practicalities of conducting experiments, it is understandable that most of this previous work has been done using simplified laboratory setups which were on a two-dimensional (2D) plane.

Yet humans and other animals spend much of their time in a more complex three-dimensional (3D) world. We move vertically as well as horizontally on slopes, in multi-level buildings or in volumetric 3D space. Three axes (x, y and z) are required to specify one's location in 3D space and at least two angles (azimuth, horizontal tilt; pitch, vertical tilt) are required to represent one's facing direction (note that one can also rotate side-to-side which is called a roll rotation) (Figure 1.1). Despite the inherent need for navigation in 3D space, there is a dearth of knowledge concerning the neural substrates of 3D spatial processing. The neural representation of 3D space might not be a simple extension of 2D space representation because of gravity (Jeffery et al.

2013, 2015). Gravity exerts effects on every living organism on earth and in so doing distinguishes one axis (the vertical) from other two axes (the horizontal) in 3D space. The vertical dimension might be differentially encoded from the horizontal dimensions, in terms of encoding precision or neural substrates. Alternatively, the brain might have an elegant mechanism for dealing with space that is independent of dimensionality (i.e. generalisable across 1D to 3D, or beyond). These mechanisms might be general for all species or species-specific.

The key aim of my PhD was to leverage our understanding of how the human brain represents the external world and allows us to navigate within it by incorporating the uncharted third dimension. Investigating how the brain deals with 3D space is important for understanding spatial cognition per se, but it could also provide new insights more generally in terms of how the brain has adapted to operate in a complex world.

In this first chapter, I summarise what is currently known about the neural building blocks of a 3D spatial map in the brain. I begin with the literature on spatial cells – place cells, head direction cells, grid cells and other spatially modulated cells – which have been mostly studied in animals. I then present behavioural findings and human neuroimaging studies on 3D space. Next, I discuss the anatomy of the hippocampal formation. Finally, I provide a brief overview of this thesis.

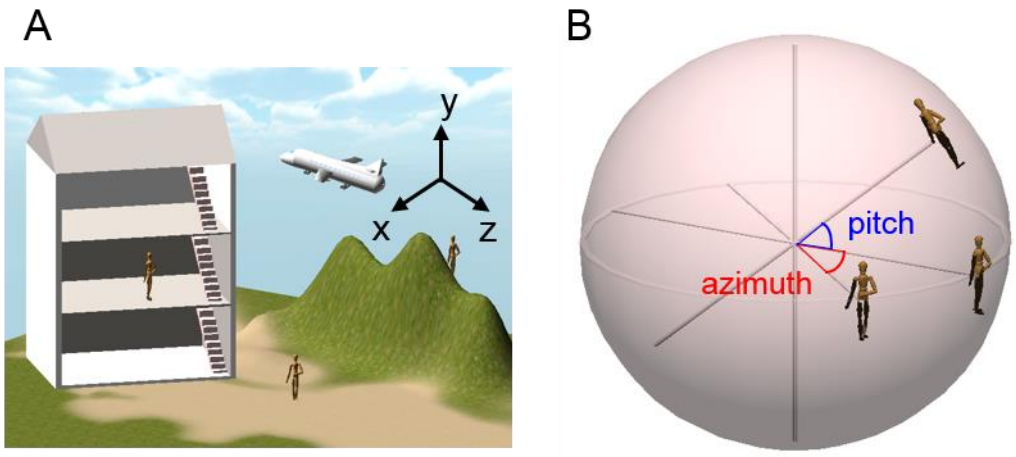


Figure 1.1 Examples of 3D space. (A) Humans and other animals operate within 3D spaces that include slopes, multi-level buildings and a voluminous sky. One's position is represented by 3 coordinates (x, y and z). (B) Orientation in 3D space is specified by the azimuth (horizontal tilt) and the pitch (vertical tilt).

1.2 Place cells in 3D space

The seminal discovery of place cells in the rat hippocampus in 1971 (O'Keefe and Dostrovsky 1971) promoted numerous empirical and theoretical studies on the internal map of the environment in the following decades, and was awarded a Nobel Prize in 2014 (for a historical review see Moser et al. (2017)). A place cell in the hippocampus fires when an animal is at one or a few unique regions of an environment called place fields (Figure 1.2B). Place cells fire at specific locations irrespective of where the animal is facing (except when animals run on a 1D trajectory repeatedly (McNaughton et al. 1983; Muller et al. 1994)) and they continue to fire even in darkness (Quirk et al. 1990). Consequently, researchers have concluded that place cells respond to abstract spatial information about where an animal is located, rather than a particular view or sensory snapshot. Different place cells fire at different locations, so an entire environment and an animal's location within it can be represented by the collective activity of place cells. Therefore, O'Keefe and Nadel suggested that place cells are the neural correlates of the "cognitive map" proposed earlier by Tolman (Tolman 1948; O'Keefe and Nadel 1978). Place cells have mostly been recorded in rodents foraging in a 2D horizontal arena (Figure 1.2A), but they have also been reported in rodents during

virtual navigation (Aronov and Tank 2014), in monkeys during physical translocation (Matsumura et al. 1999) and in humans during virtual navigation (Jacobs et al. 2010).

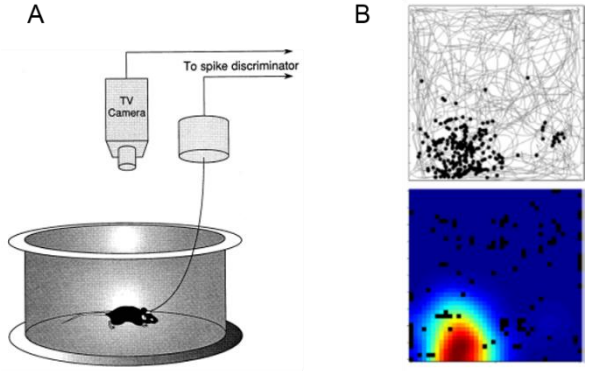


Figure 1.2 An example place cell recorded in 2D. (A) Left, place cells have typically been recorded in the rat hippocampus when animals forage on a 2D flat surface (such as a circular or rectangular arena). Reprinted from (Rotenberg et al. 1996) with permission from Elsevier. (B) Top panel, a rat's trajectory is shown as grey lines. Spikes are shown as black dots. Bottom panel, a colour-coded firing map. Red represents the high firing rate of a place cell. Reprinted from (Moser et al. 2017) with permission from Springer Nature.

In the abovementioned studies, place cells expressed information about an animal's location on a horizontal flat surface (e.g. an X-Y plane). Do place cells also encode the Z coordinate if an animal's location is not restricted to a horizontal X-Y surface? Or do they only encode horizontal location regardless of an animal's vertical coordinate? The earliest clue for answering these questions came from an experiment where rodents moved on a tiltable surface (Knierim and McNaughton 2001) (Figure 1.3A). In this study, it was asked whether a place cell's firing was bound to a 2D surface regardless of an absolute 3D coordinate within the environment. The result was not consistent with either pure 2D surface tuning or pure 3D tuning. If a place cell is perfectly bound to the local surface, place fields should be identical for both the flat and the tilted condition. Only a third of the cells showed similar place fields (e.g. cells 5 and 6 in Figure 1.3A). On the other hand, if an absolute 3D location is encoded, cells with place fields at one end of the track (which did not change its 3D location after tilting, shaded blue in Figure 1.3A) should show a higher correlation between the flat and tilted conditions compared to the cells with place fields at other end of the track (which was raised from the ground after tilting, shaded red in Figure 1.3A). However, these two groups of cells did not

show significant differences in place field correlations, contradicting a pure 3D encoding hypothesis.

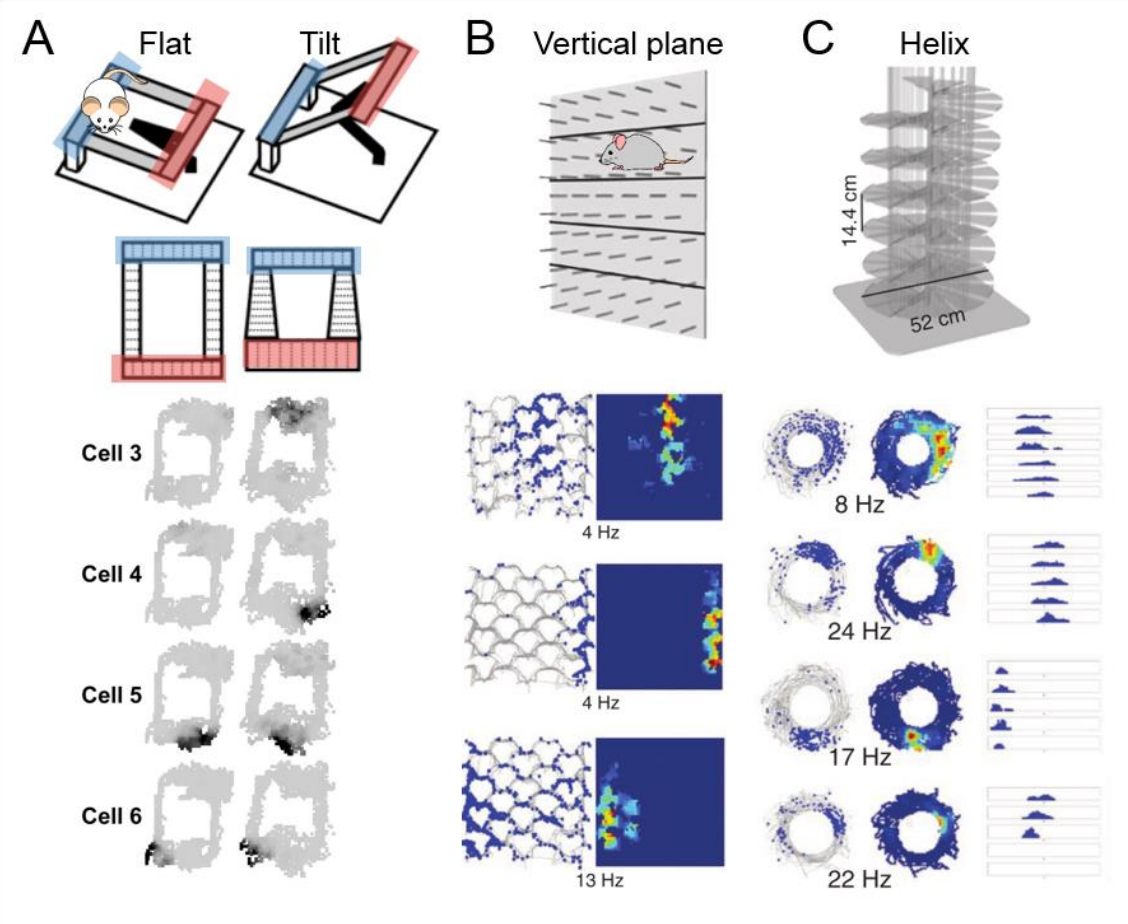


Figure 1.3 Place cells recorded on a tilted plane or on a spiral staircase. (A) Top panel, a flat rectangle maze and a tilted maze (45°). Middle panel, views of the flat and tilted maze from an overhead camera. Bottom panel, example place cells recorded in the flat and tilted conditions. Some cells (cells 5 and 6) maintained their firing fields but other place cells (cells 3 and 4) remapped and fired in either the flat or the tilted maze only. Adapted from (Knierim and McNaughton 2001). (B,C) Top panels, rats moved on a vertical pegboard or a helix structure. Bottom panels, most place cells showed vertically elongated and repeated place fields across each coil, suggesting the anisotropic encoding of 3D space. Adapted from (Hayman et al. 2011) with permission from Springer Nature.

In Hayman et al. (2011), place cells were recorded when rats foraged on a large vertical plane (a “pegboard”) or a helical staircase (Figure 1.3B,C). Most place cells showed vertically elongated place fields on the pegboard, and the place fields were repeated across each coil in the helical staircase, although there was modulation in firing rates across the coils. Spatial information encoded by each place cell was significantly smaller for the vertical axis compared to the horizontal axis (note that spatial information collectively encoded by a population of neurons might be different

(Keinath et al. 2014)). These findings suggested that place cells in rats are less sensitive to the vertical dimension and it was proposed that 3D space is anisotropically encoded, in a semi-planar manner, as most mammalian animals dwell on the surface (Jeffery et al. 2013).

In contrast to the vertically elongated place fields found in rats (Hayman et al. 2011), place cells recorded in flying bats showed isotropic place fields. This means that all three axes were equally well encoded in the bats (Yartsev and Ulanovsky 2013) (Figure 1.4). Do these findings imply that bats and rodents have fundamentally different spatial encoding mechanisms, and that bats have a complete 3D spatial map while rodents do not? It is possible that flying mammals and surface dwelling mammals have evolved differently because of differing behavioural demands. Theta oscillations in the hippocampus, which are thought to be important for place cell firing in rodents, were found to be absent in bats (Yartsev and Ulanovsky 2013). However, it is too early to conclude that rodents and other land animals do not have a 3D map.

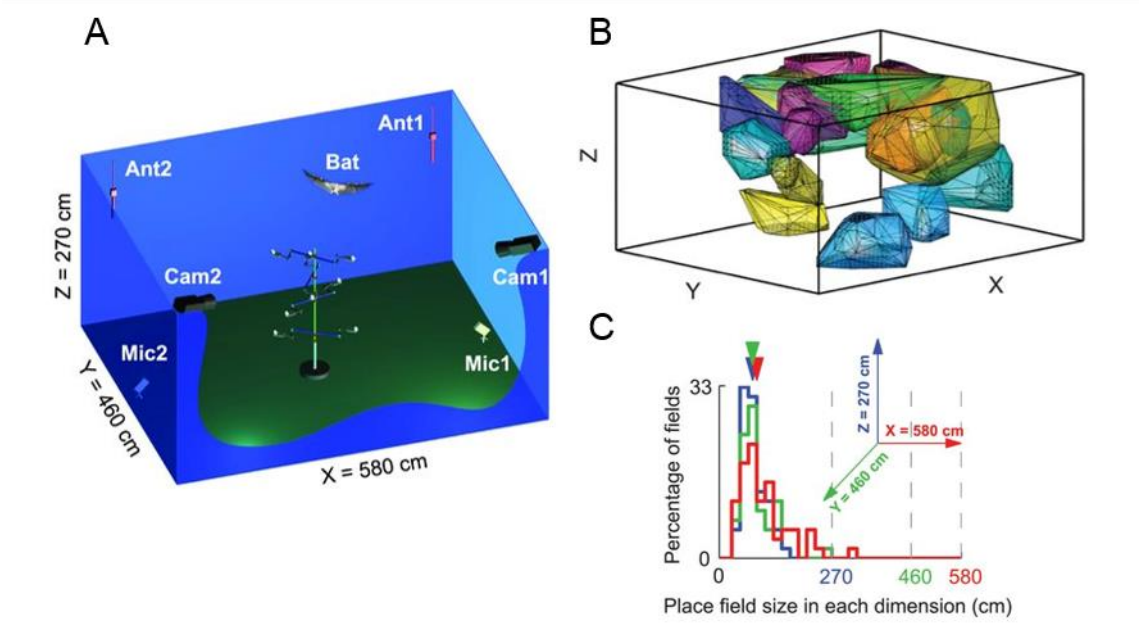


Figure 1.4 Isotropic 3D place fields in bats. (A) Place cells were recorded wirelessly in flying bats. (B) Ten place cells recorded in one bat. Different coloured blobs represent different place cells. (C) The place field size in each dimension was similar, suggesting an isotropic encoding of 3D space. Adapted from (Yartsev and Ulanovsky 2013) with permission from AAAS.

In all of the previous studies, rodents moved mainly on a surface (tilted or a vertically repeated surface), and the behavioural demand for encoding absolute 3D locations was absent. Rodents are excellent climbers (e.g. laboratory mice often hang upside down on the ceiling of a cage and wild mice swiftly climb up the pipes of the building) and they can remember the locations of food in a 3D lattice maze (Grobéty and Schenk 1992a; Flores-Abreu et al. 2014). A 3D lattice maze allows rodents to move along all 3 axes, and place cells might be equally sensitive to all three axes in this case, similar to bats (Figure 1.5A). The main methodological issue with recording place cells in such a 3D lattice is the tangling of the recording wires, and this could be the reason why only behavioural studies have been published on the 3D lattice to date. A wireless recording technique, which is rarely used currently, could circumvent this methodological problem.

Another experimental setting where pure 3D spatial mapping can be studied is a microgravity environment. One intriguing study was conducted in the Neurolab Space Shuttle mission of 1998 (Knierim et al. 2000). In the spaceship, rats propelled themselves forward by grasping the track. The track was an Escher staircase where the rats came back to the origin only after 3 right turns (Figure 1.5B). On the second day of visiting to this 3D track, rat place cells showed stable and normal firing fields. This finding suggests that even the rat hippocampus can rapidly adapt to and map a 3D environment that they never encountered before.

In humans and non-human primates, invasive recording studies have not yet been conducted during vertical movement in either physical or virtual 3D space.

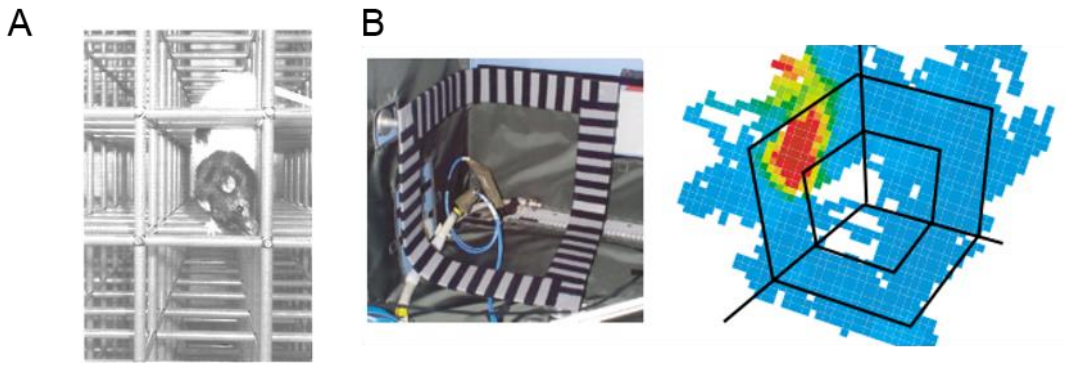


Figure 1.5 Rats in a volumetric 3D space. (A) A 3D lattice maze. Reprinted from (Grobéty and Schenk 1992a) with permission from Elsevier. (B) Rats explored an Escher staircase track during the Neurolab Space Shuttle mission. Place cells showed normal place fields on the second day of exposure to this track. Reprinted from (Knierim et al. 2000) with permission from Springer Nature.

1.3 Head direction cells in 3D space

Place cells maintain their firing even when animals move around in the environment deprived of external sensory cues, implying that an animal can know its location from self-motion inputs. If one knows the exact linear and angular movement velocity, the final destination can be calculated. For example, if I am now at the origin (x=0, y=0) and move to +x direction at a constant speed of 2 unit/s, I will be at (x=10, y=0) after 5 seconds. This is known as path integration or dead reckoning navigation. James Ranck first reported the cells in presubiculum that could provide direction information for path integration, at the annual meeting of the Society for Neuroscience (1984) and the full paper describing the basic property of these head direction cells was published in 1990 (Taube et al. 1990). As with place cells, head direction cells were first recorded in rats foraging in a circular 2D arena. This initial study revealed that the firing of head direction cells was largely independent of the animal's location and depended only on the animal's head direction (azimuth) relative to a reference direction in the laboratory (Figure 1.6).

Subsequently, multiple brain structures have been found to contain head direction cells (the lateral mammillary nucleus, anterior dorsal thalamus, presubiculum, retrosplenial

cortex, entorhinal cortex) and angular head velocity cells (dorsal tegmental nucleus, vestibular nucleus, lateral mammillary nucleus). The origin of the head direction signal has been linked to optic flow, self-movement cues, vestibular inputs and visual landmarks, while theoretical models of the head direction signal (e.g. a ring-attractor network) have also been investigated (for a review, see (Taube 2007; Cullen and Taube 2017)). However, only a handful of studies have investigated head direction cells beyond a horizontal 2D plane, and there is a dearth of knowledge about how vertical and horizontal directions are encoded in a 3D environment.

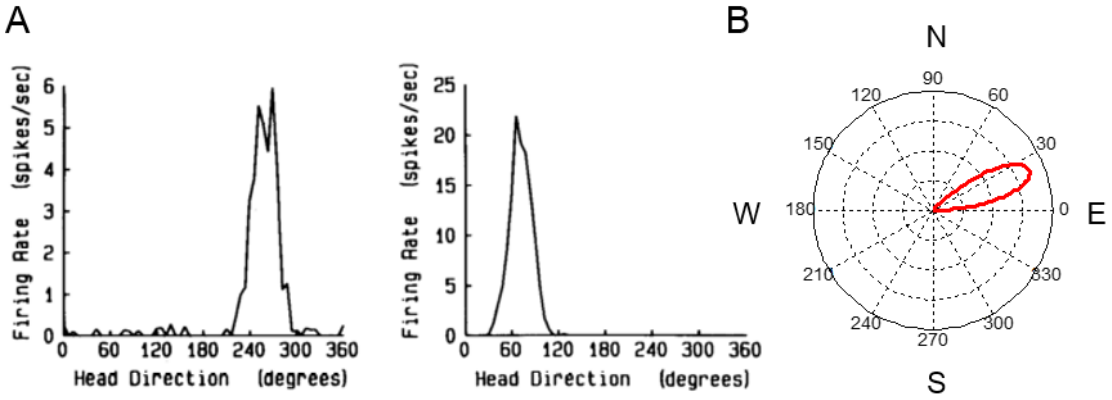


Figure 1.6 Example head direction cells in 2D. (A) One cell fired maximally when an animal was heading ~260° and another cell fired maximally when an animal was facing ~70°. Reprinted from (Taube et al. 1990) with permission conveyed through the Copyright Clearance Center, Inc. (B) A head direction cell's response can be summarized in a polar plot. This hypothetical cell's preferred direction is ~25°.

Before I summarise the limited literature on head direction cells in 3D, it is pertinent to first describe the physiological mechanisms of detecting vertical direction. Vestibular organs in the inner ear sense both vertical and horizontal tilt. The vestibular system is composed of three semicircular canals and two otolith organs called the utricle and saccule (Figure 1.7A). The three semicircular canals are fluid-filled organs that are orthogonally oriented to each other. The canals can detect the angular acceleration for yaw, roll and pitch rotations (Figure 1.7B) via the deflection of hair cells within them. The otolith organs contain small piles of calcium carbonate crystals (called otoliths) on its membrane which adds mass and increases inertia. Hair cells in the otolith organs detect linear acceleration and gravity. Due to the orientation of the utricle and saccule

within the vestibular apparatus, the utricle is more sensitive to acceleration in a horizontal plane and the saccule is more sensitive to acceleration in a vertical plane. Rotational and translational information is sent to vestibular nuclei in the brain stem. It is then conveyed to the head direction cell network via the lateral mammillary nucleus. Figure 1.8 shows a circuit diagram of key areas containing head direction cells and other spatial cells.

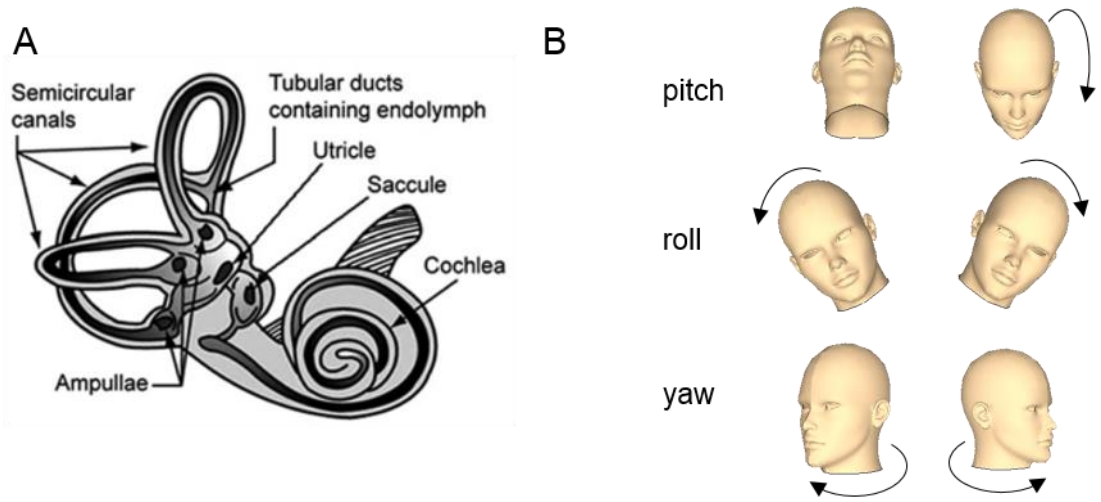


Figure 1.7 Vestibular organs and three rotations. (A) The vestibular system in the inner ear contains three orthogonally placed semicircular canals which detect angular acceleration in three directions, yaw, roll and pitch, and two otolith organs called the utricle and saccule. This figure was adapted from https://www.nasa.gov/audience/forstudents/9-12/features/F_Human_Vestibular_System_in_Space.html. (B) The three axes of rotation.

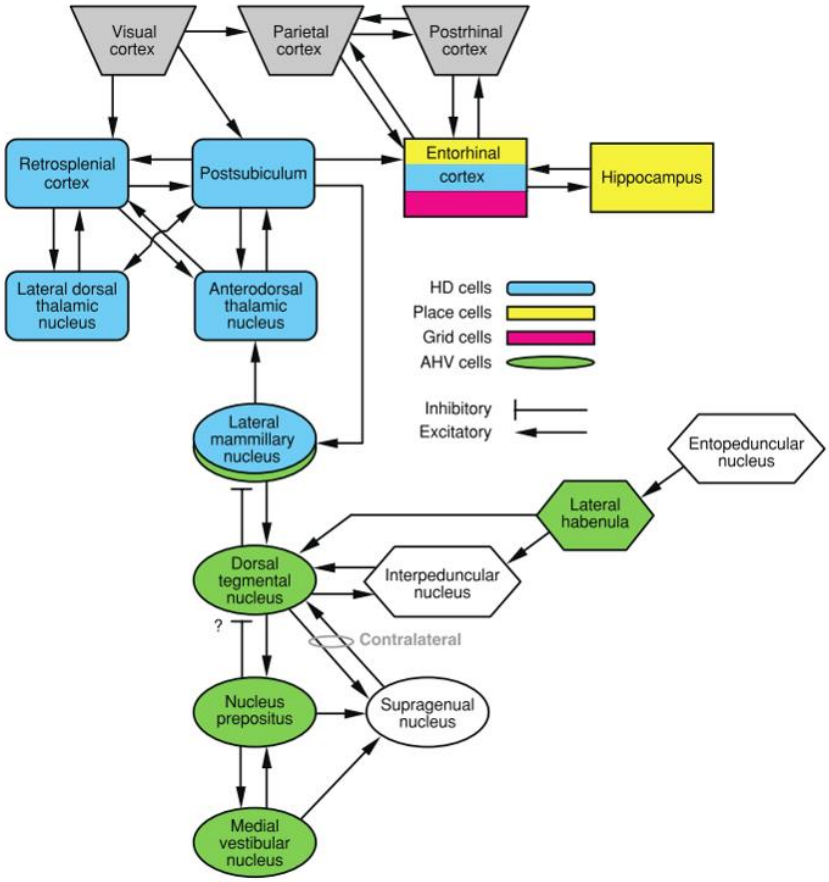


Figure 1.8 A circuit diagram of the principal connections of areas in the head direction cell network. Relevant brain regions are shown in blue. Reproduced from (Taube 2007) with permission conveyed through the Copyright Clearance Center, Inc.

In addition to vestibular inputs from the inner ear, graviceptors in the trunk provide additional tilt information which could be particularly important for body posture stabilisation (Mittelstaedt 1998). It has been suggested that the differences in otolith (in the inner ear) and abdominal viscera graviceptor dynamics might contribute to motion sickness (von Gierke and Parker 1994). Vision is another important source of verticality detection, which is particularly relevant for the virtual navigation paradigms used in this thesis. Visual objects like a tree or a tower are oriented vertically on earth, and visual inputs can provide salient cues for verticality even in the absence of physical gravity (e.g. viewing a scene photograph in a supine position) (Dyde et al. 2006).

The key question here is whether there are head direction cells that respond to 3D direction. The rodent literature does not provide reliable evidence for the existence of 3D head direction cells. An early study observed a few vertical pitch-sensitive cells in the lateral mammillary nuclei (Stackman and Taube 1998), but most pitch-sensitive cells responded only when a rat was looking up almost 90° and they were not modulated by azimuth. Cells responding to azimuth (classical head direction cells) were not modulated by pitch. It is possible that pitch and azimuth are encoded by different cells and these pieces of information are integrated in unknown downstream neurons. However, the absence of cells tuned to an intermediate vertical pitch suggests that these cells in the lateral mammillary nuclei cannot map a complete set of directions in 3D. This might be explained by the fact that the rats only moved on a flat surface and there was no behavioural need to encode vertical direction in this experiment. Azimuth encoding helps a rat to know its location on the surface for path integration, and one's location information is a fundamental element of learning and memory. In contrast, vertical encoding was unnecessary in the experiment because rats did not jump vertically or grasp anything in the air. Rats have eyes with such a wide field-of-view (Wallace et al. 2013) that they can even look up to the ceiling without tilting their head up. As long as animals are confined to a surface, direction information that is orthogonal to the surface is not so relevant.

A few studies have recorded head direction cells when rats were climbing a vertical plane or ceiling (Taube et al. 2004; Calton and Taube 2005; Taube, Wang, et al. 2013). These studies found that head direction cells responded to an animal's direction relative to the local plane of locomotion as if the vertical plane was an extension of the horizontal floor when the animals self-locomoted from the floor to the vertical walls. A dual-axis rotation rule suggests that head direction cells are updated by (1) yaw rotations around the dorsal-ventral axis of the animal and (2) rotations of the dorsal-ventral axis around the gravity axis (Page et al. 2017). Head direction cells following

the extended floor or dual-axis rule are useful when animals move between multiple planes, as they preserve global north-south orientation (Page et al. 2017). However, they do not distinguish direction when moving straight north (zero pitch) and moving 30° up north, for example. Animals moving along the two directions would arrive at very different locations if they could move freely in a 3D volumetric space. Therefore, this type of head direction cell is not suitable as a neural compass in 3D volumetric space.

A recent breakthrough in the understanding of head direction encoding in 3D came from the Ulanovsky group (Finkelstein et al. 2015). Head direction cells were recorded in the presubiculum of Egyptian fruit bats in multiple environments: a horizontal 2D plane, a vertical ring platform and a 3D arena (Figure 1.9A,E). When bats crawled on a plane, a large portion of cells (52%) was sensitive to azimuth and some were tuned to pitch (21%) or roll (12%). The recordings in the vertical ring platform allowed even sampling of pitch from -180° to 180° and cells tuned to various pitches were found, in contrast to the previously-mentioned rat study where only extreme pitch-sensitive cells were observed (Stackman and Taube 1998). About 20% of cells recorded in the crawling bats were sensitive to the combination of azimuth, pitch and roll (Figure 1.9B). The small portion of conjunctive head direction cells implies that 3D direction is collectively represented by a population of neurons rather than at the single cell level.

This study also revealed an interesting anatomical gradient in that pure azimuth cells were more abundant in the anterolateral part of presubiculum, whereas pure pitch and conjunctive cells were more numerous in the posteromedial part of presubiculum (Figure 1.9C,D). The most intriguing aspect of this study was the flight condition inside a 3D arena. Similar to the crawling condition, pure azimuth, pure pitch and conjunctive neurons were observed (Figure 1.9F), although the proportion of these cell types were not reported (the small number of cells recorded in the flight condition (n=20) might have hampered a comparison between the flight and crawling conditions). Altogether,

this study provided strong evidence that 3D direction information is present in the bat, which could be used to generate a mental map of 3D space. Is the 3D compass unique to flying animals? Probably not. Pitch and roll sensitive neurons were recently found in the macaque monkey anterior thalamus (Laurens et al. 2016). Whether a 3D compass exists in the human brain is a key question addressed in this thesis.

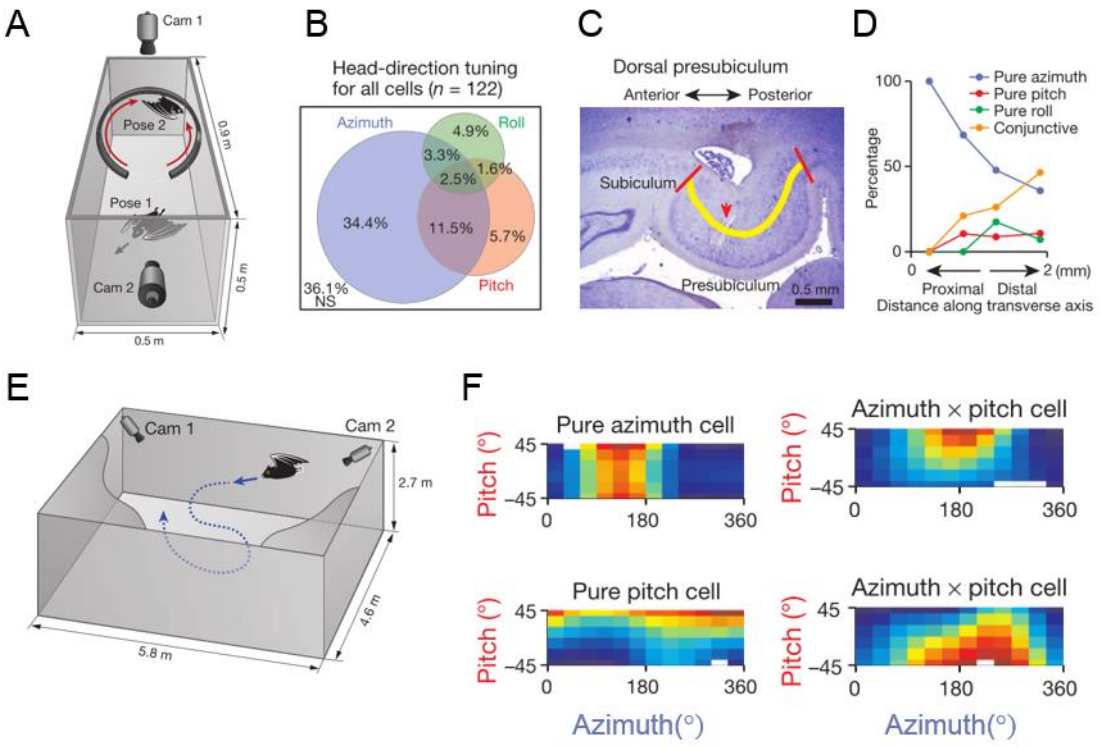


Figure 1.9 3D head direction cells in the bat presubiculum. (A) The bats crawled on a floor (Pose 1) or on a vertical ring (Pose 2). (B) The proportion of cells tuned to azimuth, pitch and roll. (C) A sagittal section through the presubiculum. (D) An anatomical gradient of head direction cells along the transverse axis of the presubiculum. (E) The flight condition setup. (F) Example head direction cells during flight. Similar to the crawling conditions, pure azimuth cells, pure pitch cells and conjunctive cells were found. Adapted from (Finkelstein et al. 2015) with permission from Springer Nature.

1.4 Grid cells in 3D space

Until 1990s, the primary region of interest in spatial neuroscience was the hippocampus. However, this changed dramatically in 2004 when the Moser group investigated dorsal medial entorhinal cortex (Figure 1.10A) (Fyhn et al. 2004). This region provides extensive inputs to the dorsal hippocampus where the most sharply

tuned place cells are found. They first observed that cells in the entorhinal cortex had stable and multiple discrete firing fields that could predict the rat's location as precisely as place cells. It turned out that the multiple firing locations are not random; rather the locations correspond to the vertex of a regular hexagonal grid, like a honeycomb structure (Figure 1.10B) (Hafting et al. 2005). This is why these cells are called grid cells. Due to the regularity of the grids, a grid cell's firing locations (called grid fields) can be described by three parameters: phase (the location or offset of grid fields), scale (the distance between the neighbouring grid fields) and orientation (the tilt of the grid axis relative to the environment) (Figure 1.10C). Grid cells maintain their firing fields regardless of an animal's movement speed and trajectory (note that direction x grid conjunctive cells were also found in the entorhinal cortex and this has important implications for probing grid cells non-invasively in humans, which will be discussed later in this thesis). That grid cell firing persists in darkness, at least in rats, suggests the importance of self-motion and path integration in this spatial representation (Hafting et al. 2005; Chen et al. 2016).

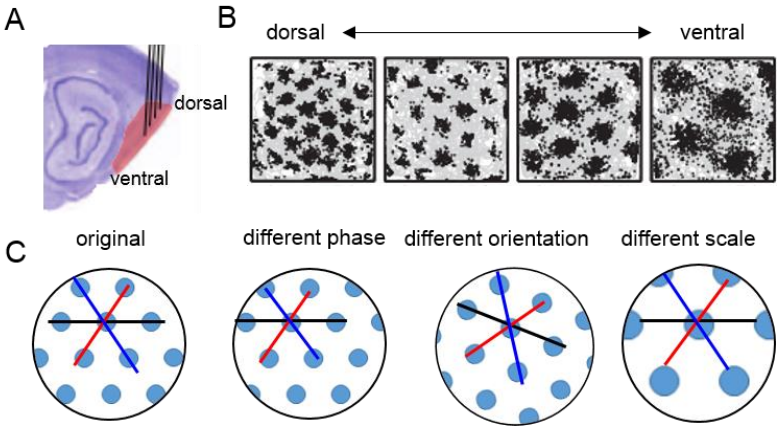


Figure 1.10 Grid cells in 2D. (A) Medial entorhinal cortex (shaded in red). (B) Example grid cells along the dorsal-ventral axis of entorhinal cortex. The grid scale is larger in the ventral entorhinal cortex. Adapted from (Stensola et al. 2012) with permission from Springer Nature. (C) Hexagonal grid fields can be described by three parameters: phase, orientation and scale.

Interestingly, grid cells are topographically and modularly organised along the dorsal-ventral axis of entorhinal cortex. Similar to place cells in the hippocampus which have

larger place fields in the ventral hippocampus (Jung et al. 1994; Kjelstrup et al. 2008), grid cells in the ventral part of entorhinal cortex have a larger grid scale than grid cells in the dorsal part of entorhinal cortex (Hafting et al. 2005) (Figure 1.10B). This difference in grid scale parallels the difference in hyperpolarization-activated cation current along the dorsal-vertical axis (Giocomo and Hasselmo 2008). An oscillatory interference model of grid cells proposed that the ionic current and resultant membrane potential oscillation is the origin of the grid signal at the single cell level (Burgess et al. 2007).

A grid scale changes discretely rather than continuously; for example, the distribution of grid scales recorded in a single rat had four peaks (modules) and the ratio of grid scales between successive modules was around $\sqrt{2}$, which implies the doubling of an area of grid fields (Stensola et al. 2012). Grid cells within a module share similar grid orientation and only differ in grid phase. Grid cells of multiple scales can serve as an efficient way of encoding a large space with high precision using a small number of cells. It can be understood as a combinatorial or modulo code (Fiete et al. 2008).

As a toy example, 8 unique locations in a simple 1D space can be encoded using only 3 grid cells that have different periodicity, if a downstream neuron can combine inputs from these grid modules (Figure 1.11A). Therefore, it seemed reasonable to assign grid cells as the origin of hippocampal place fields (Solstad et al. 2006) (Figure 1.11B). However, it was later discovered that new place fields were formed after septal inactivation which disrupts grid networks (Brandon et al. 2014), and place cells showed adult-like stable firing before grid cells showed mature firing during development (Wills and Cacucci 2014). A theoretical study showed that the characteristic hexagonal firing fields can be formed from place cell inputs in the neural network architecture that resembles a principal component analysis (Dordek et al. 2016). These findings imply that the relationship between the entorhinal grid cells and the hippocampal place cells

is not a simple unidirectional one, rather it involves a complex interaction between them.

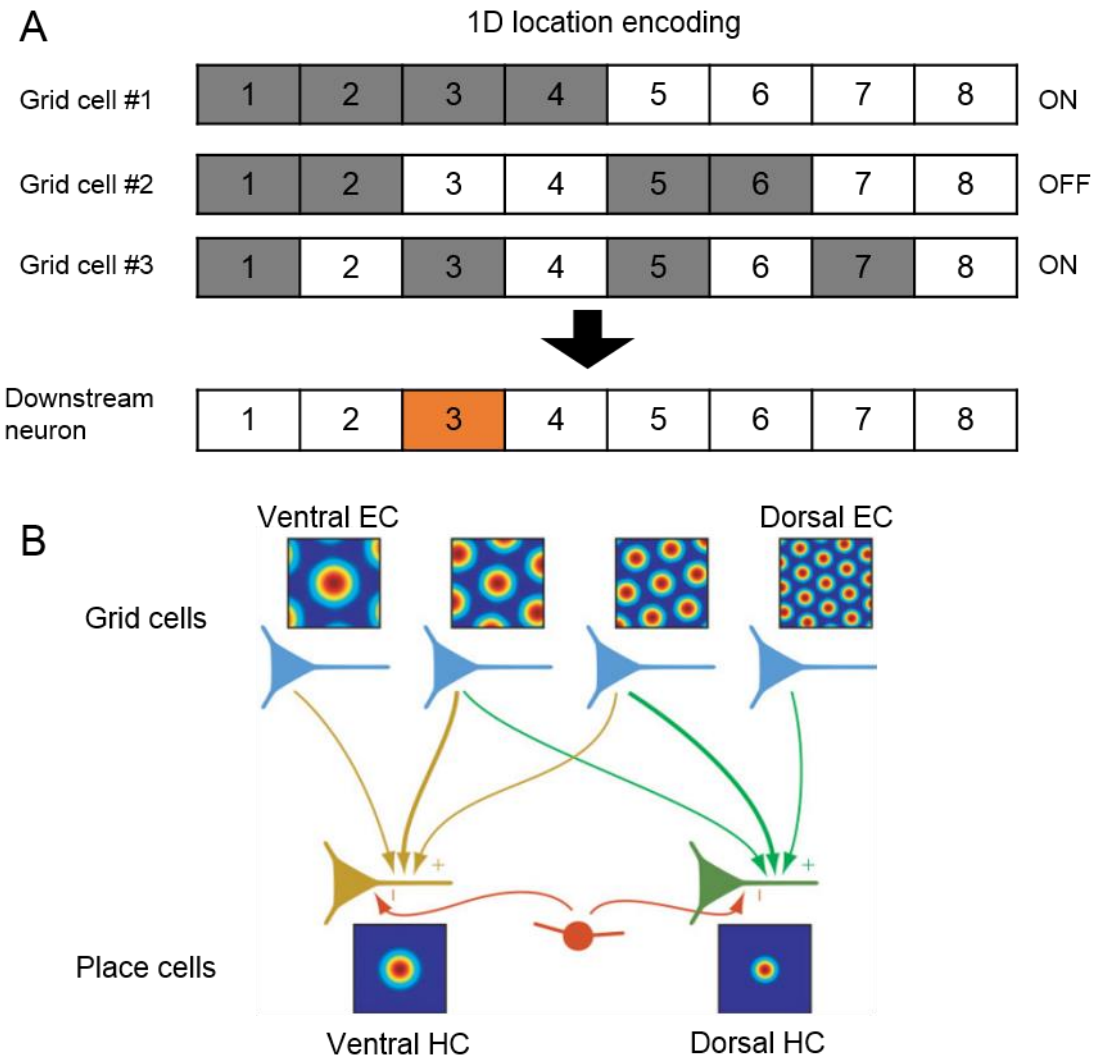


Figure 1.11 From grid cells to place cells. (A) A schematic of the combinatorial code for encoding a 1D space. Hypothetical grid cell #1 has a periodicity of 8 and it fires when an animal is at locations 1,2,3 and 4. Grid cell #2 has a periodicity of 4 and it fires at locations 1,2,5 and 6. Grid cell #3 has a periodicity of 2 and it fires at locations1,3,5 and 7. If grid cells #1 and #3 are "on" and grid cell #2 is "off", a downstream neuron can know that an animal is at location 3. (B) A proposed model for place-field formation. Ventral hippocampal place cells receive stronger inputs from grid cells in ventral entorhinal cortex which has a larger grid scale, and dorsal hippocampal place cells receive stronger inputs from grid cells in dorsal entorhinal cortex which has a smaller grid scale. Overall firing rate is kept at a physiological level by non-specific inhibitory neurons (red arrows and circle). EC, entorhinal cortex; HC, hippocampus. Adapted from (Solstad et al. 2006) with permission from John Wiley and Sons.

Like place cells and head direction cells, grid cells have mostly been investigated on a 2D horizontal plane. If grid cells are to provide an efficient spatial basis set for encoding 3D space for the hippocampus (or grid cells are the product of 3D place cell inputs), what would the firing fields of grid cells look like in 3D space? We have some clues

from an information theoretical account of grid cells (Mathis et al. 2015) and a simulation study that used a self-organizing model (Stella and Treves 2015). It was shown that the lattice with the highest packing density has the maximum Fisher information (Mathis et al. 2015). Fisher information measures how well stimuli (like the location of an animal) are decodable from the conditional firing probability of neurons. On a 2D surface, a hexagonal arrangement of circles has the highest density (compared to a rectangular array or any other regular geometric lattice). In a 3D volumetric space, the highest packing density is achieved when grid fields are arranged in a face-centred cubic (FCC) lattice, hexagonal close packing (HCP) or some combination of these two hexagonal layers (Figure 1.12).

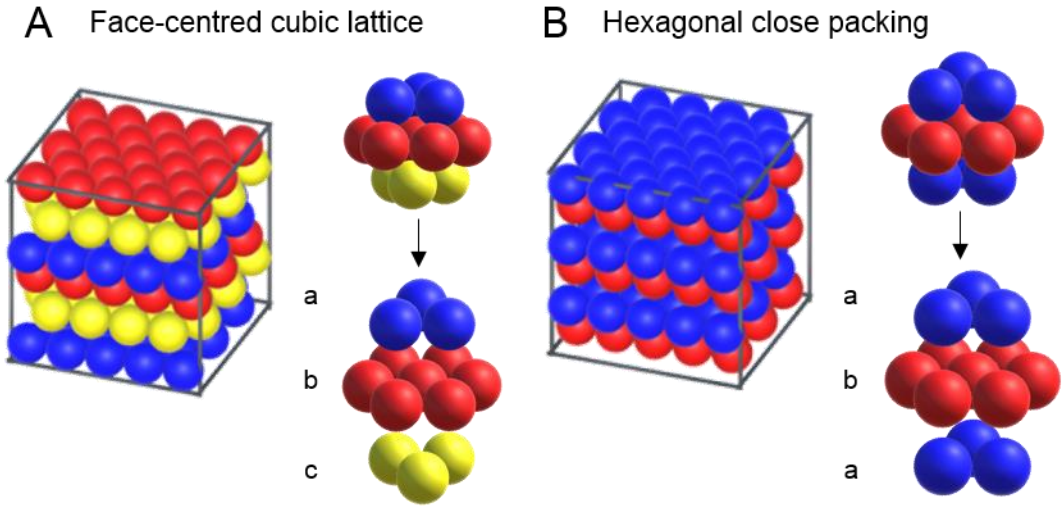


Figure 1.12 A 3D grid cell is proposed to fire at multiple locations which corresponds to the centre of spheres closely fitted in a box. (A) An FCC arrangement is composed of three hexagonal layers with a translational shift (a-b-c). (B) HCP is also composed of repeating hexagonal layers, but only two layers repeat (a-b). In both arrangements, a unit cell is composed of 1+12 fields.

Hayman et al. (2011) tested for the existence of 3D grid cells in rodents when exploring various 3D environments. On both a vertical surface ("pegboard") and a helical staircase, grid cells did not show periodic firing along the vertical dimension (Figure 1.13A,C). Grid cells only showed multiple peaks along the horizontal axis and grid fields were vertically elongated (on the pegboard) or repeated across coils (on the

helical environment), although there was some rate modulation by vertical height. This result does not fit the predicted firing pattern of grid cells in 3D described earlier (Figure 1.12; Figure 1.13B). The authors discussed the possibility that grid cells might have much larger vertical spacing than the size of the environment used in the experiment, but it is unclear whether and why grid cells should have elongated vertical spacing compared to horizontal spacing. A grid cell's firing in the helical environment resembled the repeated and fragmented grid fields that have been observed in a multi-compartment environment (Derdikman et al. 2009).

Grid cells were also recorded on a 40° slope (Hayman et al. 2015) (Figure 1.13D-F). If a grid cell has firing fields corresponding to a 3D lattice covering the entire volumetric space, grid fields should be irregular and the distance between the grid fields should be larger on the slope compared to a flat surface (Figure 1.13E). However, a grid cell's firing was almost as similar on the slope as on a flat surface. Together with the findings from the helical environment (Hayman et al. 2011), this result implies that grid cells are adapted to the local surface on which rats move, rather than developing a fully 3D map. Thus, we might infer that a surface, regardless of whether it is horizontal, tilted or stacked, is not a suitable environment where 3D grid cells can be investigated. The best animal model for testing 3D grid cells is flying bats. Some preliminary evidence of 3D grid cells in bats was presented at the Society for Neuroscience (Ginosar et al. 2016), but the grid cell showed rather irregular firing fields which were unlike the regular 3D lattice predicted by the theoretical models. Thus, the existence of a regular 3D lattice remains an important unresolved question which will be addressed later in this thesis.

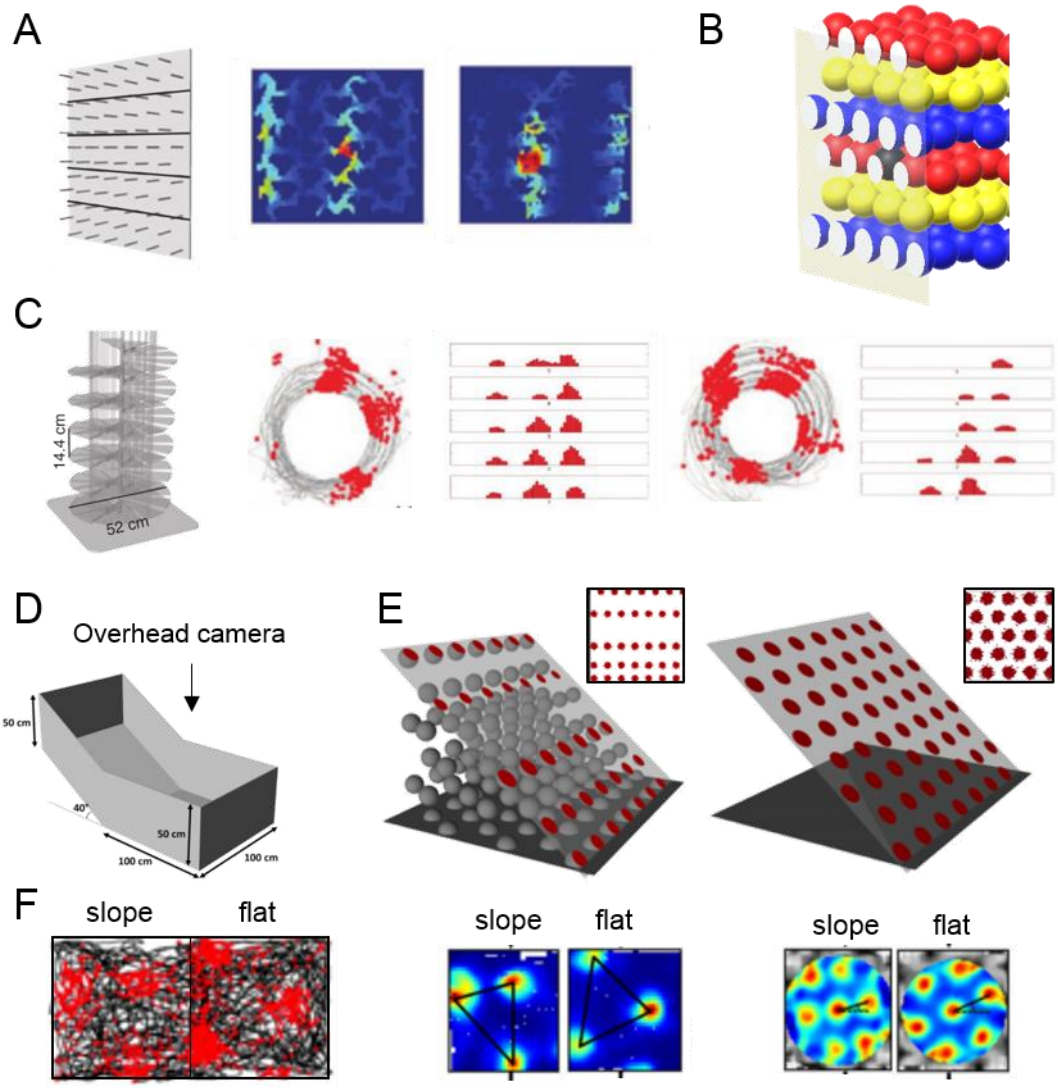


Figure 1.13 Grid cells recorded on a non-horizontal surface. (A) Two example grid cells recorded in rats moving on a vertical surface are shown in the contour plots. The grid cells did not show periodic firing for the vertical axis in contrast to the prediction from 3D grid cell models (e.g. a face-centred cubic lattice) shown in (B). (C) Rats moved on a helical staircase. The firing of two example grid cells is shown as raw data from an overhead camera and separately for each coil. Grid cell firing was repeated across multiple vertical levels (coils). (D) Rats moved on a flat surface and an adjacent slope (40°). (E) Two hypotheses of grid cells, one is a lattice permeating the 3D space (left) and the other is a planar grid on the environment surface (right). (F) An example grid cell's response shown from an overhead camera. Left panel, raw data; middle panel, a contour map; right panel, an autocorreleogram. In contrast to the 3D lattice hypothesis which predicts irregular, largely spaced grid fields on a slope, grid fields were regular and had similar spacing on the slope and flat surface. (A) and (C) adapted from (Hayman et al. 2011) with permission from Springer Nature. (D)-(F) adapted from (Hayman et al. 2015).

1.5 Boundary cells and other spatial cells in 3D space

In the previous sections, I summarised the properties of the three most widely studied cell types (place cells, head direction cells and grid cells) that are known to encode spatial information, and I included some predictions for how they might respond in a 3D

world. In this section, I consider briefly some other cell types with spatial properties that have been less well investigated, namely boundary cells and spatially periodic cells.

From very early studies it was known that a place cell's firing is modulated by the geometry and the boundaries of the environment an animal is exploring. For example, when a circular or rectangular environment was uniformly scaled along all axes, some place fields were scaled, preserving the relative location within the environment (Muller and Kubie 1987) (Figure 1.14A). This study also showed that even a visually transparent barrier altered nearby place fields. The importance of boundaries that determine the geometry of an environment has been supported by further studies. In (O'Keefe and Burgess 1996), place cells were recorded in a small square, large square, vertically elongated rectangle and horizontally elongated rectangular arena (Figure 1.14B). Place fields maintained the absolute distance or the proportion of the distance from two opposite walls, and the authors proposed a model in which a place field is formed by Gaussian tuning curves as a function of the distance from each wall. Putative cells that respond when an animal is at a particular distance and direction from a boundary ("boundary vector cells") were then proposed to provide inputs to place cells (Hartley et al. 2000) (Figure 1.14C). Subsequently, boundary vector cells were found in the subiculum (Lever et al. 2009) (Figure 1.14D).

Given that a majority of these boundary vector cells fired in the vicinity of the border of an environment, one might ask whether the cells respond to a particular view or the tactile sensation of walls. Lever et al. (2009) ruled out this possibility by showing that these cells fire at similar locations when different types of boundary (50 cm wall, a sheer drop or a low ridge) were used. The persistence of responses from these cells in complete darkness also suggested that the cells did not simply respond to particular visual cues. Cells with similar characteristics were subsequently found in the entorhinal

cortex (Savelli et al. 2008; Solstad et al. 2008) and parasubiculum (Boccara et al. 2010), and have been called boundary cells or border cells.

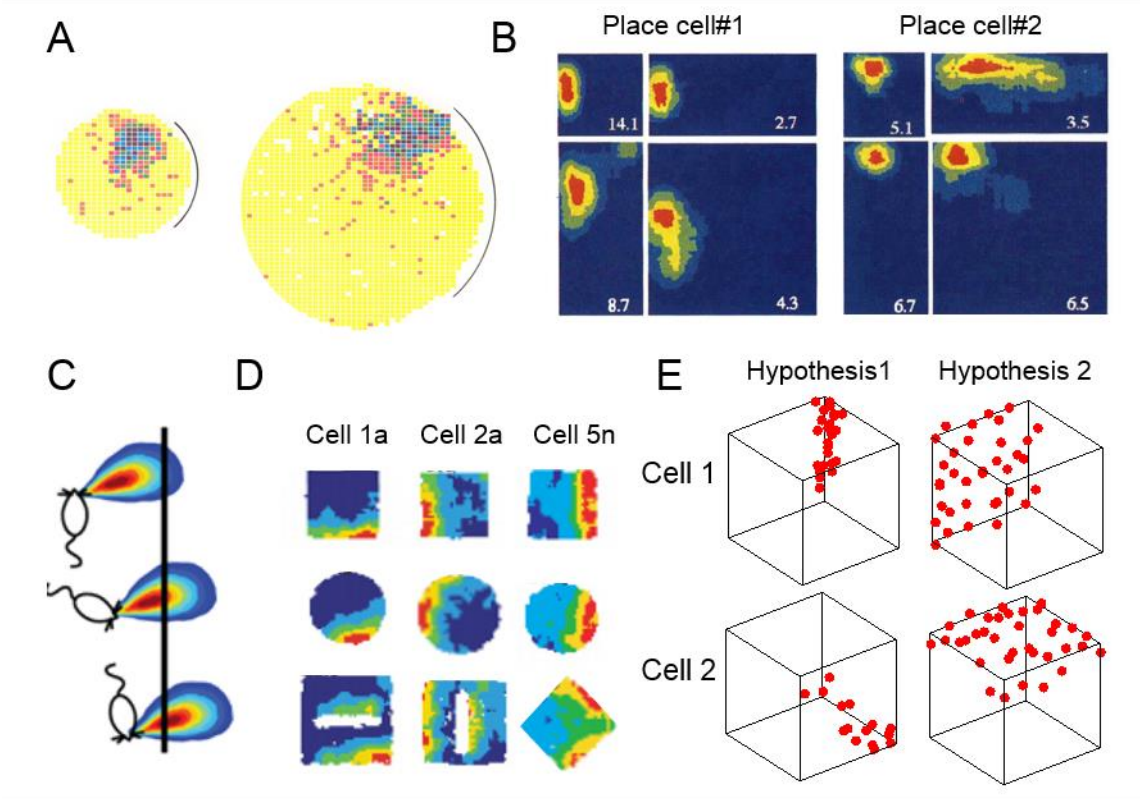


Figure 1.14 Place fields are dependent on an environmental boundary and boundary cells have been proposed to provide inputs to place cells. (A) A place cell's firing field maintained its relative location when the size of the circular arena doubled. Adapted from (Muller and Kubi 1987) with permission conveyed through the Copyright Clearance Center, Inc. (B) Place fields maintained absolute or relative distance from the walls in four different environments (small square, horizontal rectangle, vertical rectangle, large square). Adapted from (O'Keefe and Burgess 1996) with permission from Springer Nature. (C) A boundary vector cell is proposed to fire when an animal is located at a preferred distance and allocentric direction from a boundary. Adapted from (Hartley et al. 2000) with permission from John Wiley and Sons. (D) Three example boundary vector cells recorded in the rat subiculum in various environments (rectangle, circle, diamond shape and when a new barrier was inserted). Adapted from (Lever et al. 2009) with permission conveyed through the Copyright Clearance Center, Inc. (E) Hypotheses for boundary cells in 3D space. Hypothesis 1 predicts that the cells would respond to the 1D line boundary. Hypothesis 2 predicts that the cells would respond to the 2D surface boundary.

All of the experiments on boundary cells have been conducted on a flat 2D surface, which means that the boundary was one-dimensional - a flat or curved line. Would boundary cells fire at line edges (either vertical or horizontal) even when animals freely move inside a 3D box (Figure 1.14E, hypothesis 1)? I predict that boundary cells would respond to 2D surface boundaries and therefore form a plane-like firing field in 3D space (Figure 1.14E, hypothesis 2). This is because distance and direction information

from three orthogonal planes is required to locate one's 3D position. If boundary cells only encode the information about an animal's distance and direction from 1D borders, they cannot provide sufficient inputs to hippocampal place cells. We cannot even define a line boundary in a smooth 3D environment like a sphere. The best animal model to test for the existence of 3D boundary cells is flying bats. Boundary cells (responding to the line edge of the enclosure) have been found in the entorhinal cortex of bats crawling on a 2D floor (Yartsev et al. 2011), but it remains an interesting question as to whether there would be cells encoding the distance and direction from walls when bats are flying inside a 3D box.

We can make a similar prediction about the 1D line- or 2D plane-like firing fields of a cell type known as band cells or spatially periodic cells in a 3D environment. Some theoretical models suggested that the hexagonal firing pattern of grid cells originated from cells with spatially periodic cells in 1D (e.g. a stripe or band) (Hasselmo and Brandon 2012; Mhatre et al. 2012). In the rat entorhinal cortex, cells that have band-like firing fields have been observed (Krupic et al. 2012). If such band-like cells exist in a volumetric 3D space, I predict that they would look like a full planar wave, instead of 1D strips. This need to be tested in the future.

1.6 Behavioural studies in 3D space

So far, I have explained the special cell types in the mammalian brain that encode an animal's location and direction information. In this section, I summarise the behaviour of animals navigating in 3D space which can give clues about how 3D space is represented in their brains. This is particularly useful for inferring the neural representation of 3D space in humans, where neural recording data are scarce. I first describe behavioural findings in animals and then humans.

1.6.1 Fish, birds and rodents

Gravity distinguishes the vertical dimension from the horizontal dimension and all animals on earth are under the constant influence of gravity. The vertical axis can be an even more salient dimension for fish because of the hydrostatic pressure gradient. Burt de Perera's group have suggested that a fish can detect its vertical position by a change in swim-bladder volume when the fish moves vertically, and that this movement need not be large in order to be detected (Taylor et al. 2010; Holbrook and Burt de Perera 2011). They argued that a pressure change as small as 0.5 cm at the surface level can be detected (Qutob 1963). Hydrostatic pressure might be the reason why banded tetra fish *(Astyanax fasciatus)* and benthic fish *(Corydoras aeneus),* preferred a vertical cue when vertical and horizontal information were conflicted (e.g. the reward was in the up-left arm during training and the fish had to choose between the up-right and down-left arm in the test phase, Figure 1.15A) (Holbrook and Burt de Perera 2009; Davis et al. 2014).

However, this does not mean that fish are less accurate in remembering horizontal information. When fish were trained to swim freely towards a reward location, they took the shortest path, and the vertical and horizontal deviations from the ideal trajectory were comparable (Holbrook and Burt de Perera 2013) (Figure 1.15A). From these behavioural results, it was suggested that fish have an isotropic 3D spatial representation (Burt de Perera et al. 2016). Of note, spatial memory or navigation experiments in fish have been much simpler than those in rodents or humans. The fish only learned a short straight line in the experiments, and they had to be retrained multiple times during the short test phase to prevent memory extinction.

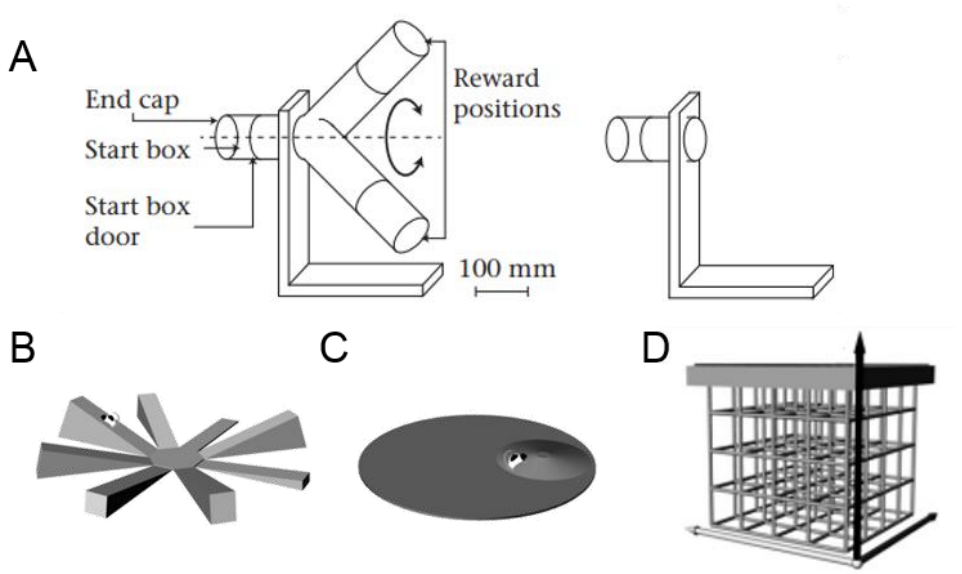


Figure 1.15 Apparatus for studying 3D navigation behaviour in animals. (A) Fish can be trained to swim inside one of the two arms to get food reward (left). The arms can be rotated so the vertical and horizontal location of the reward can be modified. Alternatively, the arms can be removed and a free 3D trajectory can be examined (right). Reproduced from (Holbrook and Burt de Perera 2013) with permission from Elsevier. (B) A radial arm maze with vertical tilts. (C) An arena with a small hill. (B) and (C) Reproduced from (Jeffery et al. 2013) with permission from Cambridge University Press (D) A 3D lattice maze for rodents and birds. Reproduced from (Jovalekic et al. 2011).

Mammals and birds do not have a sensory mechanism that can detect the absolute vertical elevation like the swim-bladder in fish (of note humans can detect high altitude from low oxygen and pressure decreases, but this is only applicable at extreme altitudes like Mountain Everest, so it is not relevant for daily navigation). However, all animals are sensitive to their orientation relative to the gravity axis for stable body posture and energy costs associated with vertical movements. Thus, a slope or vertically tilted plane can work as an additional orientation cue for navigation and spatial memory. For example, rats located a previously rewarded position better in a 45°, or 90° tilted planar maze than in a horizontal maze (Grobéty and Schenk 1992a), and performed better in a radial arm maze when some of the arms were tilted between 0 to 25° (Grobéty and Schenk 1992b) (Figure 1.15B). They also observed that the rats showed a pattern of movement that involved moving horizontally before each vertical movement in the vertical plane for the first three days of training (Grobéty and Schenk 1992a). This pattern of more frequent horizontal movements was also observed in

another study (Jovalekic et al. 2011). Animals can save energy by making this pattern of movements when exploring the vertical plane compared to moving up and down frequently.

In other experiments, rats were trained to find a reward in darkness that was located at the top of a cone hill of different heights (radius = 40cm, height = 0, 1, 2, 4 cm which corresponds to 0, 1.4, 2.9, 5.7°) within a larger circular arena (radius = 100 cm). An inverse relationship between the rats' path length and the cone height was found (Moghaddam et al. 1996) (Figure 1.15C). This means that the steeper slope was more beneficial than the shallow slope for navigation, and the rats were sensitive to subtle differences in steepness of less than a few degrees. Pigeons have also been found to use the slope (20°) of a trapezoidal arena to locate a goal position, and preferred the slope cue over a geometrical cue when the two cues conflicted (Nardi and Bingman 2009a, 2009b). Interestingly, slope preference was not affected by hippocampal lesions in pigeons (Nardi and Bingman 2009b). To the best of my knowledge, there is no hippocampal lesion study in mammals that has specifically investigated vertical or 3D navigation. From the behavioural findings of enhanced spatial memory on a slope and the high sensitivity to the steepness of a slope, I would predict that an animal's vertical tilt (pitch) on a slope is encoded in the brain at least as precisely as the horizontal direction.

In the above experiments, rodents and birds mainly navigated on a 2D surface even though a vertical element was added to the surface. The next question is, how do they navigate in a fully volumetric 3D space? Do they plan the vertical and horizontal route separately? Do they remember the vertical location better or worse than the horizontal location? There are mixed results in relation to vertical-horizontal symmetry. In (Grobéty and Schenk 1992a), rats were trained to move to a previously rewarded position in a 3D lattice (Figure 1.15D) in the following order: seven training trials with

reward (days 1-4), the first transfer test without reward (day 4), seven training trials (days 4-6) and the second transfer test (day 6). In the first transfer test, the rats spent more time on the correct vertical level than the correct horizontal locations. The opposite pattern was found in the second transfer session, which means that the rats were able to locate the horizontal location, not the vertical coordinate. The authors suggested the energy cost of vertical movement as an explanation for the earlier acquisition of vertical information. However, an energy cost account does not explain why the rats knew the horizontal coordinate better than the vertical coordinate in the second test session, and whether the rats would know the vertical or horizontal information better or to the similar degree if the test was conducted for a longer duration.

A recent study which used a similar 3D lattice setup and memory task may provide a further clue (Flores-Abreu et al. 2014). In this study, the rats chose the vertically adjacent location more than the horizontally adjacent location – this was a trend on the first test date (day 7) and significant on the second test date (day 10). This finding is in line with the result of the second test in (Grobéty and Schenk 1992a). These results suggest that rats are more confident of their horizontal location, at least after enough exposure. Interestingly, hummingbirds were also tested in a similar 3D lattice maze task in Flores-Abreu et al. (2014) and, in contrast to rats, the hummingbirds were more confident of their vertical location. Hummingbirds can also distinguish flowers by their height (Henderson et al. 2001, 2006).

In summary, differences in the navigation behaviour of fish, rats and birds highlight the importance of the mode of movement of animals (whether they are surface-bound or volumetric) for spatial cognition and navigation. The behavioural differences could originate from how the brain encodes 3D space in a species. Whether the difference is evolutionary or developmental is not easy to test, but is nevertheless an interesting

question (e.g. would a bird be more confident of its horizontal location if it was raised on a horizontal surface from birth?). In the following sections, I focus on 3D navigation behaviour in humans, the main topic of this thesis.

1.6.2 Humans

Similar to other animals, humans are also sensitive to slopes or elevations which are navigational cues. In a study by Steck et al. (2003), participants rode on a bicycle simulator and navigated a virtual reality (VR) town which contained flat or sloping (4°) streets. Navigation errors and pointing errors were significantly smaller in the sloped condition, showing the utility of a slope even though it was only 4°. The importance of vertical elevation for navigation has also been supported by a task investigating how the shortest route is found in an undulating real environment. Participants travelled longer distances in an undulated environment as they tried to avoid local hills in the travelling salesman problem (Layton et al. 2009). This natural tendency to minimise the energy cost incurred when moving vertically (climbing upward is obviously tiring and walking down can be even more tiring and dangerous with the risk of slipping and falling) might be one reason why people overestimate the angle of slopes (Proffitt et al. 1995; Creem-Regehr et al. 2004; Shaffer and Flint 2011). For example, a shallow hill on the grounds of the University of Virginia which was only 2° was perceived as 10° and a steep hill (34°) was reported as 53~55°, both verbally and visually (Proffitt et al. 1995). Of note, the participants did not greatly overestimate the steepness when they haptically reported the steepness by adjusting a tilt board with their hand.

This veridical motor judgment of slope accords with our normal walking behaviour, e.g. we do not stumble on a 30° slope by wrongly adjusting our ankles and feet for a 50° slope. This could be because of separate visual pathways for conscious perception and motor responses (Goodale and Milner 1992). The behavioural finding of the

overestimated slope raises a question of whether the neural response would be more sensitive to the vertical heading than the horizontal heading in 3D space, e.g. an abundance of pitch-sensitive head direction cells compared to the azimuth-sensitive cells in the human brain. Or, the brain might encode the vertical direction just as precisely as it encodes the horizontal direction as shown by our accurate motor responses to the slope. If there exists a neural bias in encoding the vertical and horizontal direction, is the bias dependent on the presence of an actual energy cost? For example, would the vertical direction be encoded with the same degree of sensitivity in environments where the energy cost is absent, such as a VR environment or a micro-gravity environment? The visual appearance of the slope alone might be sufficient to drive the heightened sensitivity to the vertical axis. It was shown that participants also overestimated the steepness of an escalator which does not require energy (Shaffer and Flint 2011) and hills in a virtual environment (Creem-Regehr et al. 2004).

Other than undulating terrain, the most common environment in which humans face a 3D navigation challenge is multi-level buildings. In modern buildings, from a few to over a hundred floors are stacked one on top of the other. The added vertical dimension is distinguished from horizontal dimensions not only by gravity but also by its discontinuity. Horizontal information can be described in both a continuous or discrete manner, such as “2.5 meters north from the entrance”, “the 3rd room on that floor”. In contrast, vertical information is often described in a discrete number such as “2nd floor”, “10th floor”, and vertical movement is only possible via sparsely located staircases or lifts.

Dividing a large space into smaller and more manageable regions can be an efficient navigation strategy that can overcome complexity and imprecise spatial information (Wiener et al. 2004; Balaguer et al. 2016). A multi-level building is naturally divided into

horizontal floors, and so a 3D navigation task is reduced to a 2D horizontal navigation task after an initial journey to the correct floor. It has been reported that participants who were familiar with a building, more frequently used a floor strategy than those who were unfamiliar with the building, and navigation performance was superior when the floor strategy was used compared to other navigation strategies (Hölscher et al. 2006). Difficulty with navigation across floors, or larger pointing errors for a target location on a different floor, might reflect a floor-based spatial representation (Montello and Pick 1993).

However, horizontal floors is not the only way of segmenting a 3D space; it can also be divided into vertical columns because modern buildings often have the same layout for each floor (e.g. room 301 is above room 201, room 302 is above room 202, Figure 1.16). The mental representation of, and navigation strategies within, a multi-level building can be dependent on multiple factors such as individual differences, the shape of the building and the learning experience. For instance, Büchner et al. (2007) showed that one third of participants used the horizontal layout and preferred the vertical first-horizontal later route for finding objects in a building of the University of Freiburg, whereas two thirds of participants used a columnar layout and preferred the horizontal first-vertical later route. The authors suggested that the wide building (the horizontal width was greater than the vertical height) used in the experiment might have caused more participants to choose the columnar layout. Spatial memory of multi-level buildings can also be influenced by the learning route. In Thibault et al. (2013), participants learned the layout of 9 objects in a virtual building (3 floors x 3 rooms on each floor) by watching a video that provided the sensation that they were sequentially visiting each room (Figure 1.16). Half of the participants visited the rooms along the floor route and the others visited the rooms along a columnar route, and then spatial memory was tested using both horizontal and columnar pathways in each participant. Overall, the floor learners were more accurate than the columnar learners and there

was an interaction effect of learning and testing of route. The floor learners performed better on the horizontal test trials and the columnar learners performed better on the vertical test trials, suggesting the importance of learning method.

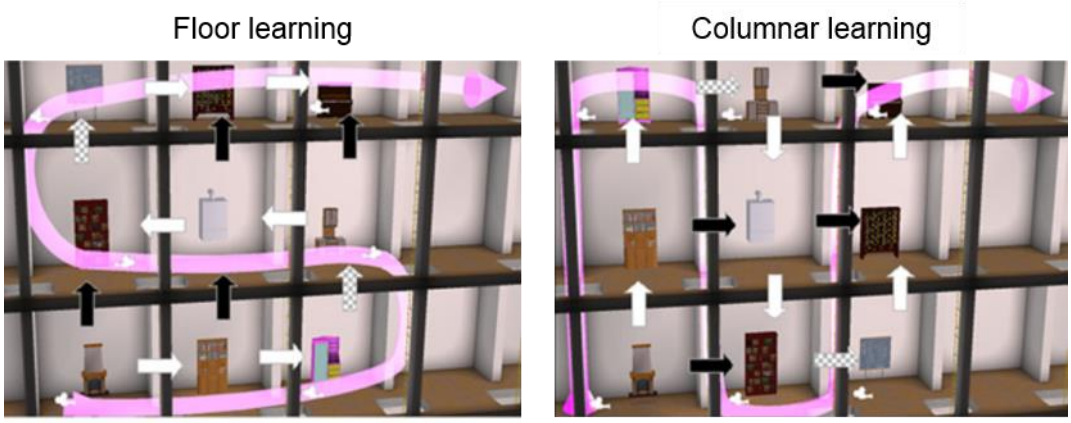


Figure 1.16 A virtual multi-level environment was learned and tested by either floor or columnar routes. Pink arrows show the route of the learning. White arrows and black arrows are the familiar and novel segments used during testing, respectively. Reproduced from (Thibault et al. 2013) with permission from Springer Nature.

Above, I summarised the importance of a slope for navigation and the multiple navigation strategies used in multi-level buildings (regionalisation, floor-base or vertical column-base). But humans sometimes navigate in a volumetric 3D space without being bound to surfaces like slopes and buildings (e.g. astronauts, pilots and divers). How do people navigate in a volumetric 3D space? As with other animals, the literature on volumetric navigation is scarce, probably due to the difficulty of implementing volumetric navigation experiments in a laboratory setting. Only a few virtual navigation experiments have investigated spatial cognition when participants made controlled 3D rotations in tunnel-like structures. In Vidal et al. (2004), participants had to identify the correct layout of 3D tunnels after being passively moved in the 3D tunnels with multiple rotations (Figure 1.17). In a “terrestrial” condition, only yaw rotations were used so the participants were always in a head upright position (Figure 1.17A). This means that the participants were moved straight up and down in the vertical section of the tunnels as if they were inside a lift. In a “weightless” condition, participants always moved forward which means that either yaw or pitch rotation was applied at each junction (Figure

1.17B). Participants were significantly more accurate in the terrestrial condition than the weightless condition when a complex 3D maze (with multiple turns) was used. However, there were some learning effects in the weightless condition because accuracy almost reached that of the terrestrial condition in the last trials. Although this learning effect was descriptive (a formal statistical test was not possible due to the randomised trial orders used in the experiment), it is an interesting observation that humans might be able to process the infrequently used non-horizontal rotation (pitch) as accurately as the horizontal rotation (yaw) within a short training period of less than an hour. Aoki et al. (2005) also found that a 3D pointing task was more difficult when participants made vertical rotations, but the learning effect was not investigated in this study.

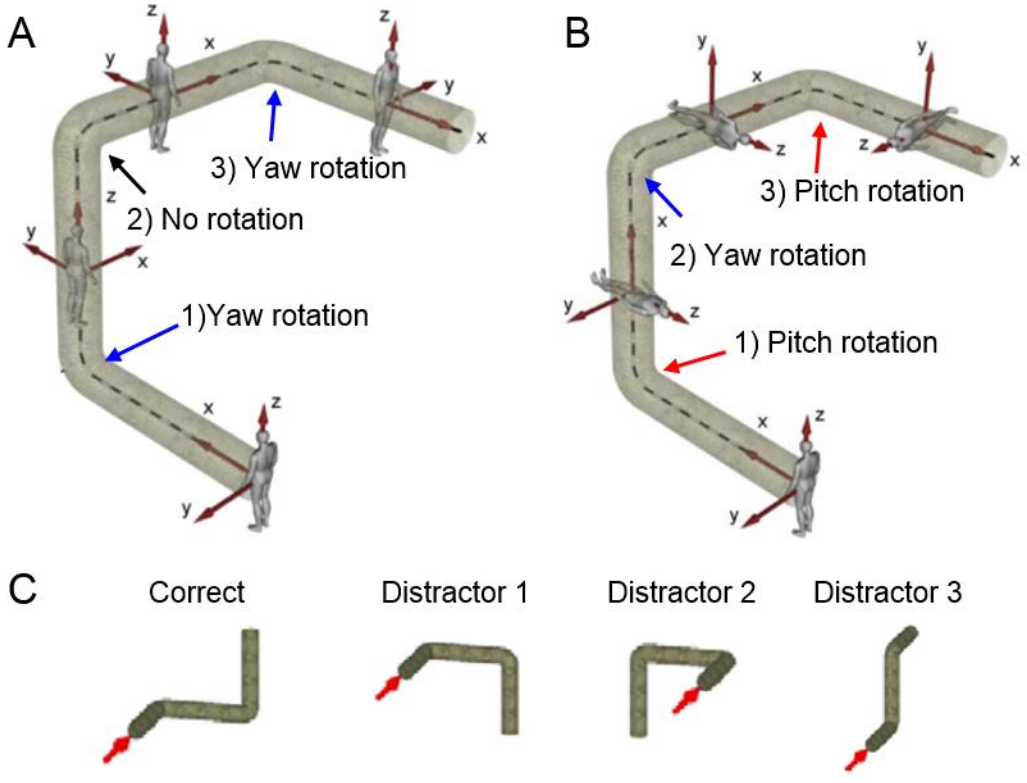


Figure 1.17 A virtual 3D tunnel experiment. (A) In the terrestrial condition, participants always kept their head upright with yaw rotation only. (B) In the weightless condition, participants were always moved forwards and either pitch or yaw rotation was applied at the junctions. (C) At the end of the 3D tunnel, the participant selected, from among lures, which option was the structure of the 3D maze. Red arrows represent the starting location. Adapted from (Vidal et al. 2004) with permission from Elsevier.

The 3D rotation experiments described above argue for the importance of the plane of rotation relative to the gravity axis. It has been suggested that participants who had difficulty spatially updating for the vertical rotation in the two studies described above might have experienced the Visual Reorientation Illusion, a common symptom reported by astronauts (Oman 2007). This is the phenomenon where the surrounding walls, ceiling, and floor suddenly exchange identities. It occurs when an astronaut works upside down or views other crew members floating upside down (compared to the "up-down" orientation on earth, where the astronauts were trained before the space mission). Jeffery et al. (2013) argued that the use of a local planar reference like "floor" and "ceiling" by astronauts in a microgravity volumetric environment shows that even volume-travelling animals might use a planar representation of 3D space. However, it is unclear how humans or other animals represent an entire volumetric space using a few planes if they constantly change the plane of movement.

1.7 Human neuropsychological, neuroimaging and electrophysiological studies in 3D space

To the best of my knowledge, very few studies have investigated neural activity, or the effect of lesions, in the human brain while participants moved both horizontally and vertically in a 3D space, and none directly assessed how different types of 3D spatial information, such as location and direction, are encoded in the brain. Before I describe the few studies on 3D spatial cognition, I summarise the neuropsychological, neuroimaging and electrophysiological evidence of 2D spatial information encoded in the human brain, which agrees with, and further extends, the previous animal electrophysiological findings on place cells, head direction cells, grid cells and border cells.

1.7.1 Neural correlates of navigation and spatial memory in humans in 2D

A century of neuropsychological studies has provided valuable information about which brain regions are responsible for spatial disorientation. As reviewed in (Aguirre and D'Esposito 1999), spatial disorientation can arise for multiple reasons that involve different brain structures. Patients with lesions in the posterior parietal lobe show egocentric disorientation - difficulty in knowing where objects or places are relative to the own body (Levine et al. 1985; Stark 1996). Patients with small focal lesions in the right retrosplenial region show a rather pure topographic disorientation ("sense of direction") with little evidence of visuospatial agnosia or egocentric disorientation (Takahashi et al. 1997). This finding might be related to head direction cells and other spatially modulated cells found in the rodent retrosplenial cortex (Cho and Sharp 2001; Alexander and Nitz 2015). Spatial disorientation could also originate from a deficit in recognising landmarks and places and this type of agnosia is often associated with medial occipital gyrus lesions (Pallis 1955).

Consistent with the well-established findings of place cells in the rodents hippocampus and impairment in water maze learning following hippocampal lesions in rodents (O'Keefe and Dostrovsky 1971; Morris et al. 1990), topographic memory impairment are observed in patients with lesions to the hippocampi or broader medial temporal lobe (Smith and Milner 1981; Maguire et al. 1996; Spiers et al. 2001). However, bilateral hippocampal lesions do not impair all aspects of spatial navigation. For instance, a taxi driver with bilateral hippocampal lesions showed intact static topographic memory such as the relationships between landmarks, and he also found his way in a virtual simulation of central London, UK that he learned 40 years ago, but only when the route involved the main artery roads. He became lost when the route depended on non-main roads, suggesting that the hippocampus is specifically required for detailed spatial representations (Maguire et al. 2006).

The involvement of the human hippocampus in spatial navigation was also observed in some of the earliest neuroimaging studies. For example, a positron emission tomography (PET) study showed that successful active navigation in a virtual town evoked more regional cerebral blood flow in the hippocampal region compared to passive arrow-following movement or unsuccessful navigation trials (Maguire et al. 1998). Functional magnetic resonance imaging (fMRI) studies also showed similar navigation performance-related activity in the hippocampus (e.g. Hartley et al. 2003). Of note, navigation is a complex cognitive function which engages more than a single neural structure (as suggested by neuropsychological studies described earlier), and the contrast between active navigation and simple route-following revealed a large brain network including parahippocampal cortex, ventral occipital cortex, medial parietal regions, cerebellum and areas of the prefrontal cortex (Hartley et al. 2003).

Spiers and Maguire (2006) revealed the complex online neural interactions between the brain areas supporting navigation, by implementing a realistic simulation of wayfinding in London, UK during fMRI scanning which included collecting retrospective verbal reports from each participant with an unprecedented degree of detail. They found that hippocampus was particularly engaged during initial planning of routes, but not during other periods of navigation such as spontaneous route planning, action planning or visual inspection. Together with the functional neuroimaging findings, a structural change associated with acquiring spatial knowledge of London's layout was found in the hippocampus of London taxi drivers (Maguire et al. 2000; Woollett and Maguire 2011).

1.7.2 Place, direction, grid, border representation in the human brain in 2D

The early neuropsychological and functional neuroimaging studies suggested a role for the human hippocampus in spatial navigation, and then the first direct evidence of place cells in the human hippocampus was reported in 2003 from an invasive recording study. Using depth electrodes implanted for pre-epilepsy surgery planning, Ekstrom et al. (2003) recorded neurons in the hippocampus, parahippocampal region, amygdala and frontal lobes while participants navigated around a virtual town. They found the cells that responded to the instantaneous location of participant, a particular view or a goal location. Place cells (that responded to the current location of a participant) were more abundant in the hippocampus compared to the other regions. Grid-like cells have also been observed in the entorhinal cortex, cingulate cortex and hippocampus in neurosurgical patients (Jacobs et al. 2013). Invasive human electrophysiological studies provide unique and detailed information on neural codes at a cellular resolution which is not achievable with current non-invasive neuroimaging methods (Parvizi and Kastner 2018). However, the invasiveness of this method, which requires a clinical setup (e.g. clinically-determined locations of electrodes, short and relatively simple experiments designed for patients, and patients with epilepsy-related pathology) renders it unsuitable for widespread use in studying human neuroscience. The challenge, then, is how we non-invasively probe place cells, head direction cells and grid cells in the healthy human brain.

This challenge was tackled in an fMRI experiment with multivoxel pattern analysis (Hassabis et al. 2009). In this study, participants visited the four corners of a rectangular room in a virtual environment during scanning. When they arrived at each corner, they looked down at the floor which was visually identical for each corner. fMRI activity was analysed for this floor-viewing period, meaning that different fMRI activity between each spatial location could not be explained by different views. The authors

could decode at which corner a participant was located from the patterns of activity across voxels in the hippocampus. This finding is important not only because it is the first fMRI evidence showing the presence of place information in the hippocampus, but because of its implications for the nature of neuronal ensembles in the hippocampus. If place cell decoding for each location relies on neurons that are mingled within each and every voxel (1.5mm cubic, which contains ~100,000 neurons), multivoxel pattern classification would be impossible. The successful decoding of location information from multivoxel patterns in the hippocampus hence suggests a spatially distributed, uneven neural representation within the hippocampus.

For example, previous animal studies showed that place fields recorded in ensembles tend to cluster (Eichenbaum et al. 1989), although this does not mean that place cells are topographically organised as in primary sensory cortex. The precise organisation of the neural code in the hippocampus is still not fully understood and there exists scepticism about the detectability of a place code at the level at which fMRI operates (Nolan et al. 2017). Nevertheless, multiple fMRI studies have since revealed place information in the hippocampal formation, as well as in the retrosplenial cortex and parahippocampal cortex (Vass and Epstein 2013; Sulpizio et al. 2014) using multivoxel pattern analysis and repetition suppression analysis. These analysis methods will be explained in detail in the next chapter.

A participant's head direction information has also been observed in multiple fMRI studies using multivoxel pattern analysis and repetition analysis. Unlike animal work, where typically head direction cells are recorded in freely and continuously moving animals, most human participants in fMRI studies have been given only static visual cues (e.g. a picture or word describing an object). This prompted them to either explicitly or implicitly think about their orientation in a virtual or real environment that they experienced before scanning (e.g. "the building is visible when I am facing north")

(Figure 1.18). Cue presentation typically lasts a few seconds and is followed by a short pause (blank screen) before the next cue presentation. This static and discrete experimental design is used to control views that are naturally associated with head direction and to control the timing (onset and offset) of each directional response for ease of fMRI data analysis. The obvious limitation of this approach is that it only evokes directional responses based on fixed landmarks. In the real world, angular velocity integration from self-motion cues (proprioceptive, vestibular, optical flow) are important sources of inputs to head direction cells. Nevertheless, fMRI researchers have found evidence for head direction information in areas that are known to contain head direction cells in animals, such as the thalamus (Shine et al. 2016), the retrosplenial cortex (Baumann and Mattingley 2010; Vass and Epstein 2013; Marchette et al. 2014; Shine et al. 2016) and the subiculum (Chadwick et al. 2015). Parahippocampal cortex (Sulpizio et al. 2014; Bellmund et al. 2016) and hippocampus (Sulpizio et al. 2014) have also been reported to contain head direction information.

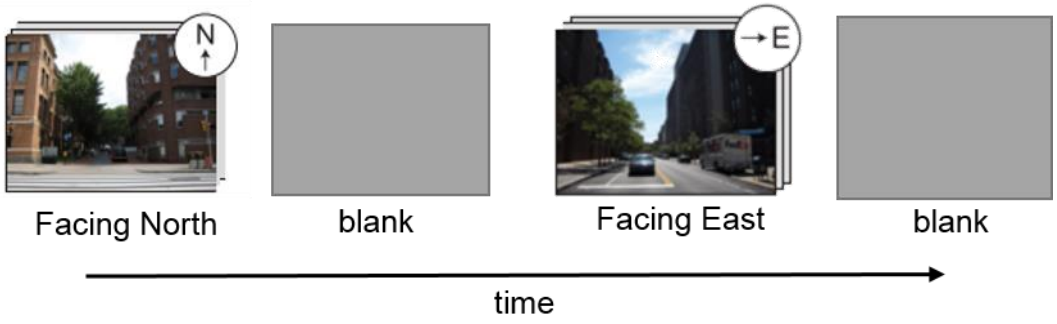


Figure 1.18 A typical fMRI experiment using a static cue (visual or verbal) for a few seconds to evoke the neural response to each direction in the environment. Figures adapted from Vass and Epstein 2013 with permission conveyed through the Copyright Clearance Center, Inc.

Compared to place cells or head direction cells, grid cells have more complex response patterns (e.g. multiple firing fields, multiple scales and offsets), making the systematic variation of a signal at the voxel level (which contains tens of thousands of grid cells) unlikely. However, properties of grid cells that allow for their detection using fMRI were

discovered based on findings from rodent electrophysiology, and the first fMRI evidence of grid cells in the entorhinal cortex was reported in 2010 (Doeller et al. 2010). The key properties are (1) that orientation of the grid axis is relatively constant across cells and (2) that a grid cell's activity is modulated by an animal's movement direction relative to the grid axis. How to probe grid cell signals using fMRI will be discussed later in this thesis. Grid cell-like fMRI signals have since been reported in a number of studies that used either actual navigation or mental imagery (Kunz et al. 2015; Bellmund et al. 2016; Horner et al. 2016). Furthermore, a grid-like representation was also found in the non-spatial domain (Constantinescu et al. 2016; Julian et al. 2018; Nau et al. 2018). Of note, fMRI studies have revealed evidence of grid-like codes not only in the entorhinal cortex but also in a broader network including the prefrontal cortex in humans (Doeller et al. 2010; Constantinescu et al. 2016).

Regarding boundary cells, a recent invasive recording study found an increase in theta power in the subiculum when a target location was close to the boundary of a circular arena (Lee et al. 2018). This effect was not present in electrodes in CA1, dentate gyrus, entorhinal or perirhinal cortex. Lee et al. (2018) examined the neural responses when participants were virtually moving towards the target, not when the participants were actually at the boundary, in order to dissociate the boundary signals from view-related signals. Boundary-related learning effects have also been reported in the posterior hippocampus (Doeller et al. 2008).

1.7.3 Neural correlates of movement and navigation in humans in 3D

There is a very limited neuroimaging literature concerning navigation in 3D space. In Indovina et al. (2016), participants were moved on a virtual rollercoaster during fMRI scanning. The rollercoaster travelled forward and then turned either left, right, up or down in a tunnel. The subsequent linear movements inside the tunnel were visually

identical regardless of the preceding turns. Brain activity was compared between each of four conditions (leftwards, rightwards, upwards and downwards) during this linear movement period. Bilateral hippocampus showed greater activity for the horizontal movements whereas the posterior portion of inferior temporal gyrus and the angular gyrus showed greater activity for vertical movements. No region distinguished the orientation within either the vertical or horizontal plane (e.g. leftward movements were comparable to rightward movements). The authors argued that the distinct brain activations for movement on the horizontal and vertical planes may allow parallel processing of 3D information, and in so doing, simplify navigation in 3D space.

It remains unknown whether these preferential responses to either horizontal or vertical motion exist even when participants move freely in a 3D space where horizontal and vertical motion are not separated, as in the rollercoaster experiment. For instance, which brain areas will be activated when a participant turns right-upwards? The more important question is how the brain encodes the specific orientation information of a participant. Although this study showed that the amplitude of brain responses differs when a participant moved horizontally or vertically, we need much more information than this to navigate in 3D. For instance, we need to distinguish whether we move up or down and how much we move up (e.g., 30° or 60°), but the neural encoding of this fine-grained 3D spatial information in humans is, as yet, unspecified.

Zwergal et al. (2016) also compared overall brain activity between vertical and horizontal navigation, but without testing for fine-grained 3D spatial information encoding. In this PET study, half of the participants searched for objects located at different levels on the staircase of a hospital (“vertical navigation”) while the other half searched for objects while on one floor (“horizontal navigation”). A similar degree of activation in the hippocampus was reported for both the vertical and horizontal navigation groups during subsequent scanning. Visual cortex and the brainstem

showed more activity during horizontal navigation. By contrast, bilateral insula, anterior cingulate cortex and cerebellum showed more activity during vertical navigation. The authors claimed that horizontal navigation is predominantly guided by visual landmarks and vertical navigation is most likely guided by vestibular inputs (insula and cerebellum). However, the claim that vertical navigation is less dependent on visual landmarks is questionable, because there was in fact a distinct lack of sufficient visual information in the vertical navigation condition in this experiment. The floor signs, important visual landmarks that distinguish visually similar multiple floors, were hidden in this experiment. The contribution of visual landmarks to navigation should be investigated in future experiments where the visual cues are well matched between horizontal and vertical conditions.

1.8 The anatomy of the hippocampal formation

In this section, I present an overview of the anatomy of the hippocampal formation which contains hippocampus, subiculum and entorhinal cortex, some of the key regions studied in this thesis. Multiple views exist on the definition and terminology of hippocampus or hippocampal formation. The difference in species and limitations in the spatial resolution of non-invasive methods makes the anatomy even more complex to describe.

Amaral and Lavenex defined the “hippocampal formation” as being comprised of the hippocampus proper, dentate gyrus, subiculum, presubiculum, parasubiculum and entorhinal cortex (Amaral and Lavenex 2006) (Figure 1.19). The hippocampus proper consists of subfields called CA1, CA2, CA3 (Cornu Ammmonis, meaning a ram’s horn due to its curved shape). The subiculum and pre and parasubiculum are collectively called the subicular complex or subicular cortices. In human fMRI studies, the subiculum label is often used to include presubiculum and parasubiculum due to limited

spatial resolution (Adler et al. 2014; Yushkevich et al. 2015). The simplified textbook account of the connectivity in this region is as follows: cerebral cortex → medial entorhinal cortex → dentate gyrus → CA3 → CA1 → subiculum → medial and lateral entorhinal cortex → cerebral cortex. The functional, anatomical and connectivity differences between the substructures are important research topics, but unfortunately delineating these substructures from in vivo MRI images with standard spatial resolution (1~3 mm) is very challenging. Numerous segmentation protocols exist with substantial discrepancies between them, especially in the anterior portion of the hippocampus (the boundary between CA1 and subicular cortices) (Yushkevich et al. 2015; Dalton et al. 2017). In most fMRI studies, the term “hippocampus” has been loosely defined and used to include the hippocampus proper (CA1/2/3), dentate gyrus and subicular cortices.

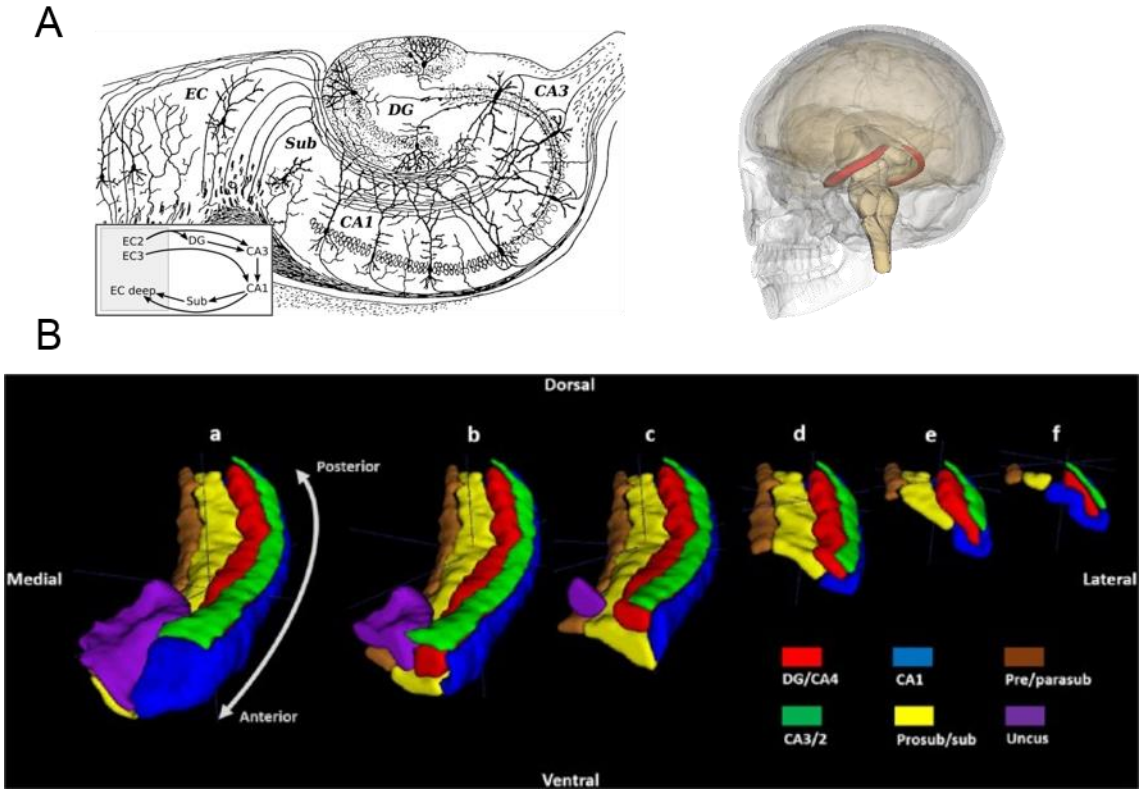


Figure 1.19 The anatomy of the hippocampus. (A) Left, a drawing of the rodent hippocampus by Cajal. Right, the location of the hippocampus (red) in the human brain. DG, dentate gyrus; Sub, subiculum; EC, entorhinal cortex. Figures from Wikimedia in the public domain. Image generated by Life Science Databases. (B) The subfields of hippocampus are shown with anterior slices moving along the longitudinal axis. Figure from Dalton et al. 2017.

1.9 Summary and thesis objectives

In this chapter, I have reviewed the literature on how specific cell types encode spatial information (e.g. location, direction, distance) and behavioural findings on spatial navigation that hint at how the 3D world is represented in the brain. To date, due to a paucity of empirical findings and the considerable differences in species, behaviour and shapes of 3D environments used in the existing literature, there is no consensus on basic issues such as whether 3D space is encoded in a volumetric manner or in a planar manner, or whether the vertical and horizontal axes are equally well encoded. In particular, there is a big gap in our understanding of 3D spatial representation in the human brain.

In my thesis, I aimed to provide insights into how 3D spatial information is encoded in the human brain when people navigate in various types of 3D environments. To do this, in the next chapter, I first introduce the general experimental methods used in this series of studies, such as the VR techniques and the basics of fMRI and data analysis. I then present the series of experiments that I conducted during my PhD in the following order:

In Chapter 3, I built a semi-volumetric virtual environment where participants could climb up and down, somewhat like a junglegym in a playground. I examined whether navigationally relevant brain structures contained vertical and horizontal spatial information (e.g. the place and direction of a participant moving in this virtual environment) equally well, or whether some brain structures were more sensitive to either the vertical or horizontal axis. I used both behavioural measurements (accuracy and response time for one's location/direction judgments, and size estimation of the 3D virtual environment) and fMRI multivoxel pattern similarity analysis to address these questions.

In Chapter 4, I developed a virtual gallery building which had multiple levels and rooms. I hypothesised that this type of 3D environment could be encoded in a hierarchical manner, e.g. where am I in a room, which room am I in within the wider building. Using fMRI repetition suppression analysis, I searched for where in the brain these different types of spatial information were encoded. I also tested whether there was a bias in vertical and horizontal information (e.g. whether people got more confused about the vertical floors than the horizontal locations within a floor).

In Chapter 5, I built a virtual zero gravity environment where participants could move along all 3 axes freely. Participants explored the virtual spaceship rendered on VR head-mounted display which provided multisensory inputs prior to scanning. I then tested how the vertical and horizontal components of 3D head direction were encoded in the head direction cell network using fMRI multivoxel similarity analysis.

In Chapter 6, I developed an fMRI analysis method to investigate a grid cell's activity in 3D based on the known direction-modulated property of grid cells. Using this method, I tested for the existence of a 3D grid-code in data I collected using the virtual zero gravity environment. I also developed interactive software that visualises 3D grid cells and predicts the activity of grid cells as a function of a participant's movement.

Finally, in Chapter 7, I draw the experimental results together to propose how the human brain encodes spatial information in various 3D environments. I discuss current methodological limitations and suggest future experiments for pursuing a fuller understanding of 3D spatial representation and navigation.

1.10 Publications

The following publications have arisen from work in this thesis:

Kim M, Jeffery KJ, Maguire EA (2017) Multivoxel pattern analysis reveals 3D place information in the human hippocampus. ***Journal of Neuroscience*** 37: 4270-4279.

Kim M, Maguire EA (2018) Hippocampus, retrosplenial and parahippocampal cortices encode multi-compartment 3D space in a hierarchical manner. ***Cerebral Cortex*** 28: 1898-1909.

Kim M, Maguire EA. 3D grid cells in human entorhinal cortex: Theoretical and methodological considerations and fMRI findings (revisions invited).

Kim M, Maguire EA. Thalamus, subiculum and retrosplenial cortex encode 3D head direction information in volumetric space (revisions invited).

Other publications arising during my PhD:

Clark IA, Kim M, Maguire EA. Verbal paired associates and the hippocampus: The role of scenes. ***Journal of Cognitive Neuroscience*** (in press).

Kim M, Barnes GR, Maguire EA. Temporal dynamics of spatial memory recall in 3D space (in preparation).

Chapter 2 General methods

2.1 Overview

In this chapter, I describe the basic principles of the data acquisition and the analysis methods that I used throughout my experiments. First, I introduce the behavioural setups and virtual environments that I developed. I then explain the principles of measuring neural responses using fMRI, preprocessing of raw data and the use of the general linear model (GLM). Finally, I outline the core fMRI analysis methods - repetition suppression analysis and multivoxel pattern analysis and the software used for statistical tests.

2.2 Participants

All participants were in the age range of 18 to 36 years, right-handed, had normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. The number of male and female participants was closely matched in each experiment. Most participants were UCL undergraduate and masters students recruited from the Institute of Cognitive Neuroscience's subject database. All gave informed written consent in accordance with the local research ethics committee. Most participants only took part in one experiment. During the recruitment process, I screened all participants for the risk of motion sickness or VR sickness. A minority of participants felt discomfort when interacting with one of my VR environments. In these cases, the experiment was ended immediately, and the participant was supervised until the feeling of sickness had passed.

2.3 Virtual environments

The use of virtual environments was central to my PhD given that I investigated the neural encoding of 3D spatial information using neuroimaging techniques where participants were physically constrained in an MRI scanner. I developed three unique 3D virtual environments for use in the experiments (a junglegym, a multi-level building and a spaceship) within which participants could navigate. The exact details of each environment are described in the experimental chapters, with general background information provided here.

I implemented the virtual environments using the 3D modelling software Sketchup (Trimble, CA, USA) and the game engine Unity 4.6 or 5.4 (Unity Technologies, CA, USA). This game engine is a complete solution for visual graphics, physics, scripting and user interface, and it supports multi-platforms (Windows: the main stimulus presentation for all my experiments; Android: the VR head-mounted display for Experiment 3; WebGL: the web-based grid visualisation software developed for Experiment 4).

Throughout my experiments, I used a first-person perspective to provide participants with the immersive feeling that they were inside a 3D space. The screen aspect ratio and camera parameters which determined the field-of-view of a participant in the virtual environment were carefully set to mimic natural human vision. A wide screen aspect (16:9 for Experiment 1 and 4:3 for Experiments 2-4) was used because human eyes have a wider field-of-view for the horizontal axis. The vertical field-of-view was ±30° for all experiments and the horizontal field-of-view was ±45.7° for Experiment 1 and ±37.6° for Experiments 2-4 (Figure 2.1A). Although humans can view up to 100° laterally (Spector 1990), our peripheral vision is not as good as our central vision, and the Unity monocular camera with a larger field-of-view induces significant distortion and

discomfort, particularly when a participant looks up or down (Figure 2.1B). This was why I used a vertical field-of-view of ±30° (a commonly used default camera setup). However, when a virtual environment is rendered on a VR head-mounted display which has binocular lenses placed closed to the eyes (the pre-scan task in Experiment 3), a wider field-of-view can be achieved without distortion. I therefore used a vertical field-of-view of ±45° (a default setup for VR head-mounted displays). The horizontal field-of-view was ±48° for the VR head-mounted display (the hardware's specification).

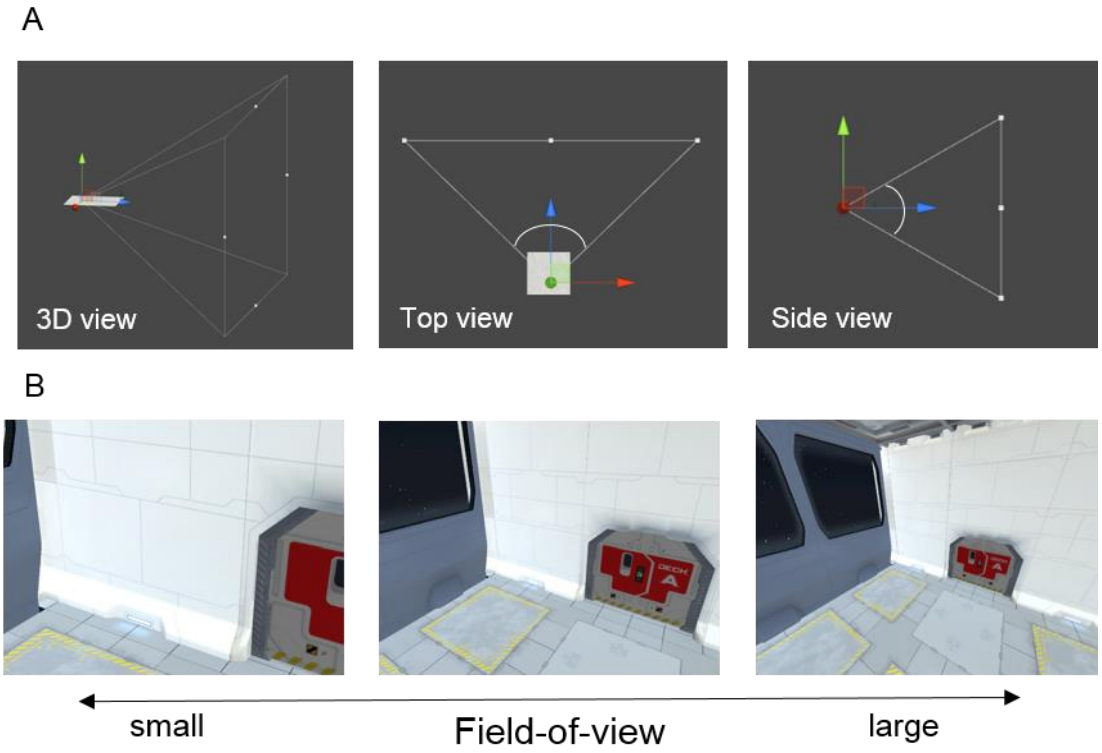


Figure 2.1 The field-of-view in an example virtual environment. (A) The top view shows the horizontal field-of-view and the side view shows the vertical field-of-view. (B) Example views for different vertical field-of-views (left, ±15°; middle, ±30°; right, ±45°). Horizontal field-of-view is proportionally scaled.

2.4 Experimental procedures

All experiments were conducted at the Wellcome Centre for Human Neuroimaging, Institute of Neurology, University College London, London, UK. Upon their arrival, participants gave written informed consent after having read an information sheet that

described the study and having asked any questions that they had. Participants performed various spatial exploration and memory tasks in a virtual environment. Pre-scan behavioural training and testing took approximately 1 hour for all experiments. Behavioural experiments took place in a quiet testing room equipped with a desktop PC. Once participants successfully completed the behavioural phase, they were thoroughly screened once again for MRI safety (having also been screened prior to their arrival at the Centre).

After the safety check, a participant wore ear protection and lay on the MRI table in a supine position. Foam pads were placed at the sides of their head to further reduce the acoustic noise and to prevent excessive movement. A breathing belt and pulse oximeter were attached to the participant so that experimenters could check for any signs of distress during scanning. In addition, a participant's status was monitored via a camera. In all of the experiments, a participant was provided with an emergency alarm which they could squeeze at any point during scanning to indicate that they wanted to come out of the scanner. They were also provided with an MRI-compatible keypad (Current Designs, PA, USA) with which to accomplish the tasks, and this was held in their right hand.

A scanning session lasted ~1.5 hours including setup time and structural scanning. Participants were debriefed after scanning and reimbursed for their time (£10/hour). In total, each participant visited the Centre for ~3 hours.

2.5 Principles of MRI and fMRI

In this section, I present a brief conceptual explanation of the biophysical basis of MRI and fMRI, given that this was the main method I used in my experiments.

2.5.1 Signal generation in MRI

MRI is a non-invasive imaging method that relies on the difference in magnetic property between tissues. Any atom that has magnetic spin can be used for MRI, but hydrogen is commonly used in MRI because of its abundance in our body (e.g. H_2O). A hydrogen nucleus has one spinning proton and this positively charged particle generates an electric current and a magnetic moment. In the presence of a strong magnetic field (the Siemens 3 Tesla MRI scanner used in my experiments creates a magnetic field that is several hundred times stronger than a typical refrigerator magnet), protons align to the external magnetic field, resulting in a net magnetization.

This net magnetization precesses around the longitudinal axis of the external field at Larmor frequency which is proportional to the external field strength and a gyromagnetic ratio of atoms. If electromagnetic energy that matches the Larmor frequency is delivered (the radiofrequency pulse provided by a head coil in the MRI scanner), protons are “excited” and change from a low energy state to a high energy state. The excitation tips the net magnetization from the longitudinal axis to the transverse plane. Once the excitation ceases, transverse magnetization decays with a time constant of T2 (called T2 relaxation) and longitudinal magnetization recovers with a time constant of T1 (called T1 relaxation). Importantly, different tissues (e.g. grey matter, white matter, cerebral spinal fluid) have different T1 and T2, resulting in different magnetic resonance (MR) signal intensity (Figure 2.2). There is also T2* which is a time constant for the decay of transverse magnetization under the presence of local field inhomogeneity. T2* is a crucial concept for fMRI and is described in a later section.

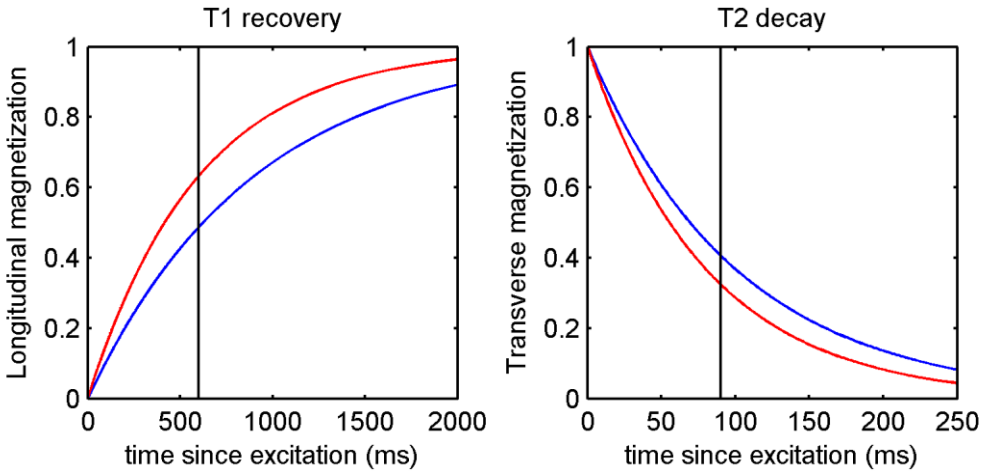


Figure 2.2 T1 and T2 contrast for MRI. Different tissues have different T1 and T2 constants which lead to a difference in MR signal intensity. For instance, longitudinal magnetization is greater in white matter (red line) than grey matter (blue line) at 600 ms from the excitation (left panel). Therefore, white matter is shown bright (higher intensity) and grey matter is shown dark (low intensity) on T1 images.

2.5.2 From raw signals to an image

Raw MR signals are measured by receiver coils in MRI in the form of time-varying electric current. How do we know where this signal comes from in order to reconstruct a 3D spatial image? This spatial encoding problem is solved by applying a gradient to the external magnetic field. If a small extra magnetic field is applied to a particular part of the brain (e.g. a horizontal slice at the level of amygdala, z = -26mm), protons within the slice will precess at a slightly higher Larmor frequency than the rest of the brain, and receiver coils can selectively measure signal from this slice (known as slice selection). Different gradients along two other axes can be applied to localise the signal within a slice (known as frequency and phase encoding). Consequently, a 3D structural image is achieved.

2.5.3 BOLD signal for fMRI

The BOLD (blood-oxygenation-level-dependent) fMRI signal is an indirect way of measuring neural activity via neural-hemodynamic coupling. A simple explanation for

the origin of the BOLD signal is as follows. Active neurons demand oxygen and glucose, and the energy is supplied by neighbouring blood vessels. There is a complex cascade of biochemical processes that increases the local cerebral blow flow when neurotransmitters are released from active synapses (Figure 2.3). Paradoxically, a supply of fresh oxygenated haemoglobin is greater than demand, resulting in a decrease in local concentration of deoxygenated haemoglobin. Deoxygenated haemoglobin is paramagnetic which means it distorts the local magnetic field and shortens T2*. A decrease in deoxygenated haemoglobin increases T2*, leading to higher MR signal. Therefore, we see “activation” (higher BOLD signal relative to the baseline) in the visual cortex when participants are shown a picture, and activation in the motor cortex when participants tap their fingers. If we see activation in some brain regions during particular cognitive or behavioural tasks, we infer that brain regions (or neurons within those brain areas) are involved in that task.

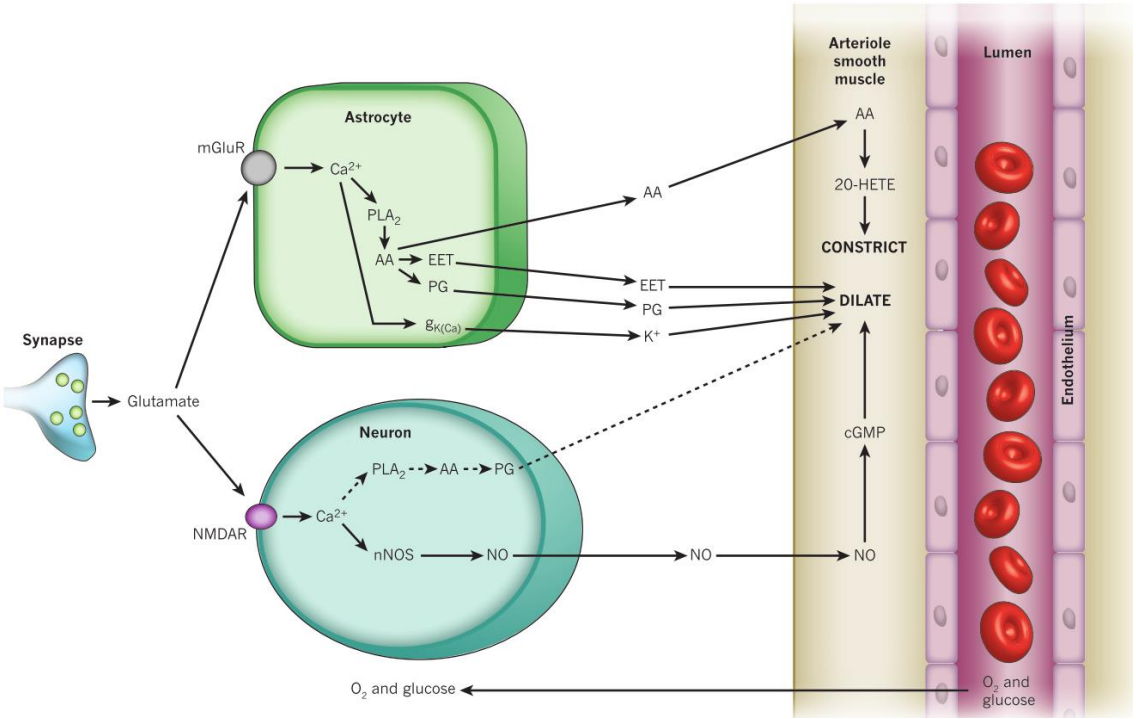


Figure 2.3 Neural-hemodynamic coupling. Glutamates released from a synapse trigger calcium signalling in both astrocytes and neurons which leads to the production of various vasoactive agents including nitric oxide (NO) and arachidonic acid (AA). The vasoactive agents constrict and dilate a blood vessel. The change in blood flow and the concentration of deoxygenated hemoglobin is the basis of the fMRI signal. Reprinted from (Attwell et al. 2010) with permission from Springer Nature.

However, “the higher the neural activity, the larger the BOLD signal” is an overly simplified description. The precise spatiotemporal relationship between electrophysiological activity of neurons and the BOLD signal is still unknown and continues to be actively researched. Previous studies that simultaneously measured fMRI and electrophysiology have mostly been conducted in primary sensory cortex in animals. These studies revealed the typical shape of the hemodynamic delay (e.g. the BOLD signals lags about 6 seconds from the peak neural activation) and that the BOLD signal is better correlated with the local field potential (LFP) than spiking activity (Logothetis 2003). This implies that inputs to the neurons contribute more to the BOLD signal than neural outputs. Although inputs and outputs are often correlated, it is possible that inhibitory inputs (which reduce spiking) could evoke a positive BOLD signal. Furthermore, different brain regions can have different neural assemblies and different neurovascular coupling.

For instance, the hippocampus was reported to show a weaker correlation between BOLD and theta-band activity than nearby parahippocampal cortex (Ekstrom et al. 2009). The sign of the BOLD signal change is not trivial to interpret because it depends on the experiment-specific baseline. In cognitive experiments, a baseline is often defined as the inter-trial-interval when stimuli are not presented and participants are not asked to perform specific tasks. Sensory cortex typically shows increased neural activity and positive BOLD signal relative to this non-stimulus baseline. However, brain regions in the default mode network, including the hippocampus, are not inactive in the absence of external stimuli, and therefore BOLD signal can be even smaller during stimulus presentation compared to the “baseline” (Ekstrom 2010). I observed this negative BOLD signal in Experiment 2.

In short, the BOLD fMRI signal measures spatiotemporally filtered neural activity (e.g. action potentials and synaptic currents induced by hundreds of thousands of neurons in

a typical fMRI voxel) via complex neural-hemodynamic coupling. I used BOLD fMRI in my experiments because it is currently the best non-invasive technology that can measure the entire brain, including deep structures such as the hippocampus, with good spatial resolution. Nevertheless, the inherent methodological limitations of fMRI should be considered when interpreting the results.

2.5.4 Scanning parameters

For BOLD imaging, T2*-weighted echo planar images (EPI) were acquired using a 3T Siemens Trio scanner (Siemens, Erlangen, Germany) with a 32-channel head coil. Scanning parameters optimised for reducing susceptibility-induced signal loss in areas near the orbitofrontal cortex and medial temporal lobes were used as follows: slice TR = 70 ms, TE = 30 ms, resolution = 3 × 3 x 3 mm, matrix size = 64 x 74, z-shim gradient moment of -0.4mT/m ms (Weiskopf et al. 2006). The number of transverse slices angled at -30° was 48 for Experiment 1, and 44 for Experiments 2-4, resulting in a volume TR of 3.36 sec for Experiment 1 and 3.08 sec for Experiments 2-4. This sequence covered the entire brain except for the very top and bottom edges in rare cases where a participant had a particularly large brain. Before the main functional scans, fieldmaps were acquired with the standard manufacturer's double echo gradient echo field map sequence (short TE = 10 ms, long TE = 12.46 ms, 64 axial slices with 2 mm thickness with a 1 mm gap yielding whole brain coverage; in-plane resolution 3 x 3 mm. The fieldmaps were later used during preprocessing for correction of magnetic field inhomogeneity. After the functional scans, a T1-weighted structural scan (MDEFT) was obtained with 1mm isotropic resolution (Deichmann et al. 2004).

2.6 Preprocessing of fMRI data

Before the main analysis, raw fMRI data need to be preprocessed to compensate for signal artefacts, head motion and individual differences in anatomy. Here I briefly describe the preprocessing applied to my data. I used the neuroimaging analysis software SPM12 (www.fil.ion.ucl.ac.uk/spm) for preprocessing.

2.6.1 Discarding the first volumes

Signal intensity is larger for the first few volumes when the TR is relatively short for T1 recovery (Figure 2.4). To allow for T1 equilibrium, I presented the task stimuli once 5 volumes had been acquired and discarded the first 5 volumes from any analyses.

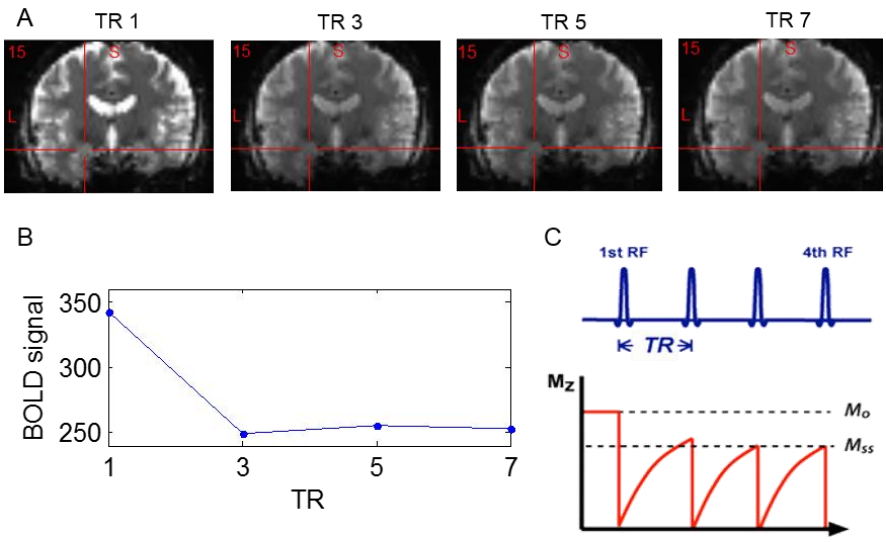


Figure 2.4 The T1 saturation effect. (A) Example images for the 1st, 3rd, 5th, 7th TR in one subject. (B) The signal intensity of one voxel indicated by the red crosshair in panel (A). Signal intensity is higher for the first volume because T1 magnetization has not yet reached the saturation. (C) A schematic explanation for T1 saturation.
Courtesy of Allen D. Elster, MRIquestions.com (http://mriquestions.com/dummy-cycles.html).

2.6.2 Realignment and unwarping

A participant's head is not completely static during scanning even if they have been instructed to remain still and the foam pads prevent excessive movements. This means that, for example, a voxel that corresponds to the anatomical coordinate of [20, 30, 28]

mm at t = 0 sec can represent [21, 31, 30] mm at time t = 100 sec. This movement can also interact with geometric distortion of images induced by local field inhomogeneity. To ensure each voxel measures the signal from the same brain structure throughout scanning with minimal geometric distortion, I used the SPM realign and unwarp function with the fieldmap images that measured the baseline field inhomogeneity. This function estimates 3 translation and 3 rotation parameters using rigid body spatial transformation. These 6 head realignment parameters were later included in the GLM to account for any residual movement that could influence the fMRI data.

2.6.3 Coregistration of functional and structural images

Functional images which have a lower spatial resolution (3 mm) need to be coregistered to a structural image which has a higher spatial resolution (1 mm) for later spatial normalisation and anatomical comparisons. I used the SPM coregistration function to match these images. This function finds the rigid transformation (affine matrix) that maximises the mutual information between the functional and structural image which have different signal intensity profiles (grey matter is brighter in functional images and white matter is brighter in structural images).

2.6.4 Spatial normalisation

In all of my experiments, I made inferences about neural activity at the population level. To achieve group-level inference, brain images from multiple participants need to be aligned in the same space, in my case using the “template brain” from the Montreal Neurological Institute (MNI) (Mazziotta et al. 1995). This inter-subject alignment process is known as spatial normalisation. I used the SPM segment and normalisation function to normalise individual brains to MNI space. This function segments the structural images into different tissues (white matter, grey matter, CSF, bone) based on

signal intensity. A Bayesian generative model is used to combine a prior tissue probabilistic map and intensity distribution in individual data. Local shrinkage and stretch (a deformation matrix) is applied in order to match individual brains to the MNI template brain.

2.6.5 Spatial smoothing

Spatial smoothing is often applied in fMRI research for three reasons. First, smoothing can maximise the signal-to-noise ratio (SNR) if the width of smoothing matches the size and shape of the true signal. For example, if the spatial extent of neural activity is expected to be 8mm, random spatial noise will be averaged out if smoothing of 8mm is applied. Second, smoothing makes the distribution of fMRI signal close to the normal distribution which is a common assumption for standard statistical tests. Third, smoothed data have a smaller number of independent signals, which can alleviate the multiple comparison correction problem. In SPM, a Gaussian smoothing kernel is used. The full-width at half maximum (FWHM) determines the size of smoothing and 1 to 3 voxel sizes (which is 3 to 9 mm for my fMRI data) is commonly used (Figure 2.5).

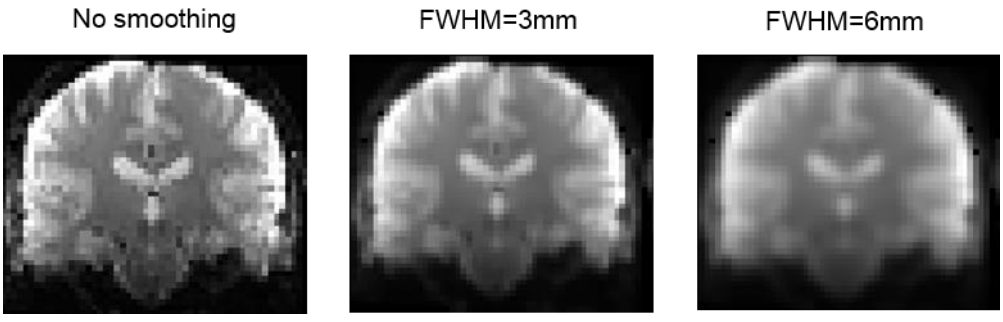


Figure 2.5 Spatial smoothing.

However, an obvious disadvantage of spatial smoothing is that it sacrifices the spatial resolution of fMRI data and it carries the risk of washing away the true signal when we do not know the size of the true signal. For example, if each voxel is tuned to slightly

different stimuli (e.g. voxel A responds more when a participant is facing North and its neighbouring voxels B, C, D respond more when a participant is facing South, East, and West respectively), averaging these voxels' responses would eliminate the chance of detecting a direction-specific response. Thus, it is a common not to spatially smooth the data in the context of multivoxel pattern analysis, which seeks neural information that distinguishes numerous experimental conditions. However, one study has suggested that smoothing does not harm multivoxel pattern analysis (Op de Beeck 2010).

In my experiment, I did not smooth the data when I used multivoxel pattern analysis (Experiment 1 and Experiment 3) and smoothed the data when I conducted univariate analyses (Experiment 2 and Experiment 4).

2.6.6 Temporal filtering

The fMRI signal can contain noise such as a slow fluctuation of the magnetic field in the scanner and also physiological noise. To remove any slow fluctuations, SPM includes a set of discrete cosine transform basis functions when it analyses fMRI data using the GLM which is described in the next section. I used a cut-off frequency of 128 sec for all of my experiments, which is the default in SPM.

2.7 The GLM and mass-univariate analysis

After preprocessing, fMRI data are typically analysed in a GLM framework. This is due to hemodynamic delay that renders temporal correspondence difficult. In other words, the fMRI signal at time t does not reflect the instantaneous neural response at time t, rather it reflects the blood oxygenation state which is influenced by temporally extended neural responses from up to 15 seconds prior to time t. Considering this slow

neural-hemodynamic coupling, researchers use a GLM to estimate neural responses to multiple experimental conditions.

Each GLM contains multiple regressors that predict fMRI signal for given experimental conditions. This regressor (Figure 2.6C) is a boxcar function that models the period of interest (e.g. when a participant is facing 30° – black line – and when a participant is facing 60° – red line, Figure 2.6A) convolved by a hemodynamic response function (Figure 2.6B). Throughout my experiments, I used the SPM canonical hemodynamic response function which is a mixture of gamma functions which model the BOLD signal as reaching a peak about 6 seconds after the neural impulse (Figure 2.6B). A GLM assumes that the fMRI signal is a weighted linear mixture of neural signal.

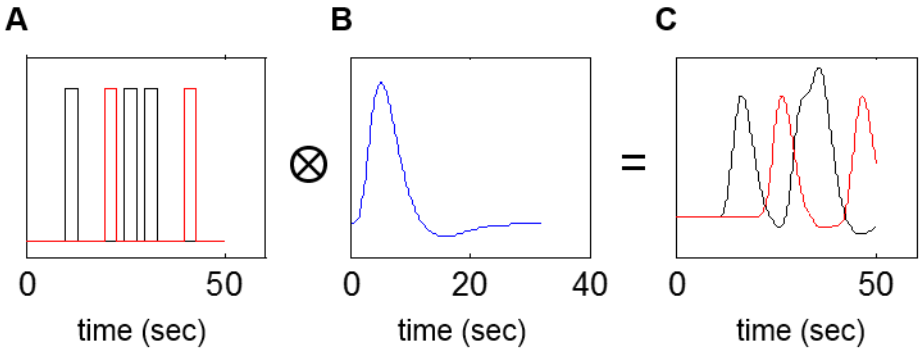


Figure 2.6 The hemodynamic response function. The fMRI signal (C) is modelled as a convolution of neural responses (A) and hemodynamic response function (B).

This can be formulated with the following equation:

$$y = X\beta + \varepsilon$$

y is the fMRI data at each voxel (a column vector where each element is the fMRI signal at each time point). X is called a design matrix which contains multiple columns (regressors) modelling each experimental condition such as “facing 30°” and “facing 60°” (black and red lines in Figure 2.6C). A design matrix X often contains regressors of no interest (“nuisance regressors”) such as the head realignment parameters or

session-specific constant terms. These nuisance regressors are included to rule out the influence of non-interesting variables on neural data. β is the weight of each regressor and it can be interpreted as the amplitude of fMRI signal for each experimental condition (e.g. how much fMRI signal varies when a participant is facing 30° and 60° respectively). The beta parameters need to be estimated from the data. ε is the residual error.

2.7.1 Parameter estimates

Under the assumption that the residual error variance is independent and homogeneous:

$$\varepsilon \sim \mathrm{N}(0, \sigma^2)$$

the best parameter estimate for the above equation is:

$$\hat{\beta} = (X'X)^{-1}X'\mathrm{y}$$

The above solution is known as Ordinary Least Squares. However, in real data residual error is unlikely to be independent and homogeneous. For instance, the fMRI signal at time t is likely to be similar to the signal at time t-1 and t+1, due to temporal autocorrelations. This violation of IID (independent and identically distributed) assumption complicates the parameter estimates and later hypothesis testing (e.g. the degrees of freedom). To overcome this problem, SPM first estimates the noise variance structure in the data using Restricted Maximum Likelihood (ReML) (Friston et al. 2002). This process recursively searches a parameter using sample covariance and a prior model covariance structure. I used a first-order autoregressive model, AR(1), in my experiments, which is the default in SPM.

In this process, SPM pools the sample covariance over voxels, assuming the noise covariance is constant across voxels. This pooling process was motivated by the fact that estimation of the hyperparameters for a covariance structure in a single voxel is

very noisy and pooling across multiple voxels allows more precise estimation of a covariance structure. Of course, a covariance structure is not perfectly identical across all voxels, so the estimated covariance structure might differ slightly depending on the voxels included in the analysis (e.g. whether only a small region of entorhinal cortex or the whole brain is included). In my experiments, the whole brain was included in the analysis. Once the covariance structure is estimated, a whitening matrix which compensates for the covariance structure is generated and applied to the design matrix. Then the parameters can be estimated using a Least Squares solution. This estimate is now called Weighted Least Squares.

The parameters are estimated for each voxel in the brain, generating a Statistical Parametric Map (SPM). This parameter map can be later used in hypothesis testing using linear contrasts.

2.7.2 T-statistics

It is known that under the null hypothesis $\mathrm{H0}: \mathrm{c}\hat{\beta} = 0$, where c is a row vector of constants, the following statistic

$$\mathrm{t} = \frac{c\hat{\beta}}{\sqrt{\sigma^2 c(X'X)^{-1}c'}}$$

has a t-distribution with degrees of freedoms equal to the number of samples minus the number of parameters. For example, when a GLM contains two regressors or experimental conditions (e.g. “visited a different room” and “visited the same room again”), regression coefficients are estimated for each condition, beta1 and beta2. By using a contrast, c = [1, -1], we can test whether the neural parameter (or response strength) is significantly greater when a participant visited a different room compared to when the participant visited the same room again, using a simple t-test. If a contrast of c = [1, 0] is used, it tests whether the regression coefficient for “visited a different room”

is zero or not. This is effectively a parameter estimate normalised by the residual error (the denominator in the above equation) and I used this as a proxy for neural representations in the later multivoxel pattern similarity analysis.

So far, a GLM was applied to individual participants and parameter estimates were computed for each participant's fMRI time series in each voxel. To make an inference at a population level (e.g. that brain region X shows a greater response when people visit a different room than when they visit the same room again), I used a random effects analysis in SPM (called a second-level analysis). This analysis treats each participant as a random sample drawn from a wider population where between-subject variance is typically large. It simply tests whether an effect (e.g. the difference between visiting the same or a different room, or a neural response when facing 60°) is significantly above zero at the group level using a t-test:

$$\mathrm{t} = \frac{\mu}{SE}$$

where μ is the mean of (linearly weighted) betas across all participants and SE is the standard error of the (linearly weighted) betas across all participants. The degrees of freedom is the number of participants minus 1.

2.7.3 Statistical thresholding

When we report whether some voxels show a significant effect of the experimental manipulation, alpha level of 0.05 is commonly used which controls the false positive rate to be less than 5%. However, the GLM approach described above is applied to multiple voxels in the brain (which is why it is called a "mass-univariate" analysis) which means that about 3,000 voxels can be identified as false positive if the same t-test were used 60,000 times (the number of 3 mm voxels in the brain). This problem is known as the multiple comparison problem. To calculate the p-value adjusted for

multiple comparisons, SPM uses Random Field Theory to estimate the spatial smoothness of the data and measure the number of independent tests within the brain. Because neural responses are often spatially smooth (e.g. if one voxel is activated by face stimuli, its neighbouring voxels are also likely to be activated by face stimuli), the number of independent voxels (called Resel in SPM) is smaller than the number of total voxels. The resultant p-value is known as Family-Wise Error rate (FWE) which is the probability of having at least one false positive among all voxels tested. Throughout my experiments, I report the p-value corrected for the number of voxels within regions of interest (ROIs) or across the whole-brain.

2.8 Repetition suppression analysis

The main methodological limitation of fMRI (and all other non-invasive neuroimaging methods used in humans) is that it measures the gross activity of millions of neurons. When we want to investigate fine neural representations at the cellular level (e.g. head direction cells that fire when a participant is facing North and neighbouring cells that fire when a participant is facing South), the fMRI signal from a single voxel is unlikely to reveal subtle differences between experimental conditions because the sum of mixed responses would be almost identical (Figure 2.7A), unless most cells within a voxel are tuned to particular stimulus, like the primary visual cortex favouring visual stimuli compared to auditory stimuli. However, the spatial limitation of fMRI can be partially relieved by using repetition suppression analysis (or multivariate analysis which I will describe in the next section).

Repetition suppression analysis takes advantage of the phenomenon that neural activity is reduced when the same neural population is repeatedly activated (for a recent review, see (Barron et al. 2016)). This phenomenon, also called adaptation, has been observed in single unit recordings in multiple brain regions including

inferotemporal cortex (Gross et al. 1969), primary visual cortex (Maddess et al. 1988), perirhinal cortex and entorhinal cortex (Xiang and Brown 1998) in animals. Repetition suppression can reveal the tuning of neurons (stimulus selectivity), and not just the effects of novelty. For instance, neurons in inferotemporal cortex showed repetition suppression for an object of varying size and location, suggesting that these neurons encode the identity of the object rather than the pictorial representation of the object on the retina (Lueschow et al. 1994). Importantly, this repetition suppression effect predicts that the fMRI signal would be lower when a stimulus is preceded by the same stimulus (Y→Y) compared to when a stimulus is preceded by a non-identical stimulus (X→Y) even though fMRI responses to the single presentation of stimulus X and stimulus Y are not distinguishable due to the mingled neural responses within a voxel (Figure 2.7A). Thus, we can test whether neural activity contains information about these stimulus types by comparing the response to repeated stimuli and non-repeated stimuli (Figure 2.7B).

Of note, as shown in the previous study which used object stimuli with varying size and locations on the retina (Lueschow et al. 1994), repetition suppression analysis can be used to compare the neural representation between non-identical stimuli that share some features. For example, researchers can manipulate multiple features of stimuli (e.g. object and floor; object A on the 1st floor, object A on the 2nd floor, object B on the 1st floor) and test which feature evokes repetition suppression (e.g. object or floor) to reveal what is encoded in a brain region. Repetition suppression analysis has been widely used in fMRI experiments; to name just a few, invariant object representation in the lateral occipital complex (Grill-Spector et al. 1999), mnemonic interference in the hippocampal formation (Reagh and Yassa 2014) and construction of novel memories during decision making in prefrontal cortex (Barron et al. 2013).

Repetition suppression analysis can be applied to various brain regions and multi-modal techniques (single unit recordings, fMRI, magnetoencephalography (MEG)), but there is a limitation on the experimental designs where this analysis can be used. This analysis is most suitable for a simple experimental design where experimenters can present stimuli in a discrete and specific sequence. For example, when experimenters want to compare the similarity of neural representations of stimuli A, B, and C, the stimuli should be ideally presented in all possible permutations (e.g. A→B, A→C, B→C). This is not possible for spatial navigation tasks in a naturalistic context where a participant's location and direction changes continuously (unless a virtual wormhole is used). Therefore, I only used a repetition suppression analysis in Experiment 2, where the neural representation of each discrete location was discontinuously evoked in a specially designed sequence.

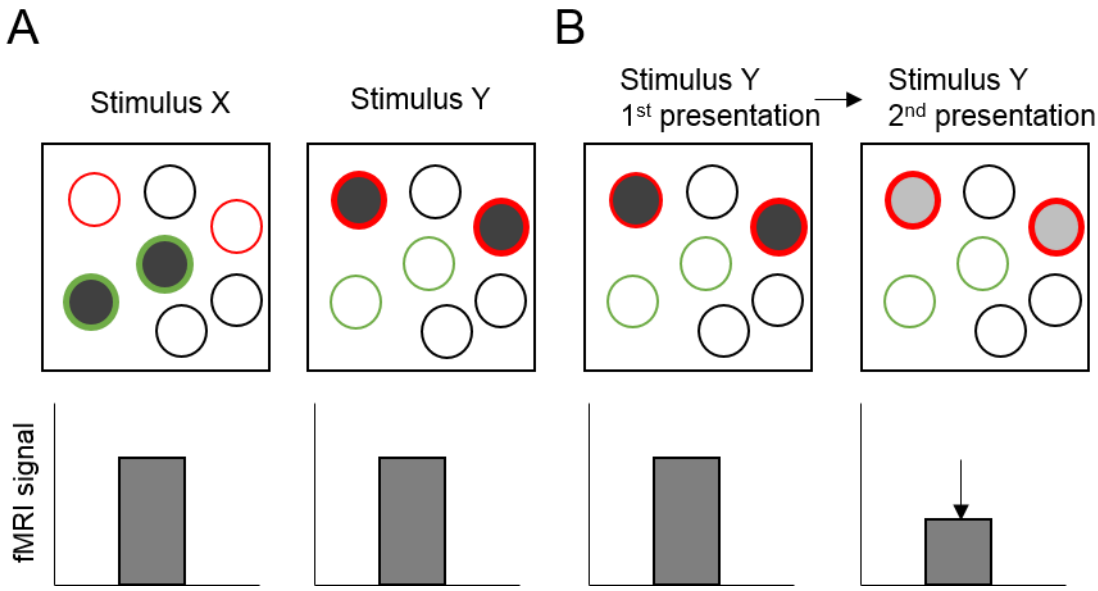


Figure 2.7 The principle of repetition suppression analysis. (A) A typical fMRI study measures the sum of activity of numerous neurons within each voxel. For simplicity, let us assume that there are seven neurons (circles) in one voxel (box) of which two voxels respond to stimulus X (green circles) and stimulus Y (red circles) respectively. In this case, the sum of the neural responses within a voxel as measured by fMRI is identical for stimulus X (the bottom left bar graph) and stimulus Y (the bottom right bar graph) because two neurons are activated for both stimuli. Therefore, if we simply compare the neural responses to these two conditions, we cannot tell whether this voxel contains information about stimulus type. (B) When a population of neurons is reactivated in a short time interval, the neural activity weakens (shown as bright grey in the right graph) due to neural adaptation. Consequently, the fMRI signal is expected to be lower when a stimulus is preceded by the same stimulus (Y→Y) compared to when a stimulus is preceded by a non-identical stimulus (X→Y). Therefore, we can test whether neural activity contains information

about stimulus type by comparing the response to repeated stimuli and non-repeated stimuli, instead of directly comparing the response to one stimulus type or the other.

2.9 Multivariate pattern analysis

Traditionally, fMRI studies have searched for neural activity that depends on experimental manipulations at a single voxel. This is analogous to single cell electrophysiology, which focuses on the selectivity of a neuron to a few experimental conditions, e.g. a place cell that responds when an animal is at a particular location, a neuron that responds most to the picture of Jennifer Anniston (Quiroga et al. 2005). However, with technological advances in multi-electrode recordings, the idea that the brain encodes information in terms of population activity rather than single cell activity has received growing empirical and theoretical support (Averbeck et al. 2006). For instance, individual place cells in ventral hippocampus have larger place fields than the cells in dorsal hippocampus, and consequently were thought to contain less spatial information than the dorsal hippocampus. However, it was found that the location of an animal reconstructed from the population activity of ventral hippocampus was as precise as those reconstructed from the dorsal hippocampus (Keinath et al. 2014).

In parallel to a growing interest in population coding, multivoxel (or multivariate) pattern analysis became popular in fMRI research. Multivoxel pattern analysis assumes that neural information (or representations) is present in the spatially distributed activity patterns across multiple voxels. Biases in neural populations in each voxel are a prerequisite for a multivariate analysis, e.g. voxels 1, 2, 5 in primary visual cortex contain slightly more neurons tuned to a 30° visual grating and voxels 3, 4 contain slightly more neurons tuned to a -30° grating.

One application of multivariate analysis is decoding or classification of experimental conditions, sometimes known in the popular media as “mind reading” (Norman et al.

2006). This classification approach uses machine learning techniques (e.g. support vector machines, neural networks, nearest-neighbour classifiers) to test whether each experimental condition (e.g. object X and object Y) evokes a distinct multivoxel pattern (Haxby et al. 2001). Another approach is representational similarity analysis, or RSA (Kriegeskorte et al. 2008). RSA focuses on how the neural representational similarity between experimental conditions is related to the similarity measures predicted from multiple theoretical models. These two approaches are not mutually exclusive. One can test multiple theoretical models by using different classifiers, and one can classify the experimental conditions using a binary model representational similarity matrix. However, these approaches have different philosophies; the classification approach can be viewed as a first-order isomorphism that seeks a one-to-one correspondence between the neural representation and experimental conditions, whereas RSA uses a second-order isomorphism. RSA is particularly suitable for testing common hypotheses in the data acquired from different modalities (e.g. electrophysiology, fMRI, MEG) which have a different number of sensors and resolution so a simple one-to-one correspondence cannot be achieved. An implementation of RSA is simpler than the classification approach which involves advanced machine learning techniques. Furthermore, it has been suggested that RSA, which uses a continuous similarity metric between neural representations, can more reliably reveal neural information than the classification approach which uses discretised classification accuracy (Walther et al. 2016). For these reasons, I used a pattern similarity analysis to test multiple spatial encoding hypotheses in Experiment 1 and Experiment 3. I explain the basic analysis flow of RSA below.

The first step is to estimate the neural representation for each experimental condition of interest (e.g. location A, location B, facing 30° vertically, facing 30° horizontally). I defined the neural representation as multivoxel patterns associated with each experimental condition. I used a first-level GLM to account for the hemodynamic delay

when estimating neural responses to each experimental condition from the fMRI data. I estimated the responses for each scanning session separately, e.g. a neural pattern for location A in session 1, a neural pattern for location A in session 2. Either the raw regression coefficients (beta) or variance normalised T values can be used as a proxy for neural representations, but a methodological review paper has suggested that T values outperform beta values in multivoxel pattern classification (Misaki et al. 2010). This could be due to normalisation of noisy voxels that have high beta values. Consequently, I used T values in my experiments. Some studies subtract mean patterns of multiple experimental conditions to eliminate the common activation effect, but the mean subtraction can induce arbitrary correlations and obscure the interpretation of results (Garrido et al. 2013), so I used raw T values without any mean subtraction.

Once multivoxel patterns for each experimental condition are estimated, similarity values between every pair of experimental conditions are computed. The two most commonly used similarity measures are Pearson correlation and Euclidean distance. I used Pearson correlation because I assumed that the relative signal intensity between voxels contains more important information than the absolute signal intensity difference captured by Euclidean distance. Absolute signal intensity is also more vulnerable to baseline signal fluctuations between scanning sessions. I always calculated the Pearson correlation between experimental conditions in different scanning sessions (e.g. location A in session 1 – location B in session 2) to minimize the potential temporal confounds when comparing the neural responses to different experimental conditions within one scanning session.

I averaged the correlation values between all pairs of scanning sessions; e.g. similarity between location A and location B was the average of similarity between [location A in session 1, location B in session 2], [location A in session 1 – location B in session 3],

[location A in session 2 – location B in session 1], [location A in session 2 – location B in session 3], [location A in session 3 – location B in session 1], [location A in session 3 – location B in session 2]. I also computed the similarity within the same condition across scanning sessions; e.g. location A in session 1 – location B in session 2. This within-condition similarity measures how reliable the neural representation is across multiple scanning sessions, and this within-similarity should be higher than similarities between different experimental conditions if the multivoxel response patterns contain information about the experimental conditions.

Next, neural representational similarity values are compared to similarity values predicted from multiple encoding hypotheses. There are numerous ways of comparing empirical data to theoretical predictions. One is to compare the raw pattern similarity values between different experimental conditions using t-tests, based on hypotheses (e.g. the horizontal-weighted anisotropic 3D encoding hypothesis in Experiment 1 predicts the pattern similarity would be as follows: within-same location > between vertically adjacent locations > between horizontally adjacent locations). I used this simple comparison approach in Experiment 1.

Another approach is to compare the neural similarity values to theoretical similarity values using a correlation or regression analysis. This approach is more appropriate when continuous variables or multiple levels of similarity values are predicted from hypotheses. This was the case in Experiment 3, where neural similarity values were modelled as a function of the difference in angle between each heading direction (the differences ranged from 0 to 120°). In my experiment, and many previous studies (Carlin et al. 2011; Vass and Epstein 2013; Marchette et al. 2014; Chadwick et al. 2015), regression coefficients or correlations that are significantly greater than zero at the group level were taken as evidence showing that multivoxel responses in a brain region contain information that matches a theoretical model. However, one should be

aware that this is only a low bar for a complete explanation of neural representations in a brain region (Nili et al. 2014). Multiple models can explain some part of the variance in the data, and neural data always have many unexplained sources of variance, especially in non-sensorimotor areas like the hippocampus. Thus, when I claim “this brain region X’s activity was significantly or best explained by hypothesis Y” in this thesis, I do not rule out the possibility that alternative models that have not been specifically tested in my experiment could better explain the data. This issue, that only a limited hypotheses space can be tested in one experiment, applies to any empirical research field.

2.10 Statistical tests

Statistical tests for fMRI GLM were conducted as described in Section 2.7.3. All other statistical tests were computed in MATLAB (R2014a, MathWorks, MA) or IBM SPSS 22 (IBM, NY) using threshold of 0.05. Data are presented with the mean ± 1 standard deviation (SD), unless otherwise specified. Error bars are standard error of the mean (SEM), unless otherwise specified.

In the next four chapters, I describe my experiments where participants explored various virtual 3D environments and their brain responses were analysed using the methods described here.

Chapter 3 Experiment 1: A semi-volumetric space

3.1 Introduction

In Chapter 1, I summarised how 2D place and direction information in animals is encoded in spatially-sensitive cells such as place cells and head direction cells (O'Keefe and Dostrovsky 1971; Taube et al. 1990), and the evidence for this spatial information in humans in brain regions including hippocampus (Hassabis et al. 2009; Sulpizio et al. 2014), retrosplenial cortex (Baumann and Mattingley 2010; Vass and Epstein 2013; Marchette et al. 2014; Sulpizio et al. 2014; Shine et al. 2016) and entorhinal cortex (Chadwick et al. 2015). In contrast, I also noted in Chapter 1 that 3D space encoding in the brain has been considered in only a few animal studies, with discrepancies between their findings. For example, Hayman et al. (2011) found that place cells in the hippocampus expressed less information about the vertical axis compared to the horizontal axis when rats moved on a vertical wall and helix staircase. This neural finding, and asymmetry in vertical and horizontal navigation behaviours (Grobéty and Schenk 1992a; Jovalekic et al. 2011), led some to propose that 3D space representation might be fundamentally anisotropic, such that 3D space is divided into the main plane of locomotion (e.g. horizontal) and the axis orthogonal to the plane (e.g. vertical) (Jeffery et al. 2013). In contrast, most place cells recorded in the hippocampus of flying bats were equally sensitive to all three axes (Yartsev and Ulanovsky 2013), and head direction cells recorded in the presubiculum of crawling or flying bats were tuned to either horizontal (azimuth), vertical (pitch), or a combination of azimuth and pitch, thus providing a 3D compass required for a complete, isotropic 3D map (Finkelstein et al. 2015, 2016).

But what about humans - is 3D space anisotropically represented as it appears to be in rats, or isotropically represented as in bats? Are different brain regions involved in

vertical and horizontal spatial processing as suggested by two previous neuroimaging studies that I described in Chapter 1 (Indovina et al. 2016; Zwergal et al. 2016)? To answer these questions, in this first experiment I investigated whether and how 3D location and direction information is represented in the human brain using an fMRI virtual navigation paradigm and multivoxel pattern analysis. The aim was to adjudicate between the following hypotheses: (1) vertical and horizontal information is similarly represented in a brain structure (isotropic encoding); (2) either vertical or horizontal information is represented with greater sensitivity in one or more brain regions (anisotropic encoding); or (3) vertical and horizontal information is represented in separate brain areas (2D planar encoding).

When testing these hypotheses, the choice of 3D environment and experimental paradigm is pertinent because the mode of movement can determine the dimensionality and the shape of neural representations (Finkelstein et al. 2016). Based on the degrees of freedom of movements, there is a spectrum of 3D environments ranging from a multi-level building, where one mainly moves horizontally and only occasionally moves vertically, to a microgravity environment where one can freely move along all 3 axes. I elected to start with an intermediate type of environment, a semi-volumetric lattice virtual reality environment, where participants could move along flat, tilted-up or tilted-down pathways. I chose this approach because, first, both vertical and horizontal information is equally well accessible in this lattice environment, allowing for a fair comparison between them. The helical staircase (Hayman et al. 2011) or multi-level building (Zwergal et al. 2016) used in the previous experiments in rats and humans respectively, did not have matched sensory inputs between the vertical and horizontal conditions. This was because the different vertical locations were visually similar due to the repeated nature of the structures (note that a repeated structure is in itself an interesting topic for consideration in relation to 3D spatial representation, and I will address this further in Experiment 2). Second, under normal gravity conditions, a

human's movement is bound to a surface rather than a complete volumetric space. Therefore, the semi-volumetric 3D lattice environment was a more ecological and reasonable starting point to study 3D spatial representation in the human brain (note that a volumetric environment will be tested in Experiment 3).

As a priori ROIs in this first experiment, I focussed on the hippocampus, retrosplenial cortex and entorhinal cortex because place cells and head direction cells have been found in these regions in animals in 2D space, and numerous fMRI studies have also reported place and direction information in these structures in 2D virtual environments.

3.2 Methods

3.2.1 Participants

Thirty six healthy right-handed adults took part in the experiment (18 females, mean age 24.2 years, standard deviation (SD) 4.25 years, range 19-34 years). All had normal or corrected-to-normal vision and gave informed written consent in accordance with the local research ethics committee.

3.2.2 The virtual reality environment

The virtual environment was a lattice structure which conveyed the sense of an open 3D space, like a junglegym in a playground (Figure 3.1A). It comprised four levels and each level contained 4 x 4 nodes which were linked to neighbouring nodes by narrow pavements (horizontal or sloped) or wooden pillars. From where a subject stood on a node, they could move along the pavements to one of six neighbouring nodes: four on the same horizontal plane and two on different floors. From an egocentric perspective, they could move straight forwards or straight backwards (in the latter case they would

turn 180° to approach the node behind them), diagonally on the same floor forwards or backwards, or they could move up or down via a slope (Figure 3.1C). The lattice was enclosed by tall concrete walls without a ceiling. All walls looked identical except that one contained a green door which acted as a unique landmark. During the experiment, the participants could only occasionally see the green door, and so they could not use a simple landmark matching strategy to know where they were; instead they had to carefully keep track of their location throughout the experiment.

The horizontal (x-axis, see Figure 3.1A for the axes convention) and vertical (z-axis) distance between two adjacent nodes was made identical in order to test the isotropy of vertical and horizontal space representation in the human brain. The distance along the y-axis was set to be 1.9 times larger than the other two distances to make the slope of the pavement 29°. This angle was chosen in order to preserve ecological validity, because this has been reported to be the steepest slope a human can walk up at a normal pace (Kinsella-Shaw et al. 1992). The virtual environment was implemented using Unity 4.6 (Unity Technologies, CA, United States). A first-person perspective was used and the field-of-view was ±30° for the vertical axis and ±45.7° for horizontal axes. A snapshot of the 3D lattice as seen from a participant's perspective is shown in Figure 3.1D. During pre-scan training, the stimuli were rendered on a standard PC (Dell Optiplex 980, with an integrated graphic chipset) and presented on a 20.1 inch LCD monitor (Dell 2007FP) with a screen resolution of 1600 x 900. The same PC and resolution were used during scanning. The stimuli were projected on a screen using an Epson EH-TW5900 projector at the back of the MRI scanner bore, and participants saw the screen through a mirror attached to the head coil. The screen covered a field of view of ~21° horizontally and ~12° vertically.

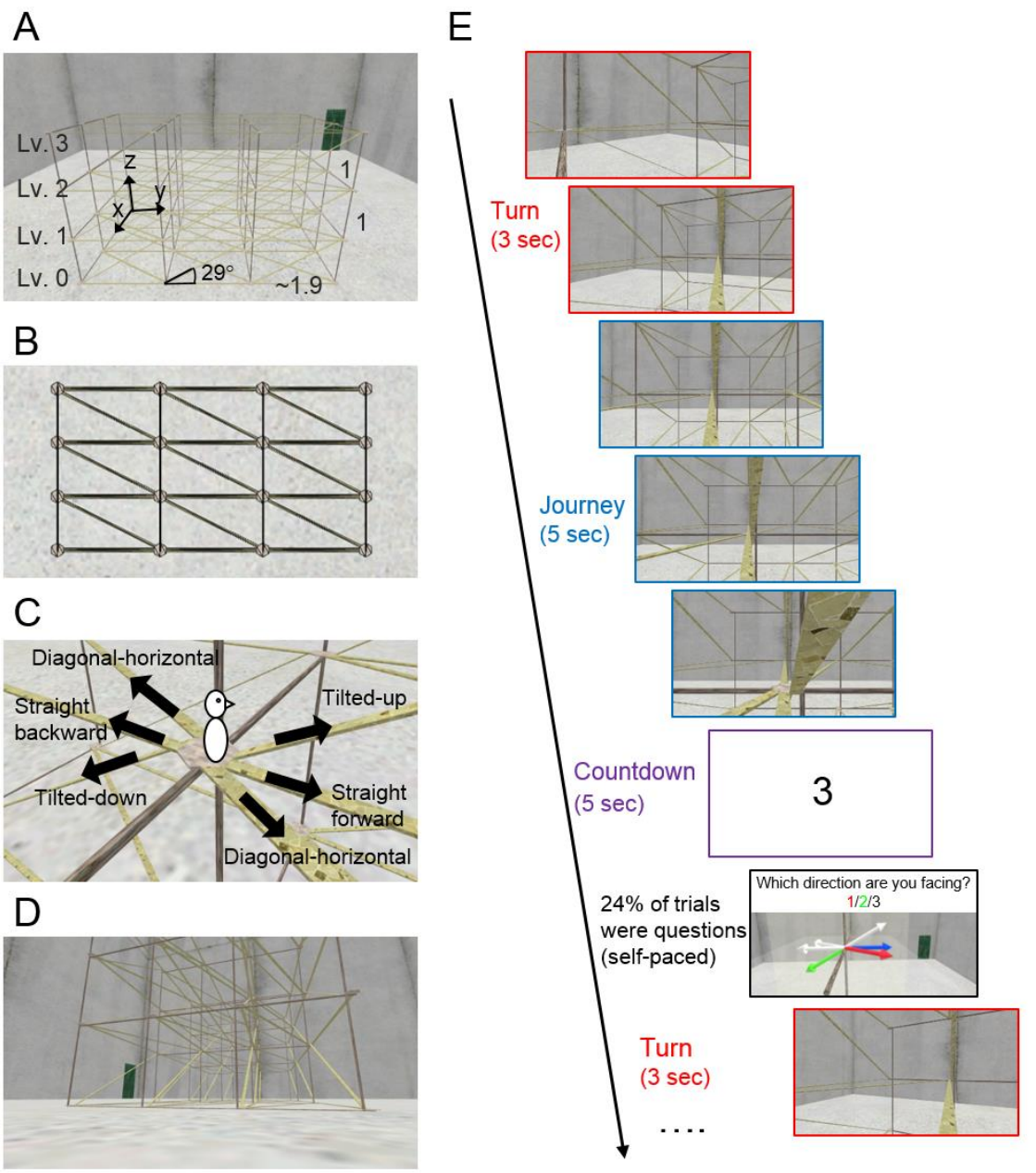


Figure 3.1 Experimental design. (A) An overview of the virtual environment. A 4 x 4 x 4 lattice structure was enclosed by tall concrete walls. One of the walls contained a green door as a unique landmark. The horizontal (x-axis) and vertical (z-axis) distance between two adjacent nodes was made identical to test the isotropy of 3D space encoding. The distance along one horizontal axis (y-axis) was 1.9 times longer than other distances because the vertical slope was designed to be walkable at 29°. (B) An overhead view of the lattice. (C) A close-up view of 6 pavements around a centre node. From where a subject stood on a node, they could move along the pavements to one of six neighbouring nodes: four on the same horizontal plane and two on different floors. (D) An example view of the lattice structure from a subject's perspective during the free exploration phase prior to scanning. (E) Example of a trial during scanning. Participants continuously moved from one node to another as if they were riding a rollercoaster. Each trial began with a turn at the node, followed by a linear movement on the pavement (journey), then a countdown screen. In some trials, a place or direction question was presented before the next trial began. An example direction question is shown here. See Figure 3.2A for place questions.

3.2.3 Procedure

Each participant completed the experimental tasks in the following order: free exploration prior to scanning, one practice of the experimental task before scanning, the experimental task during scanning and a post-scan debriefing session.

Free exploration prior to scanning

Having watched a short demonstration of the experimental task, participants freely explored the virtual environment using a keyboard in a testing room. During this self-paced exploratory period (mean duration 734 sec, SD 300 sec), participants volitionally moved along the pavements and visited all four floors of the lattice structure. The position and heading direction of participants were recorded every 0.1 sec. Because the pavements linking each node were designed to be narrow, most participants “fell” down from the lattice at least once during the exploration. I welcomed this experience because it allowed participants to appreciate the height, maximizing the sense of 3D space. However, to prevent height-related anxiety from influencing the task, participants were told that they would move along a pre-programmed route during the scanning experiment without falling off. I later confirmed in the debriefing session that 94% of the participants were not at all anxious during the scanning experiment. Participants did not practice in advance the exact routes that would be experienced in the scanner because different pseudo-randomised routes that were optimised for sampling each direction and place were used during scanning.

Scanning task

In the scanner, participants moved along a pre-programmed route in the 3D lattice structure as if they were riding a rollercoaster (this was practised before entering the scanner). This constrained-movement approach had advantages over unconstrained free exploration or the use of static picture stimuli. Compared to the latter, the

rollercoaster experience provided participants with a strong sense of being in a 3D space (see the debriefing results), thereby allowing a more ecological investigation of space representation in the human brain. This approach also permitted precise control of the movement trajectory for every subject, which cannot be achieved if participants are moving freely.

Although there were 4 levels in the virtual environment, only the inner 8 nodes on the middle two floors were used for the analysis because the ground level and the top floor were quite distinctive in physical appearance. To increase the number of visits to these inner nodes, the rollercoaster moved between these inner nodes on 76% of the trials. I used movement sequences in which each of the inner eight nodes and directions of interest were sampled with similar frequency (see section 3.2.7), allowing an unbiased and reliable estimation of 3D place and direction representations. The routes were presented in a randomised fashion across subjects. To ensure subjects paid attention during the task, they were occasionally asked about their current position and direction (see below, and Figure 3.1E for the timeline of an example trial).

A pre-programmed route during one scanning session was composed of 50 consecutive movements (trials) from one node to an adjacent node in the lattice. On each trial the rollercoaster prepared to move at the initial node by turning towards the next node (‘turn’). A constant angular velocity was applied during this 3 sec turn and the instantaneous direction was a linear interpolation between the initial 3D direction vector and the next directional vector. After the turn, the rollercoaster underwent a linear movement along the pavement (‘journey’) at a constant speed to the next node, which took 5 sec. Participants’ viewing angle, equivalent to their head direction, was parallel to the pavement. This meant that when they were moving up by 29°, head pitch was also 29°. Having arrived at the destination node, the virtual environment was temporarily hidden by a white countdown screen for 5 sec (‘countdown’). In the majority

of trials (76%), the next trial started straight after the countdown. In 24% of the trials, a question was presented before the next trial and the subject indicated their current position or direction on a map using a keypad with the right hand (Figure 3.1E). These occasional questions were included to maintain participants' attention and to compare the behavioural sensitivity of encoding the vertical and horizontal dimensions (see Section 3.2.4 below). The question period also helped participants to maintain the correct sense of direction throughout the experiment because when participants answered incorrectly, the correct place or direction was shown on the screen. In total, one session of 50 consecutive trials lasted ~13 minutes. Participants completed four scanning sessions with a short break between each session, making a total functional scanning time of ~50 minutes.

Post-scan debriefing session

After scanning, participants were asked about how much they felt immersed in the virtual environment with the following options: "I felt like I was really there", "I occasionally thought about the environment as being on a computer screen, but overall the environment was convincing and I felt I was moving straight, up or down", "I was often distracted by the feeling that I was not in a real environment". They also reported whether the height made them anxious or nervous during the scanning task with 3 options: "Not at all", "Somewhat", "Very".

Importantly, participants were also asked (without prior notice) to estimate the length and angle of the virtual environment in terms of meters and degrees based on their experience of navigating within it. Although the vertical and horizontal distance between two nodes were made equivalent, participants' subjective perception of distance or direction could be different from the true physical distances due to the horizontal-vertical visual illusion (Avery and Day 1969), and this might influence the neural encoding of the vertical and horizontal dimensions. For instance, the neural

representation of the straight heading direction and the vertically 29° tilted direction could be more dissimilar than the straight direction and horizontally 29° tilted direction if participants overestimate the vertical slope compared to the horizontal angle. In addition to the quantitative estimate of vertical/horizontal size, I also asked about participants' qualitative impression of the size of the whole environment (options: small/medium/large) because spatial scale-dependent representation has been associated with the hippocampus (Evensmoen et al. 2015).

3.2.4 Behavioural analysis

Performance during the scanning task

Both place and direction questions were three-alternative forced choice, meaning that chance accuracy was 33%. In the place question, the positions of the two distractors varied systematically enabling me to compare the behavioural sensitivity of horizontal and vertical encoding (Figure 3.2). In the 'with-V' condition, one distractor was above or below the correct position (vertical distractor) and the other distractor was on the same floor, adjacent to the correct position. In the 'without-V' condition, all three choices were on the same horizontal plane such that one distractor was adjacent to the correct position along the short horizontal axis and the other along the long horizontal axis. If the vertical and horizontal axes were equally well encoded, performance for both conditions should be similar. On the other hand, if the vertical axis was poorly encoded relative to the other two horizontal axes such that the participants were more confused by a distractor above or below the true position, performance for the with-V conditions would be worse than for the without-V condition, which did not involve a vertical distractor. I compared the response time and accuracy of these conditions using paired t-tests.

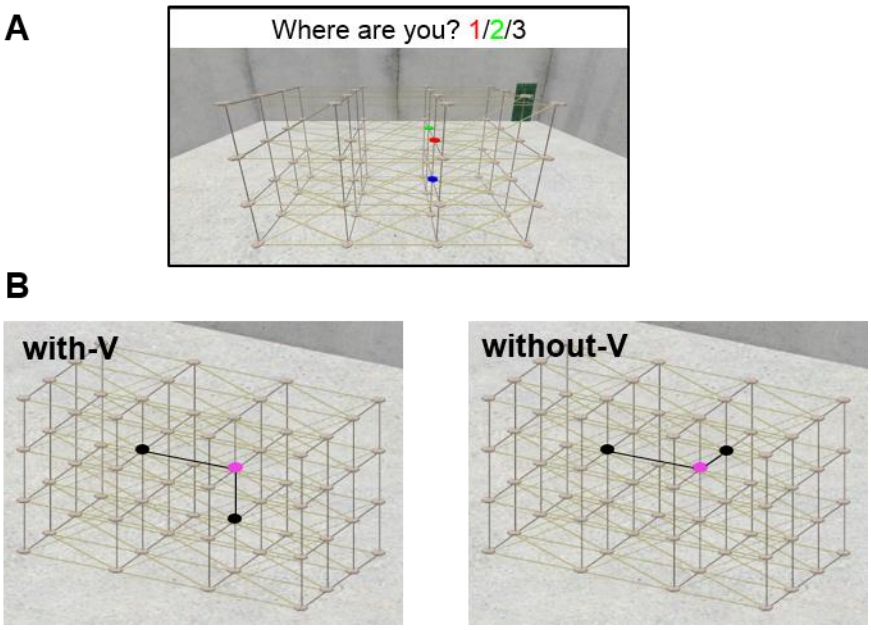


Figure 3.2 Behavioral analysis of vertical and horizontal place encoding. (A) Participants occasionally indicated their location from 3-alternative forced choice questions as shown here. (B) The positions of the distractors in the place questions were varied systematically. For simplicity of explanation, a correct node here is shown in pink and the two distractors are shown in black in orthogonal projections of the 3D lattice. In the with-V condition, one distractor was above or below the correct node and the other distractors were adjacent to the correct position on the same horizontal plane. In the without-V condition, both distractors were on the same floor as the correct location.

The direction questions were included to motivate the participants to concentrate on both place and direction, and there was no variation in the distractors. Rather, I compared the response time and accuracy when the correct direction had a non-zero vertical pitch component (direction J/M in Figure 3.4B) and when the pitch of the correct direction was zero (direction I/K/L/N in Figure 3.4B) to test whether vertical pitch was more distinguishable.

Post-scan debriefing session

I counted the number of responses for each option in the multiple choice debriefing (i.e., participants' engagement in the virtual environment, emotional state and qualitatively perceived size of the environment). A quantitative size estimate of the vertical and horizontal dimensions was analysed using a t-test. A ratio of the perceived vertical and horizontal distance and angle was tested against a true ratio of 1 using a one sample t-test.

3.2.5 Scanning and image processing

Functional MRI scans (voxel size = 3 x 3 x 3 mm, 48 slices, TR = 3.44 sec) and a structural MRI scan (1 x 1 x 1 mm) were acquired using a 3T Siemens Trio scanner and preprocessed using SPM12. The scanning parameters and preprocessing are described in detail in Section 2.5.4. To summarise, after removing the first 5 volumes to allow for T1 equilibration effects, functional images were realigned to the first volume of each scanning session and geometric distortion was corrected by the SPM unwarp function using the fieldmaps. Each participant's anatomical image was then coregistered to the distortion corrected mean functional images. Functional images were normalized to MNI space, and were left unsmoothed for multivoxel pattern analysis to preserve the fine scale activity patterns.

3.2.6 Anatomical ROIs

I defined three anatomical ROIs for areas known to contain spatial information: hippocampus, retrosplenial cortex and entorhinal cortex. Each ROI was manually delineated on the group average structural MRI scan (1 x 1 x 1 mm) using ITK-SNAP (www.itksnap.org). The ROIs were then resampled to the resolution of the functional scans (3 x 3 x 3 mm).

The hippocampus was divided into anterior and posterior ROIs given the literature showing anatomical and functional variation along its long axis (Poppenk et al. 2013; Strange et al. 2014). A coronal coordinate (y = -19 mm) which approximates the position of the apex of the uncus on the group average structural MRI scan was used to divide anterior and posterior hippocampus. Of note, the hippocampus ROI in this study included CA1/2/3, dentate gyrus and the subiculum proper, but it did not extended medially to the pre/parasubiculum (see Section 1.8 for the neuroanatomy).

Although some navigation fMRI studies have defined a “retrosplenial complex” which includes Brodmann areas 29-30, occipitotemporal sulcus and posterior cingulate cortex using a functional localizer (Vass and Epstein 2013; Marchette et al. 2014), I used a more precise anatomical definition of retrosplenial cortex, based on cytoarchitecture, which includes only Brodmann areas 29-30 (Vann et al. 2009). Functionally defined ROIs vary from one study to another depending on the statistical threshold and individual differences. Anatomical retrosplenial cortex and functionally defined retrosplenial complex overlap, and whether an anatomical or functional definition is more appropriate depends on the specific research question. In this study, I was interested in searching for 3D spatial information with precise anatomical priors in retrosplenial cortex, where spatially modulated cells have been reported in animals, so I elected to use this conservative and threshold-free anatomical definition.

The entorhinal ROI was defined following the protocol in Pruessner et al. (2002). I defined the caudal end of entorhinal cortex as 2mm posterior to the uncal apex (y = -21 mm) following this protocol, but note that some studies have used a more posteriorly-extended definition of entorhinal cortex (Chadwick et al. 2015). While the entorhinal cortex mask was derived so as not to extend into the neighbouring pre/parasubiculum, there existed a chance that one or two voxels in the neighbouring pre/parasubiculum were included in the entorhinal ROI due to spatial resolution and individual differences. The entorhinal cortex is challenging for fMRI researchers due to substantial signal loss induced by susceptibility artefact in this region. I assessed the temporal signal to noise ratio (tSNR) defined as the mean of the fMRI time series divided by the standard deviation in every voxel of the ROI (after spatial normalisation). As expected, tSNR was much lower in entorhinal cortex (19.8 ± 6.6) compared to the whole hippocampus (61.1 ± 8.9) and retrosplenial cortex (56.2 ± 9.9), implying that entorhinal cortex could be disadvantaged in expressing its function. Anatomical localisation of the ROIs can be seen in Figure 3.3. The number of functional MRI voxels within each ROI was: left

anterior hippocampus 54; right anterior hippocampus 61; left posterior hippocampus 98; right posterior hippocampus 91; left retrosplenial cortex 158; right retrosplenial cortex 135; left entorhinal cortex 47; right entorhinal cortex 49.

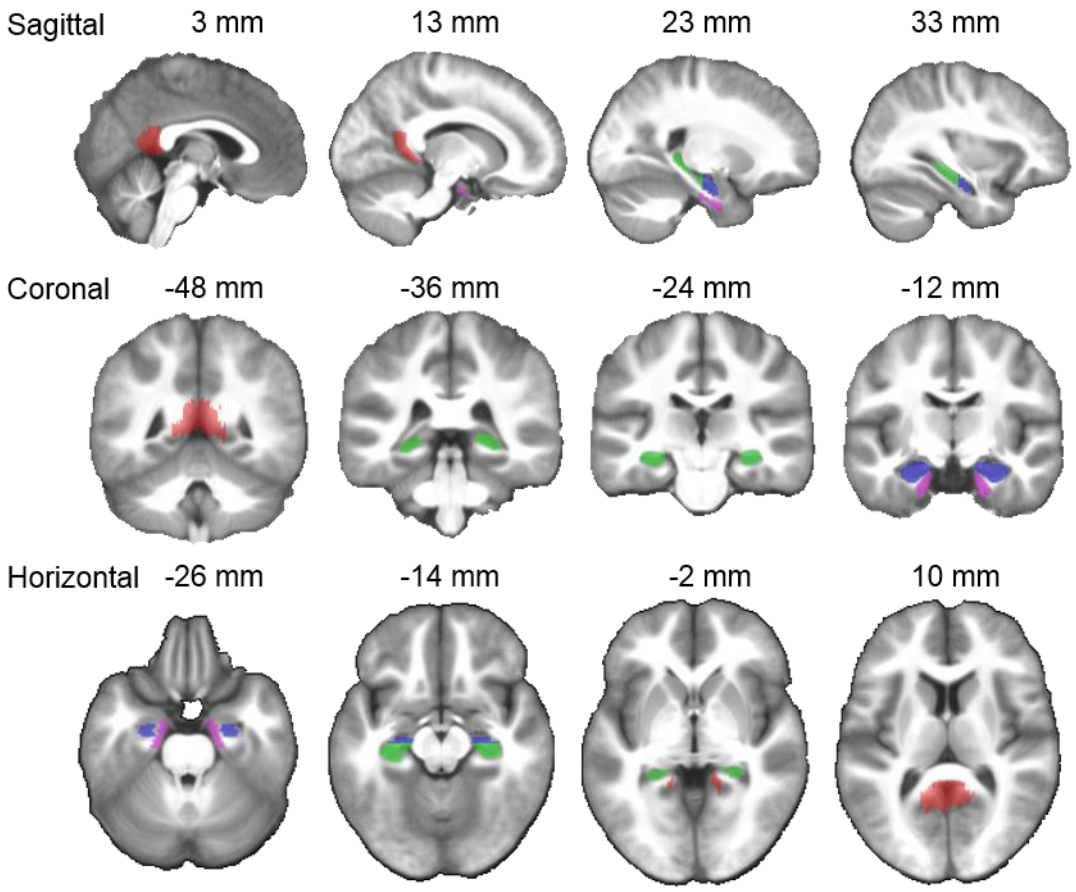


Figure 3.3 Anatomical ROIs. Anterior hippocampus (blue), posterior hippocampus (green), entorhinal cortex (pink) and retrosplenial cortex (red) masks are overlaid on the group averaged structural scan.

3.2.7 Multivoxel pattern analysis

3D space encoding hypotheses

To adjudicate between the different 3D space encoding hypotheses (isotropic 3D, anisotropic 3D or planar), I compared the amount of vertical and horizontal spatial information in each ROI using multivoxel representational similarity analysis (Haxby et al. 2001; Kriegeskorte et al. 2008). My assumption was that if both vertical and horizontal dimensions were encoded with equal sensitivity (*isotropic*, Figure 3.4C), the

neural representation of two points along the vertical axes should be as distinguishable as those of two points along the horizontal axes, given that the distance between these two points is equivalent.

Therefore, in Figure 3.4A, fMRI multivoxel pattern similarity between place A and B (different vertical, 'diff-V') would be comparable to the similarity between place A and C (different horizontal, 'diff-H'), and obviously both should be lower than the within-place (A and A) pattern similarity ('same'). If the vertical axis is poorly encoded compared to the horizontal axis (*anisotropic, horizontal weighted,* Figure 3.4C), the two places along the vertical axes would be less distinguishable than the two places along the horizontal axes. Therefore, the pattern similarity between place A and B (diff-V) would be larger than the similarity between A and C (diff-H). By contrast, if the horizontal axis is encoded with low sensitivity (*vertical weighted, not shown*), the pattern similarity of the diff-H condition would be larger than the diff-V condition. In the extreme case when only the horizontal (or vertical) dimension is encoded (*pure horizontal or pure vertical,* Figure 3.4C), the neural representation of two places that share the same horizontal (or vertical) coordinates would be completely indistinguishable, so the pattern similarity for the two positions along the vertical axis, diff-V condition, (or the two positions along the horizontal axis, diff-H condition) would be comparable to same place condition. If neither horizontal nor vertical information is encoded, neural responses to each location will be random and inconsistent, so there would be no systematic differences between the same, diff-V, diff-H conditions. An analogous analysis was used to test for the existence and quantity of horizontal (azimuth) and vertical (pitch) direction information in the ROIs.

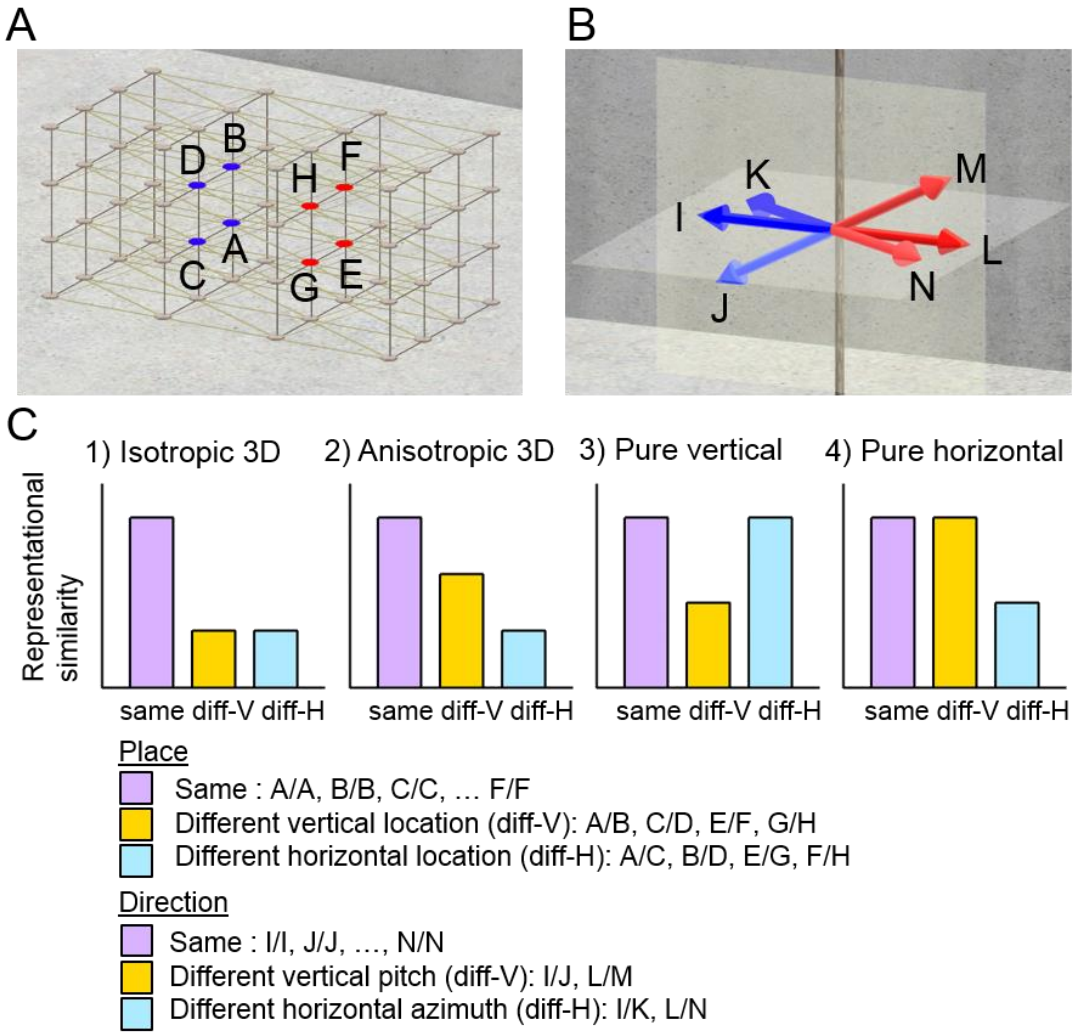


Figure 3.4 Place and direction encoding hypotheses. (A) The places of interest. The inner 8 nodes on the middle two floors were used because the ground level and the top floors were visually distinctive. (B) There were six heading directions. (C) Hypotheses: 1) The isotropic 3D encoding hypothesis predicts that the vertical and horizontal axes are symmetrically encoded. Thus, two places along the vertical axes (diff-V, e.g. A and B) are equally distinguishable as the two points along the horizontal axes (diff-H, e.g. A and C), resulting in equal representational similarity for diff-V and diff-H conditions that are smaller than the same location condition (same, e.g. A and A). In the case of direction encoding, two directions that have different vertical pitch components (diff-V, e.g. L and M) would have similar pattern similarity as two directions that have different horizontal azimuth components (diff-H, e.g. L and N). 2) The anisotropic, horizontal-weighted hypothesis predicts higher pattern similarity for diff-V than diff-H because the neural response is less sensitive to the vertical change than to the horizontal change. 3) A pure vertical encoding hypothesis predicts that as long as the vertical coordinate is the same, the neural pattern will be equivalent even if the horizontal coordinate is different, therefore diff-H is comparable to the same. 4) A pure horizontal encoding hypothesis predicts the opposite, that diff-V is comparable to the same.

Analysis protocol

The first step for the representational similarity analysis was to estimate neural representations (multivoxel patterns) for each place and direction in the virtual 3D

lattice structure. As mentioned previously, although there were 4 levels in the virtual environment, only the inner 8 nodes on the middle two floors were used for the analysis because the ground level and the top floor were quite distinctive in physical appearance (Figure 3.4A). Therefore, the 4 nodes marked in blue were usually approached from 3 directions marked in blue and the other 4 nodes marked in red were approached from the 3 directions in red (Figure 3.4A,B). I estimated the unique multivoxel pattern activity for each place x direction pair (8 x 3 = 24) for each scanning session and each participant using the SPM. My GLM contained 24 place x direction regressors which modelled the journey+countdown period of 10 sec for each of the 4 scanning sessions. I used the whole journey+countdown period because the participants reported that they thought about where they were moving from the beginning of the journey period and they had to maintain this spatial information until the end of countdown.

In addition to the 24 regressors of interest for each session, nuisance regressors were included in the GLM; one for modelling the trials when the participants visited outside the inner eight nodes, two regressors for modelling the occasional place and direction question periods, six regressors for head motion realignment and a constant regressor for each scanning session to account for mean signal variation. In summary, the resulting t-statistics for each voxel in the ROIs were the estimates of multivoxel activations when the subjects were facing those directions and moving towards and standing in those locations, with the hemodynamic delay being taken into account. The second step was to calculate the similarity of multivoxel patterns for each place and direction combination using Pearson's correlation coefficient and to compare the similarity to the 3D encoding hypotheses described in Figure 3.4C.

Because the 4 nodes shown in blue and the 4 nodes shown in red in Figure 3.4.A were approached by 3 different directions as described above, I restricted the analysis to

within either blue or red nodes to fully control the direction and place factors, and then averaged the two similarity matrices later. As a result, a 12 x 12 pairwise correlation matrix was created for each subject. Importantly, I cross-validated the similarity measure across scanning sessions to ensure the independence of each dataset and to estimate a non-biased similarity measure (e.g. the similarity between place A-direction 1 and place B-direction 2 was the mean of the correlation between the place A-direction 1 in session 1 and place B-direction 2 in session 2 and the correlation between session 1 and session 3, between session 1 and session 4, etc.) Each pairwise similarity measure was then grouped into 3 categories: same (e.g. A and A), different vertical (diff-V, e.g. A and B), different horizontal (diff-H, e.g. A and C) (Figure 3.4C).

For the place encoding analysis, the pairs of place x direction combinations which shared the same direction were excluded in order to control for the direction factor. By excluding the same direction pair, the neural representation similarity between the same place could not be due to mere visual identity. Rather, if the neural representation was more similar for the same place compared to another place, it could be interpreted as evidence for place encoding that is generalisable across different directions and different scanning sessions. For the direction encoding analysis, the same place pairs were excluded to test for the existence of direction information that was independent of place.

Finally, the mean pattern similarities for each of the three categories (same, diff-V, diff-H) were compared at the group level with a one-way repeated measures ANOVAs and post-hoc t-tests with Bonferroni correction. I report one sided p-values for comparisons between same and diff-V or same and diff-H because the pattern similarity for the same spatial condition should be higher than those for a different condition if place or direction information is present. I then plotted the bar graphs of this mean pattern

similarity so that they can be easily compared with the 3D place and direction encoding hypotheses that are shown in Figure 3.4C.

3.2.8 Control analysis: visual texture similarity

I designed the virtual environment with a limited palette of colours and textures and only one salient landmark (even this green door landmark was rarely visible during the experiment) in order to minimise the influence of visual inputs when investigating place or direction information. Nevertheless, visual input differed second-to-second because participants solely relied on visual information to track their position and direction. The floor in particular was important as it provided an unambiguous and reliable sense of 3D space. Thus, I applied a slightly brighter colour to the floor, and the proportion of the floor included in the field of view varied depending on whether the participants were heading straight, or up/down (Figure 3.5). To examine whether the place or direction information found in the brain in the main analysis was explained by these visual differences, I conducted a control analysis that used partial correlations to compare the neural similarity data to the place or direction encoding hypotheses, while controlling for this visual component.

To do this, I measured pairwise visual similarity between each place and direction using a simple visual texture model (Renninger and Malik 2004), in the same way as neural pattern similarity was calculated (e.g., visual similarity between place A-direction 1 and place A-direction 2, the similarity between place C-direction 1 and place D-direction 2). Images captured at every half a second during the 5 sec journey period of each place x direction combination were averaged and entered into the texture model. The model applied Gabor filters of varying orientation and size to extract the common textures, and then a distribution of textures across the different pairs of images was compared using the chi-square distribution. The chi-square distance metric was

converted to a similarity measure by subtracting each chi-square value from the maximum chi-square value. Consequently, I obtained "observed neural similarity" and "visual similarity" variables for each subject.

Next, I created “predicted neural similarity” variables based on the *isotropic place encoding* or *pure vertical direction encoding* hypotheses (Figure 3.4C). The isotropic place encoding model predicts high neural similarity for the “same” condition and low similarity for “diff-V” and “diff-H” condition, therefore 1 was assigned for the pairwise similarity in the “same” condition and 0 was assigned for the “diff-V” and “diff-H” conditions. As long as the rank order was preserved, any number could be assigned. The pure vertical direction encoding model was assigned 1 for “same” and “diff-H”, and 0 for “diff-V”. I then calculated the partial Spearman correlation between the “observed neural similarity” and “predicted neural similarity”, while controlling the “visual similarity” for each subject. If this partial correlation is significantly above zero across subjects, it would be evidence of place or direction encoding in the neural data that is not fully accounted for by low level visual features. This partial correlation approach is similar to that employed by Carlin et al. (2011). I used a one-sided t-test to test significance.

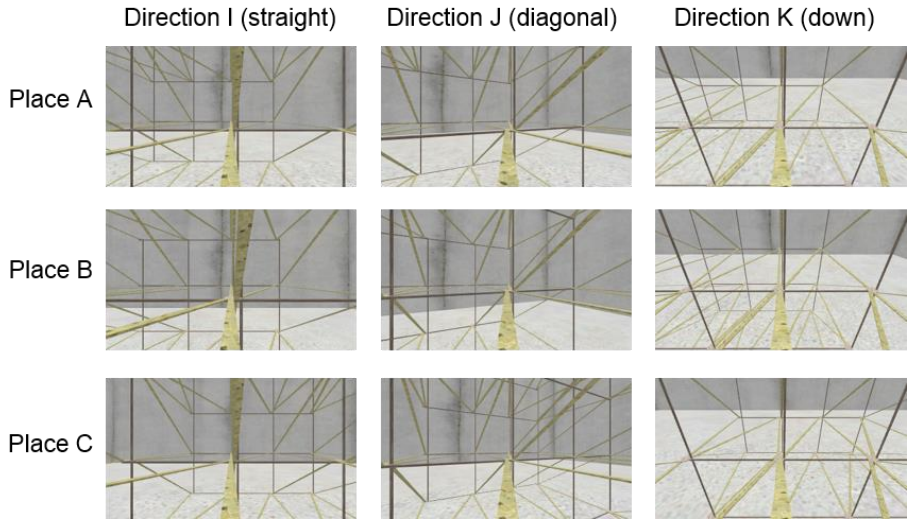


Figure 3.5 Example views from multiple locations and directions. Although the virtual environment was designed to contain a limited palette of colours and textures to prevent simple associations between spatial information and visual cues, the views were not completely orthogonal to the spatial information. The view particularly differed when participants moved up or down due to the proportion of the floor in the field of view.

3.2.9 Control analysis: head motion

One might ask whether participants physically moved their head during this virtual navigation experiment inside the MRI scanner, and this physical movement-related artefact could confound the analysis that seeks evidence of place and directional encoding in the (virtual) 3D space. To avoid this problem, I carefully instructed participants to remain still and used the standard head immobilisation equipment (foam pad) to minimise head movement. As Table 3.1 below shows, this was effective because head motion within each scanning session, as estimated by the SPM image realignment protocol, was considerably smaller than voxel size. On average, the maximal movement (Z translation) was only 1.38 mm compared to the 3 mm voxel size.

Table 3.1 Head motion parameters estimated by the SPM realignment process.

	X (mm)	Y (mm)	Z (mm)	Pitch (°)	Roll (°)	Yaw (°)
Mean (SD) across 36 subjects	0.32 (0.19)	1.08 (0.42)	1.38 (1.07)	0.03 (0.02)	0.01 (0.005)	0.01 (0.005)
The subject with the most movement	0.76	2.33	6.07	0.09	0.02	0.03
The subject with the least movement	0.11	0.30	0.30	0.01	0.003	0.003

More importantly, I confirmed the absence of confounding head motion related to the virtual rollercoaster ride in a pilot scanning study before the main experiment was conducted. In this pilot study involving the VR task, a small optical marker was attached to upper teeth of one volunteer via a personalised dental holder. While in the scanner, a high resolution motion tracking camera tracked the position and orientation of the optical marker with resolution of 80 Hz and 0.001 mm. After bandpass filtering (0.0078 to 0.2 Hz) which removes various sources of noise (e.g. fMRI slice acquisition frequency 14.3 Hz, heartbeat ~1.3 Hz and breathing ~0.3Hz), I examined the time course of the optical marker's information (3 linear movements and 3 rotation

parameters, a proxy for head position) time-locked to the onset of the journey (where the regressor for the GLM was constructed in the main experiment).

As can be seen from Figure 3.6 below, none of 6 head motion parameters showed abrupt changes at the onset of journey (pink line), countdown (cyan) or turn (green). The time course was baseline corrected by -4 to -3 sec to the journey onset. I also checked whether there were overall amplitude changes of head motion for each location or heading direction in the virtual environment. A squared sum of head motion parameters during the journey and countdown period was calculated for each trial and included in a one-way ANOVA with the 8 locations and 6 directions as the main factors. These tests further confirmed the absence of systematic differences in head motion for virtual location or direction ($p>0.2$ for all six parameters).

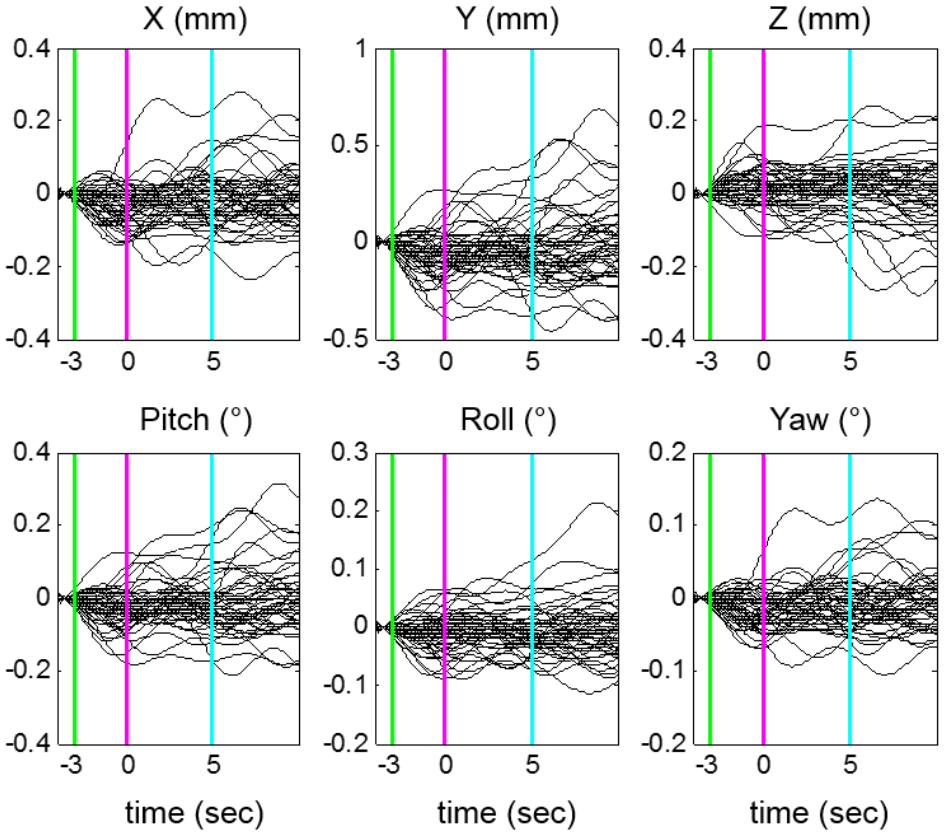


Figure 3.6 Six head motion parameters measured by a high resolution motion tracking system in a pilot subject. The black line is the time course of each trial; the pink line (0 sec) refers to the onset of the journey and the cyan line (5 sec) indicates the onset of the countdown period. The green line (-3 sec) indicates the onset of the turn prior to the journey. No sudden head motion was detected in the turn, journey or countdown periods.

3.3 Results

3.3.1 Behavioural performance during scanning

Overall, participants correctly kept track of their position and direction within the 3D lattice during the virtual rollercoaster ride (place question accuracy = 86.6 ± 12.3%, direction question accuracy = 93.4 ± 8.6%, chance level = 33.3%). The place questions were divided into two categories depending on the existence of a vertical distractor (Figure 3.2B). Accuracy did not differ between the categories ($t(35)=0.0$, $p=1.0$, Figure 3.2B), however response time differed significantly. Participants responded faster when there was no distractor along the vertical axis (without-V, mean = 2.91 ± 0.71 sec) compared to when a vertical distractor was present (with-V, mean = 3.21 ± 0.84 sec; $t(35)=3.7$, $p<0.001$). These results imply that the participants precisely identified themselves within a horizontal plane, and the process of distinguishing the vertical coordinate (‘Am I on the first floor or second floor?’) slightly slowed down the response without affecting accuracy.

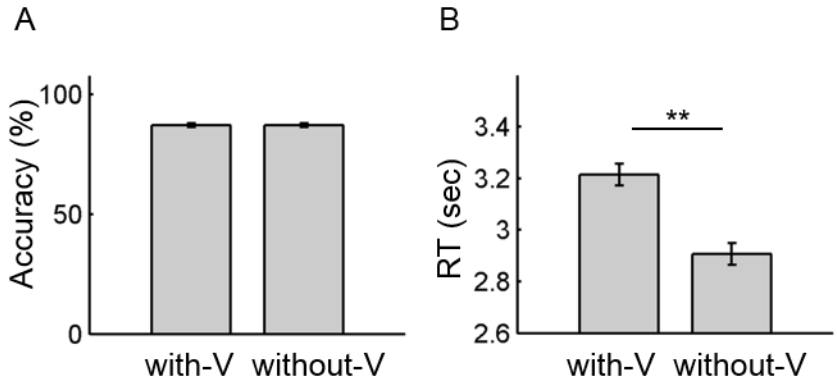


Figure 3.7 Place question results. (A) Accuracy did not differ between the with-V and without-V conditions. (B) Response time (RT) was significantly shorter in the without-V condition indicating that distinguishing the vertical position took longer than locating one’s position within a horizontal plane. Error bars are standard error of the mean adjusted for a within-subjects design (Morey 2008) **p<0.01.

For the direction question, the presence of a slope speeded up responses. Response time was significantly shorter when facing direction was tilted up or tilted down (mean =

2.26 ± 0.57 sec) compared to when facing direction was on a horizontal plane (mean = 2.69 ± 0.63 sec; $t(35)=-6.3$, $p<0.001$). Accuracy was higher in the vertical question trials (mean = 97.7 ± 4.9%) compared to the non-vertical trials (mean = 91.1 ± 12.6%; $t(35)=3.3$, $p=0.003$). This finding is consistent with previous results in humans and rats where the slope of a maze facilitated spatial memory (Grobéty and Schenk 1992a; Steck et al. 2003).

3.3.2 Post-scan debriefing session

When asked about their engagement with the task, 19% of participants chose the option “I felt like I was really there” and 69% chose “I occasionally thought about the environment as being on a computer screen, but overall the environment was convincing and I felt I was moving straight, up or down”, implying that the virtual environment used in this experiment provided an effective, if not complete, sense of being in 3D space. As mentioned previously, I also confirmed that height-related anxiety was not the confounding factor, as 94% of participants reported being “not at all anxious” during the scanning experiment.

Testing of the perception of the environment’s size revealed that the majority of subjects regarded the overall size of the virtual environment as medium or large (large: 31%, medium: 58%) and only 11% of the subjects reported it as small. Quantitatively, the ratio of vertical and horizontal distance estimates was not significantly different from the true ratio of 1 (mean ratio = 1.01 ± 0.27; $t(35)=0.2$, $p=0.8$). This result is suggestive of unbiased, isotropic 3D space perception in the virtual environment. However, the participants estimated the angle between the slope and the horizontal plane as significantly larger than the angle between two pavements on the horizontal plane (vertical angle = 40.3° ± 8.9, horizontal angle = 36.7° ± 9.0, mean ratio = 1.15 ± 0.31; $t(35)=2.8$, $p=0.008$) even though basic geometry would imply that the two angles

should be identical given that the vertical and horizontal distances are equal. This finding accords with the literature on human observers' tendency to overestimate the steepness of a slope (Proffitt et al. 1995) and vertical pointing in a 3D building (Brandt et al. 2015). The symmetric distance perception and asymmetric angle perception raised the question of whether the brain would encode vertical/horizontal place and direction symmetrically or asymmetrically, which I examined next.

3.3.3 Multivoxel pattern analysis – place encoding

Among the ROIs, the right anterior hippocampus showed evidence of significant place information ($F(2,70)=7.6$, $p<0.001$; Figure 3.8). In the hippocampus, both vertically displaced locations (diff-V) and horizontally displaced locations (diff-H) were significantly distinguishable from the same locations (post-hoc pairwise comparison, same > diff-V, $t(35)=3.1$, $p=0.006$; same > diff-H, $t(35)=3.6$, $p=0.002$, Bonferroni corrected). fMRI pattern similarity of vertically displaced locations (diff-V) and horizontally displaced locations (diff-H) were not significantly different from each other ($t(35)=0.7$, $p=1.0$). Although the absence of significant difference between the vertical and horizontal place encoding is not direct evidence of equivalence between the two, my prior encoding hypotheses suggest that this finding best fits with the isotropic 3D place encoding hypothesis where the horizontal and vertical dimensions are encoded with similar sensitivity (Figure 3.4C, *isotropic 3D*). The right entorhinal ROI showed a trend for pure vertical encoding ($F(2,70)=2.3$, $p=0.1$, same > diff-V, $t(35)=2.1$, $p=0.07$). No other ROIs showed either vertical or horizontal place information.

Of note, the absolute fMRI pattern similarity value of 0.02~0.03 may appear to be low, but given that neural signals are highly variable, and the similarity was calculated across different scanning sessions to ensure complete independence of the datasets (while also controlling the direction factor for place, and the place factor for direction),

small values are to be expected. Indeed, these values are perfectly in line with extant studies using this approach (Hsieh et al. 2014; Chadwick et al. 2015; Hsieh and Ranganath 2015; Schapiro et al. 2016; Schuck et al. 2016). More importantly, the absolute similarity value within a single condition has little meaning and the existence of place information should be tested by the difference in pattern similarity value between the conditions.

A control analysis confirmed that this isotropic place encoding in the anterior hippocampus pertained even after controlling for low level visual features (t(35)=3.5, p=0.0006).

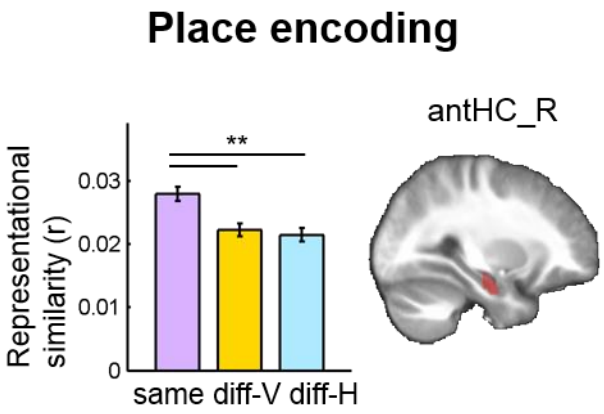


Figure 3.8 Place encoding results. The right anterior hippocampus (antHC_R) contained significant place information. Locations along the vertical axis (diff-V) and locations along the horizontal axis (diff-H) were equally distinguishable (same > diff-V or diff-H), suggesting an isotropic 3D representation. The anatomical ROI is overlaid on the group average structural MRI scan. Error bars are standard error of mean adjusted for a within-subjects design (Morey 2008); **p<0.01, post-hoc Bonferroni-corrected.

3.3.4 Multivoxel pattern analysis - direction encoding

The direction encoding analysis revealed different results from the place encoding results. Significant direction information was only found in the right posterior hippocampus (F(2,70)=4.8, p=0.01) and the right retrosplenial cortex (F(2,70)=3.8, p=0.04). In both ROIs, different vertical directions (diff-V) were distinguishable whereas

the different horizontal directions (diff-H) were not (post-hoc pairwise comparison, same > diff-V, t(35)=2.5, p=0.03; same ≈ diff-H, t(35)=-1.1, p=1.0 for the right posterior hippocampus; same > diff-V, t(35)=2.6, p=0.02; same ≈ diff-H, t(35)=-0.6, p=1.0 for the right retrosplenial cortex, Bonferroni corrected) (Figure 3.9). This result suggests a pure vertical encoding scheme in the right posterior hippocampus and right retrosplenial cortex, and this resonates with previous animal studies which found head direction cells that were only sensitive to the vertical pitch (Stackman and Taube, 1998; Finkelstein et al., 2014).

A control analysis confirmed that this result pertained even after controlling for visual texture similarity (t(35)=2.17, p=0.02 for the right posterior hippocampus; t(35)=1.9, p=0.03 for the right retrosplenial cortex).

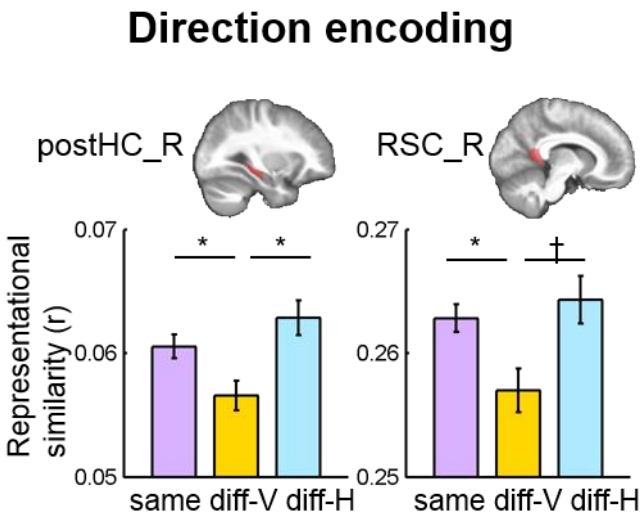


Figure 3.9 Direction encoding results. The right posterior hippocampus (postHC_R) and right retrosplenial cortex (RSC_R) ROIs contained significant direction information. In both regions, different vertical directions were distinct (same > diff-V) but the horizontal directions were not (same ≈ diff-H). The anatomical ROI is overlaid on the group average structural MRI scan. Error bars are standard error of mean adjusted for a within-subjects design (Morey 2008); *p<0.05, †p=0.12, post-hoc Bonferroni-corrected.

3.4 Discussion

In this first experiment, I investigated the neural representation of 3D spatial information in the human brain when participants explored a semi-volumetric space, a lattice structure. There were three main results. First, behaviourally, participants had similarly accurate memory for vertical and horizontal locations, whereas vertical tilt facilitated performance on the direction judgement task. Second, I found that the right anterior hippocampus contained place information that was sensitive to both horizontal and vertical axes. Finally, vertical directional information was found in the right posterior hippocampus and retrosplenial cortex.

The hippocampus has been long known to encode an animal's 2D location (O'Keefe and Dostrovsky 1971; Ekstrom et al. 2003; Hassabis et al. 2009; Sulpizio et al. 2014), and the current experiment extends previous knowledge by taking into account the third spatial dimension. Together with the behavioural findings of high accuracy irrespective of the presence of a vertical or horizontal distractor (except for the small response time difference), and the symmetrically perceived length of the 3D environment, multivoxel representational similarity in the right anterior hippocampus supports an isotropic 3D space encoding hypothesis, as has been observed in bats (Yartsev and Ulanovsky 2013), rather than the anisotropic planar encoding hypothesis based on rodent findings (Hayman et al. 2011; Jeffery et al. 2013).

However, it would be premature to conclude that the human hippocampus is more similar to that of the bat than the rat in relation to 3D spatial encoding. The shape of the environment and behavioural demands should be considered before coming to a firm conclusion. Place cells show a repeating firing pattern when an environment comprises multiple recurring compartments (Nitz 2011; Spiers et al. 2015) and some 3D environments have repeated structures along the vertical axis, e.g. the helical staircase

used in a previous rat study (Hayman et al. 2011). Therefore, an apparent lack of vertical information in rats could be due to the repeating nature of the 3D environment. Furthermore, the physical boundaries that discretise a 3D space along the vertical dimension (e.g. ceiling) might have contributed to the anisotropy of 3D space in Hayman et al. (2011). In contrast, my 3D lattice environment did not contain any walls and each location of interest was separated by the same vertical and horizontal distance, physically and perceptually. It would be interesting to test whether the human hippocampus shows a symmetric representation of 3D space even when the space is divided by clear physical boundaries like walls and ceilings, and this will be examined in Experiment 2.

Regarding behavioural demands, it should be noted that participants in my experiment were explicitly asked to encode both vertical and horizontal coordinates, whereas most animal studies do not impose such a requirement. A place cell's response can be modulated by reward and attention (Markus et al. 1995; Hölscher et al. 2003) and it is possible that a place cell adapts to encode and remember the space better when it is behaviourally relevant. It could be that place cells in rats would show an isotropic firing pattern if they were explicitly required to distinguish every location in 3D space. A recent behavioural study showed that rats were able to learn a 3D maze as well as a 2D maze, at least over a short timescale (Wilson et al. 2015). Conversely, the vertical and horizontal axes might be differentially encoded in humans where explicit spatial awareness is absent or a more demanding goal-oriented navigation task is used.

The other findings from this first experiment concerned the encoding of vertical and horizontal directions. At the behavioural level, participants indicated their heading direction faster and more accurately when they were facing up or down, and the vertical angle was overestimated to a greater extent than the horizontal angle. This result fits with the idea that the gravity (vertical) axis is a reference direction (Barnett-

Cowan and Bülthoff 2013). Knowing one's direction relative to the gravity axis is essential for maintaining the stability of body posture, and all animals have a tendency to maintain an upright head posture. Physical gravity did not play a part in the current experiment because subjects were in a supine position in the MRI scanner. However, the vertical axis can be defined not only by gravity but also by visual cues (Dyde et al. 2006). I believe the visually conveyed vertical axis in the current virtual environment was a reasonable proxy for the gravity vertical axis in real life, and that the experience of 'falling' during the pre-scan free exploration also supported this analogy.

At the neural level, the right posterior hippocampus and retrosplenial cortex exhibited only vertical direction information. This might reflect potential head direction cells that are only sensitive to vertical pitch similar to those found in animals (Stackman and Taube 1998; Finkelstein et al. 2015). The direction encoding result remained significant after controlling for visual texture similarity, but I acknowledge that view and head direction, in particular for the vertical component, were not perfectly orthogonal. It is possible that the posterior hippocampus and retrosplenial cortex findings may also be related to view encoding. Retrosplenial cortex has connections to many cortical and subcortical regions that map space in different reference frames including the hippocampus, posterior parietal cortex and thalamic nuclei (Vann et al. 2009). It encodes not only head direction but also turning behaviour (Alexander and Nitz 2015), place and view (Vass and Epstein 2013; Marchette et al. 2014), and stable landmarks (Auger et al. 2012, 2015). Future work should seek to disentangle these factors and isolate 3D head direction information.

It should also be noted that the current experimental design might not have had sufficient sensitivity to detect horizontal direction encoding in the brain. I set the angular difference between the directions as 29° because that is known to be the steepest slope humans can walk up. However, most previous fMRI studies detected direction-

specific responses using much larger angular differences (90°), e.g. north, south, east and west) (Baumann and Mattingley 2010; Vass and Epstein 2013; Marchette et al. 2014; Chadwick et al. 2015; Shine et al. 2016). A previous fMRI study treated directions that differed by 30° as identical directions (Bellmund et al. 2016) based on the animal literature that the tuning width of head direction cells is much larger than 30° (Cullen and Taube 2017). Therefore, it might be possible to detect both vertical and horizontal direction information if larger angular differences are used. Head direction encoding will be more extensively studied in Experiment 3 using a volumetric space.

The finding of different types of spatial information (place versus vertical direction) in the anterior and posterior hippocampus accords with other evidence of functional variation down its longitudinal axis. Based on the evidence from animal electrophysiology (Kjelstrup et al. 2008), lesion studies (McTighe et al. 2009) and neuroimaging (Evensmoen et al. 2015), it was proposed that the anterior hippocampus may encode a large-scale or generalisable representations of the environment, whereas posterior hippocampus may encode a fine-scale and local representation (Poppenk et al. 2013; Zeidman and Maguire 2016). In this experiment, the lattice structure eliminated the demand for fine-scale encoding of locations, and most subjects perceived the size of the environment as medium or large rather than small. Thus, the anterior hippocampus may have been suitable for representing this location information independent of the direction. In contrast, the posterior hippocampus could be associated with vertical direction because the detail of a view was more distinguishable when participants were heading up or down. Posterior hippocampus is connected to the parahippocampal and retrosplenial cortices that are known for scene processing (Kobayashi and Amaral 2003; Blessing et al. 2016) and are activated during scene discrimination tasks (Lee et al. 2008).

In conclusion, this first experiment provided evidence suggesting that vertical and horizontal location information can be equally well encoded in the hippocampus in a semi-volumetric 3D lattice environment, although the vertical gravity direction seems more salient than the horizontal direction. In the next experiment, I tested how a compartmentalised 3D space is encoded in the brain.

Chapter 4 Experiment 2: A multi-compartment 3D space

4.1 Introduction

In the previous chapter, I showed that the human anterior hippocampus expressed 3D location information that was similarly sensitive to the vertical and horizontal dimension when participants explored a semi-volumetric space. The 3D lattice structure used in the previous experiment did not contain a ceiling or walls which might potentially influence the representation of 3D space. In the next experiment, I investigated how a multi-level building, the most common type of working and living spaces for humans today, is represented in the brain. The spatial navigation and encoding strategies for multi-level buildings could be different to those for semi-volumetric space.

As noted in Chapter 1, regionalisation is a key characteristic of modern buildings - multiple floors stacked on top of each other and multiple rooms located side-by-side on a floor. When we navigate within multi-floor buildings, we can use hierarchical planning rather than using a 3D vector shortcut or volumetric 3D map. For example, we can decide which floor to go to (“second floor”), and which room on that floor (“the first room nearest the stairs”), then the location within the room (“the inside left corner of the room”). Regionalisation and the hierarchical representation of space involving multiple scales has been consistently observed (Hirtle and Jonides 1985; Han and Becker 2014; Balaguer et al. 2016), but it is not fully understood how spatial information about multiple scales is encoded at the neural level, particularly in a 3D context.

One obvious question is whether a common neural representation is used for each compartment (room). Using a generalised code to register local information is an efficient strategy compared to assigning unique codes for every location in an entire environment in the context of repeating substructures. Moreover, a common local

representation can be seen as a “spatial schema” that captures the essence of an environment and helps future learning of relevant environments or events (Tse et al. 2007; Marchette et al. 2017). The retrosplenial cortex and hippocampus are candidate brain regions for the encoding of within-compartment, local information. Place cells in the hippocampus are known to repeat their firing fields in a multi-compartment environment (Derdikman et al. 2009; Spiers et al. 2015). Moreover, human fMRI studies have shown that the hippocampus contains order information that generalises across different temporal sequences (Hsieh et al. 2014), and the retrosplenial cortex contains location codes that generalise across different virtual buildings (Marchette et al. 2014).

Another question is where in the brain each compartment is represented within the larger environment in order to complement the local room information. To the best of my knowledge, no study has simultaneously interrogated the neural representation of local spatial representations and the compartment information itself. The hippocampus might contain both types of information. It has been suggested that the hippocampus represents spatial information of multiple scales down its long axis. For example, the size of place fields is larger in ventral hippocampus than dorsal hippocampus in rats (Kjelstrup et al. 2008). In a human fMRI study, increased activation in posterior (dorsal) hippocampus was associated with a fine-grained spatial map, whereas the anterior (ventral) hippocampus was linked with coarse-grained encoding (Evensmoen et al. 2015).

It is also important to ask whether vertical and horizontal information is equally well encoded in a compartmentalised 3D environment. In other words, is it the case that when a room is located directly above another room, are they as equally distinguishable as two rooms that are side-by-side on the same floor? Although I showed that the human hippocampus encoded vertical and horizontal location

information equally well in a 3D lattice structure (Experiment 1), it might be different in the presence of physical barriers like walls and ceilings. As reviewed in Section 1.6.2, there are mixed behavioural findings in relation to vertical-horizontal symmetry/asymmetry in a multi-level building. For example, a group who learned the location of objects in a virtual multi-floor building along a floor route had, overall, better spatial memory than a group who learned along a vertical columnar route, suggesting a bias towards the floor-base representation (Thibault et al. 2013). However, another study reported that twice as many participants reported a columnar representation of a building than a floor representation (Büchner et al. 2007).

I sought to address the issues outlined above in order to provide much-needed information about how regionalisation of space is realised at the neural level, in particular in a 3D context. Participants learned the locations of paintings in a virtual multi-floor gallery building by volitionally navigating within it prior to scanning. Also before scanning, I compared their spatial judgments within and across vertical and horizontal boundaries. Then participants performed an object-location memory test while being passively moved in the virtual building during fMRI scanning. Repetition suppression analysis was used to ascertain which brain regions represented the local information within a room, or the room information within the building. In addition, I asked whether vertical and horizontal room information in the brain was symmetrically or asymmetrically represented.

4.2 Methods

4.2.1 Participants

Thirty healthy adults took part in the experiment (15 females; age 23.7 ± 4.6 years; range 18-35 years; all right-handed). All had normal or corrected-to-normal vision and

gave informed written consent to participation in accordance with the local research ethics committee.

4.2.2 The virtual environment

The virtual environment was a gallery building. There were 4 identical-looking rooms within the building, two rooms on each of two main floors (Figure 4.1). There was a unique painting located in each of the four corners of a room, resulting in 16 unique locations in the building. The paintings were simple and depicted animals or plants such as a dog, rose or koala bear. Painting locations were randomised across the participants, therefore spatial location was orthogonal to the content of the painting associated with it. The virtual environment was implemented using Unity 4.6 (Unity Technologies, CA, USA). A first-person-perspective was used and the field-of-view was ±30° for vertical axes and ±37.6° for horizontal axes. During pre-scan training, the stimuli were rendered on a standard PC (Dell Optiplex 980, integrated graphic chipset) and presented on a 20.1 inch LCD monitor (Dell 2007FP). The stimuli filled 70% of the screen width. The same PC was used during scanning, and the stimuli were projected (using an Epson EH-TW5900 projector; resolution 1024 x 768) on a screen at the back of the MRI scanner bore and participants saw the screen through a mirror attached to the head coil.

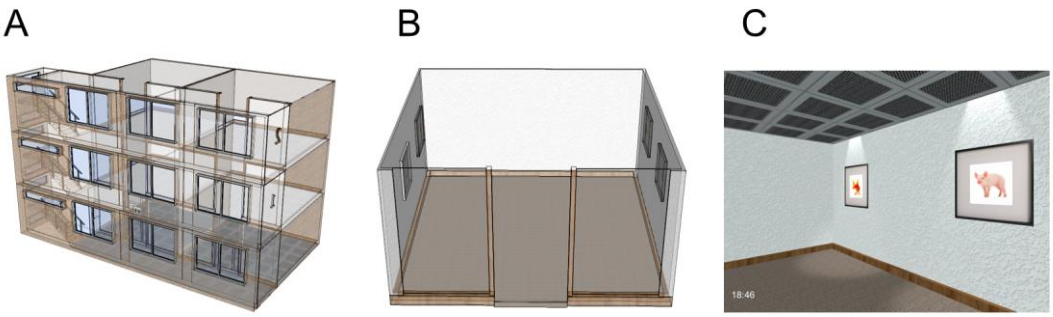


Figure 4.1 The virtual gallery. (A) Overview of the virtual gallery building with transparent walls for display purposes here. (B) Overview of one room with transparent walls for display purposes here. (C) An example view from a subject's perspective during the learning period prior to scanning.

4.2.3 Procedure

Having watched a demonstration of the experimental tasks, each participant completed the tasks in following order: movement practice and learning prior to scanning, a pre-scan egocentric judgment task and the object location memory task during scanning (which was preceded by a short practice of the scanner task).

Movement practise

Participants practised moving inside the virtual building using a keyboard (the W/A/S/D and arrow keys). They could move forwards, backwards, left and right. They could turn to the right or left side, and look up or down. They were asked to visit all floors via a staircase. During this movement practise, the paintings were hidden. All participants quickly learned how to control their movements in the virtual environment.

Learning prior to scanning

Participants were instructed to freely explore the virtual gallery building and to memorise the locations of all the paintings. Note taking was not allowed. I allocated 20 minutes for the initial learning phase (a timer was shown in a corner of the screen), but allowed participants to proceed to the test phase before 20 minutes had elapsed if they felt that they had learned the layout very well. The purpose of this self-determined criterion was to prevent participants from becoming bored. Seventeen out of the 30 participants moved on to the test phase before 20 minutes has passed (mean 16 min, SD 2 min). This subjective criterion is used because experimenters could check the participants' objective memory performance afterwards and let the participants revisit the building and learn the layout again if performance was sub-optimal, before they proceeded to the scanner. Four out of the 30 participants (only one of whom was among the 17 participants who asked to move on to testing prior to the 20 minutes

elapsing) had to re-visit the virtual building for up to 5 additional minutes because their accuracy either in the egocentric judgment test or the short practice for the scanning object-location test was below 70%. These four participants' accuracy during scanning was between 78% and 83%. The mean accuracy of the 30 participants for the scanning task was 93% (SD 5.5%). I am confident, therefore, that every participant had good knowledge of the spatial layout.

Pre-scan: egocentric judgment task

Immediately after the learning phase, there was a spatial memory test which required participants to make egocentric spatial judgments in the gallery. This test was used to examine the influence of vertical and horizontal boundaries on the mental representation of 3D space (see the Section 4.2.4 below).

On each trial of this test, participants saw a short dynamic video which provided the sensation of being transported to one of the 16 paintings (locations) from the corridor (duration 2.5 sec; Figure 4.2A). These videos were used to promote the impression of navigation whilst providing full experimental control. On half of the trials, participants started from one end of the corridor facing a floor sign on the wall, while on the remaining trials, they started from the other end of the corridor facing the stairs. In both cases they would terminate in the same location within a room, regardless of whether they began the journey facing the floor sign or stairs.

On every trial, participants were transported to one of two rooms on the floor where they started - thus the videos did not contain vertical movement via the staircase. Of note, except for the target painting, the other three paintings in a room were concealed behind curtains. Once a participant arrived at the target painting within a room, a question appeared on the screen (Figure 4.2A). The question asked about the position of another painting relative to the participant's current position, e.g. "Is the pig on your

left?", "Is the sunflower on your right?", "Is the dog above you?", "Is the duck below you?". Participants responded yes or no by pressing a keypad with their index or middle finger. Similar to a previous study by Marchette et al. (2014), I instructed participants to interpret left, right, above or below broadly, "including anything that would be on that side of the body" and not just the painting directly left, right, above or below. For example when a participant was facing the turtle painting in Figure 4.2B, the sunflower was on their right and the duck was above. The time limit for answering the question was up to 5 sec. The inter-trial interval (ITI) was drawn from a truncated gamma distribution (mean 2.9 sec, minimum 2.0 sec, maximum 6.0 sec, shape parameter 4, scale parameter 0.5) and there were 64 trials. Participants were provided with their total number of correct and wrong answers at the end of the test, but did not receive feedback on individual trials.

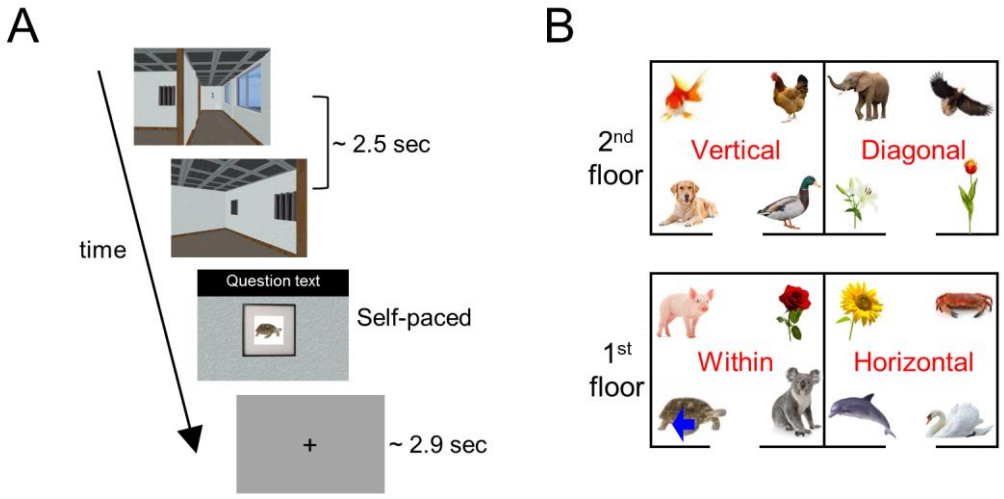


Figure 4.2 Experimental tasks. (A) On each trial, participants were virtually transported to one of the paintings from a corridor, and then participants performed spatial memory tasks. During the pre-scan egocentric judgments test, they were asked to make spatial judgments about the locations of other paintings, e.g. "Is the pig on your right?" (mean response time = 3.2 sec). During the scanning test, they were asked to indicate whether the painting was the correct one or not for that location, "Is this picture correct?" (mean response time = 1.3 sec). (B) An example layout of 16 paintings located in the 4 rooms of the gallery. The within, vertical, horizontal and diagonal rooms were defined relative to a participant's current location. In this example, the participant was standing in front of the turtle painting (blue arrow).

fMRI scan: object location memory task

On each trial of the scanning task, participants were transported to one of the paintings from the corridor as in the pre-scan memory test (duration 2.5 sec; Figure 4.2A). All

four paintings in the room were concealed behind curtains. Once a participant arrived at a painting, the curtain was lifted. The participant then indicated whether the painting was the correct one or not for that location by using a keypad. On 80% of the trials, the correct painting was presented and on 20% of trials a painting was replaced by one of the other 15 paintings. The response to the question was self-paced with an upper limit of 4.5 sec (mean response time 1.3 sec, SD 0.7 sec), and the inter-trial intervals were the same as those in the pre-scan memory test. There were 100 trials for each scanning session and each participant completed 4 scanning sessions with a short break between them, making a total functional scanning time of ~50 minutes. Participants were told the total number of correct and wrong answers at the end of each scanning session, but individual trial feedback was not given. The order of visiting the paintings (locations) are important for repetition suppression analysis, and this will described in shortly. I used a sequence that balances first-order carry-over effects (Aguirre 2007; Nonyane and Theobald 2007). This meant that one location was followed by every other location with similar frequency. A similar sequence has been used in other fMRI studies (Vass and Epstein 2013; Sulpizio et al. 2014).

4.2.4 Behavioural analyses

Pre-scan egocentric judgement test

To test the influence of compartmentalisation by vertical and horizontal boundaries on spatial judgments, I compared the accuracy and response time of egocentric spatial judgments between four conditions (Figure 4.2B): (1) within; when the painting in question was in the same room as a participant; e.g. a participant was facing the turtle and made a spatial judgment about the pig, rose, or koala; (2) vertical; when the painting in question was in the room above or below a participant; e.g. a participant facing the turtle was asked about the dog, gold fish, duck or chicken; (3) horizontal; when the painting in question was in the adjacent room on the same floor as a

participant; e.g. a participant facing the turtle was asked about the sunflower, crab, dolphin or swan; (4) diagonal; when the painting in question was in a diagonal room; e.g. a participant facing the turtle was asked about the elephant, eagle, lily or tulip.

If participants had a holistic mental representation of 3D space irrespective of physical boundaries within the building, performance for all four conditions should be similar. In contrast, if their mental representation was segmented into each room, spatial judgments within the same room (within) would be facilitated and therefore higher accuracy and/or faster response times would be expected compared to spatial judgments across different rooms (vertical, horizontal or diagonal conditions). If space is predominantly divided into a horizontal plane, as suggested by some previous studies (Jovalekic et al. 2011; Thibault et al. 2013; Flores-Abreu et al. 2014), spatial judgments about paintings on different floors (vertical, diagonal) would be more difficult than paintings on the same floor (within, horizontal). I used a repeated one-way ANOVA and post-hoc paired t-tests to compare the accuracy and response time for the four conditions, with a significance threshold of $p<0.05$.

Object location memory test during scanning

I tested whether spatial knowledge of 3D location was organised into multiple compartments by measuring a behavioural priming effect. Each trial was labelled as one of four conditions depending on the room participants visited in the immediately preceding trial (Figure 4.3A,B). Figure 4.3B shows an example trial sequence and the room label for each trial in red: (1) same; when participants visited the same room in the previous trial, e.g. the 2nd trial; (2) vertical; when participants previously visited the room above or below the current room, e.g. the 3rd trial; (3) horizontal; when participants previously visited the adjacent room on the same floor, e.g. the 5th trial; (4) diagonal; when participants previously visited neither a vertically nor horizontally adjacent room, the e.g. 4th trial.

A holistic, volumetric representation of space would result in similar behavioural performance for all four conditions. If representations were compartmentalised, participants would make more accurate and/or faster judgments when spatial memory was primed by the representation of the same compartment (room). If spatial representations were further grouped along the horizontal plane, visiting the adjacent room on the same floor (horizontal condition) will also evoke a behavioural priming effect. Alternatively, the space might be represented in a vertical column, leading to the prediction of a priming effect for the vertical condition. I compared accuracy and response time for the four conditions using a repeated-measure ANOVA and post-hoc paired t-tests.

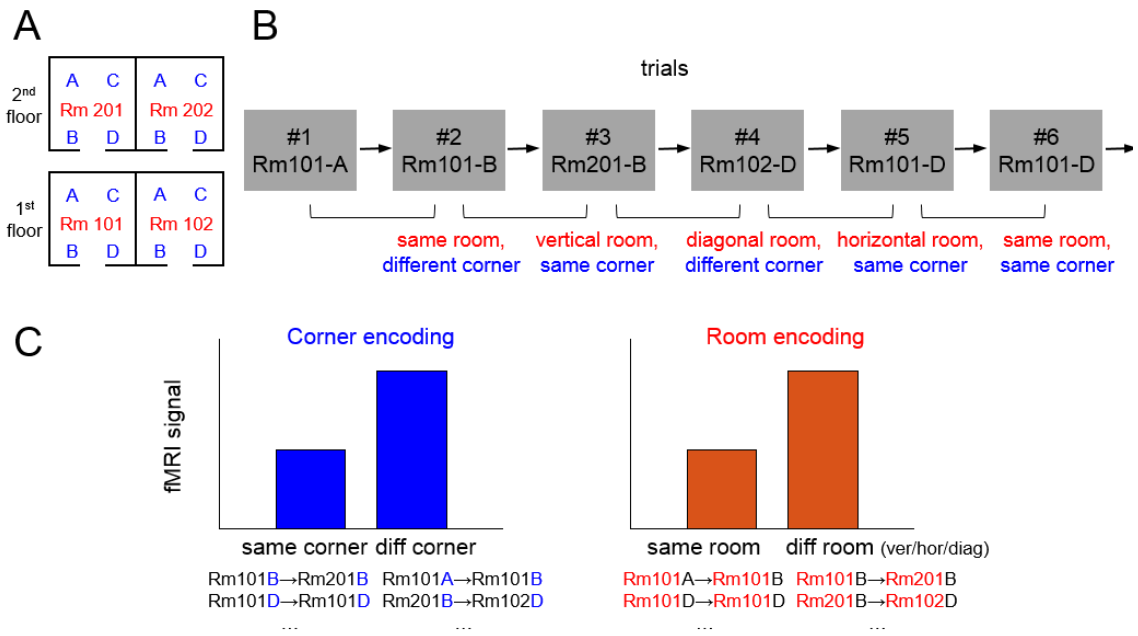


Figure 4.3 Analysis overview. (A) A floor plan of the virtual building. The 4 rooms are labelled as "Rm101", "Rm102", "Rm201", "Rm202" and the 4 corners as "A", "B", "C", "D" for the purposes of explanation here. Participants were not told of any explicit labels during the experiment. (B) An example trial sequence. For the behavioural and fMRI repetition suppression analyses, each trial was labelled based on its spatial relationship with the preceding trial, e.g., the 2nd trial belongs to the "same room, different corner" condition. Of note, this trial definition is used for analysis only and participants were not asked to pay attention to the preceding trial. (C) Predictions for the fMRI signals. If some brain regions encoded corner information, a lower fMRI signal was expected for the same corner condition compared to the different corner condition. If room information was encoded, fMRI signal was expected to be lower for the same room condition compared to the different room condition.

4.2.5 Scanning and image processing

Functional scans (voxel size = 3 x 3 x 3 mm, 44 slices, TR = 3.08 sec) and a structural scan (1 x 1 x 1 mm) were acquired using a 3T Siemens Trio scanner and preprocessed using SPM12. The scanning parameters and preprocessing are described in detail in Section 2.5.4. To summarise, after removing the first 5 volumes to allow for T1 equilibration effects, functional images were realigned to the first volume of each scanning session and geometric distortion was corrected by the SPM unwarp function using the fieldmaps. Each participant's anatomical image was then coregistered to the distortion corrected mean functional images. Functional images were normalised to MNI space, then spatial smoothing (8 mm) was applied.

4.2.6 fMRI analyses

Main analysis: room and corner encoding

I used an fMRI repetition suppression analysis to search for two types of spatial information in the brain: (1) *corner*; a participant's location within a room, and (2) *room*; which room a participant was in within the building. fMRI repetition suppression analysis is based on the assumption that when a similar neural population is activated across two consecutive trials, the fMRI signal is reduced during the second trial. Therefore, if a brain region encodes corner information, visiting the same corner in a consecutive trial would result in reduced fMRI signal compared to visiting a different corner (Figure 4.3C). For example, visiting Rm201-B after Rm101-B (3rd trial in the example sequence, Figure 4.3B) or visiting Rm101-D twice in a row (6th trial in the example) would evoke reduced fMRI signal than visiting Rm101-B after Rm101-A (2nd trial in the example) or visiting Rm102-D after Rm201-B (4th trial). On the other hand, if a brain region encodes room information, visiting the same room in consecutive trials (e.g. Rm101-A → Rm101-B, 2nd trial) would result in reduced fMRI signal compared to visiting a different room (e.g. Rm101-B →Rm201-B, 3rd trial).

I also tested whether vertical and horizontal boundaries similarly influenced neural similarity between rooms. If there was a bias in encoding horizontal information better than vertical (floor), then the two rooms on top of each other (e.g. Rm101 and Rm201) would be less distinguishable than the two rooms on the same floor (e.g. Rm101 and Rm102). Therefore, visiting a vertically adjacent room (e.g. Rm101-B → Rm201-B, 3rd trial) would result in more repetition suppression, leading to a reduced fMRI signal, than visiting a horizontally adjacent room (e.g. Rm102-D → Rm101-D, 5th trial). I was also able to ask whether two rooms in a diagonal relationship (e.g. Rm201-B → Rm102-D, 4th trial) were more distinguishable than vertically or horizontally adjacent rooms.

To answer these questions, I constructed a GLM which modelled each trial based on its spatial relationship to the preceding trial in terms of two factors: corner and room. The corner factor had 2 levels: same or different corner, and the room factor had 4 levels: same, vertical, horizontal or diagonal room. This resulted in a 2 x 4 = 8 main regressors. Each regressor was a boxcar function which for each trial modelled the entire stimulus duration including the virtual navigation period (2.5 sec) and subsequent object-location memory test (mean response time 1.3 sec, SD 0.7 sec) (Figure 4.2A top three panels) convolved with the SPM canonical hemodynamic response function. I modelled the entire period as a single boxcar function because information about spatial location was cumulatively processed throughout the navigation video and continued until participants decided whether the painting was the correct one or not for the location. The first trial of each scanning session, which did not have an immediately preceding trial, or the trials where participants were incorrect (mean 6.8%, SD 5.5%) were excluded from the main regressors and modelled separately. The GLM also included nuisance regressors: six head motion realignment parameters and the scanning session-specific constant regressor.

First, I conducted a whole-brain analysis to search for corner and room information using two contrasts: (1) “same corner < different corner”, collapsed across the room factor, and (2) “same room < different room” (the average of the vertical room/horizontal room/diagonal rooms), collapsed across the corner factor. Each participant’s contrast map was then fed into a group level random effects analysis. Given my a priori hypothesis about the role of hippocampus and retrosplenial cortex for encoding spatial information, I report voxel-wise p-values corrected for anatomically defined hippocampus and retrosplenial cortex ROIs. I used the same anatomical ROIs that I delineated in Experiment 1 because my fMRI images were spatially normalised to a standard template brain so that anatomical differences between groups of participants in each experiment were negligible (see Section 3.2.6 for details). For the rest of the brain, I report regions that survived a whole-brain corrected family-wise error (FWE) rate of 0.05.

Having identified brain regions that contained significant corner information from the whole brain analysis, I examined the spatial encoding in these regions further by extracting the mean fMRI activity. As a proxy for the mean fMRI activity, beta weights for every voxel within the spherical ROIs (radius 5 mm, centred at the peak voxel) were averaged for each participant, and then compared at the group level by paired t-tests. For this functional ROI-based analysis, I divided the “same corner” condition into “same corner, same room” and “same corner, different room” and compared each condition to “different corner”. This analysis was used to rule out the possibility that the corner encoding was driven purely by the repetition suppression effect of “same corner, same room” < “different corner”. If a brain region encodes each of the 16 locations (or associated paintings) without a spatial hierarchy, repetition suppression would only occur for the “same corner, same room” condition and there would be no difference between “same corner, different room” and “different corner”.

I conducted a similar control analysis in the brain regions that contained significant room information (“same room < different room”). I compared the mean activity of “same room, same corner” and “same room, different corner” to “different room” to rule out the possibility that the room encoding was driven by the repetition suppression of the exactly same location. Crucially, I also compared mean activity of different room conditions (vertical/horizontal/diagonal rooms) to test for any potential bias in encoding vertical or horizontal information.

Of note, this experiment was specifically designed to examine the main effect of corner and room information, rather than to test a pure non-hierarchical encoding model where only the exact same location shows repetition suppression (“same corner, same room” < ”different corner, same room” = “same corner, different room” = “different corner, different room”). Such an encoding hypothesis cannot be separated from the painting encoding hypothesis, given that each location was associated with a unique painting. The data could also be examined in terms of 3D physical metric distance from the preceding trial modelled as a linear parametric regressor. However, the highly discretized nature of the environment makes inferences about metric encoding difficult in this context, and this issue would be better addressed with a different type of environment.

4.2.7 Supplementary analysis: room versus view encoding

In this experiment, room information was cued by a distinctive view such as a wall containing a floor sign, therefore the room encoding effect could arise due to view encoding and/or more abstract spatial information about a room that was not limited to a particular view. I was able to test for these possibilities because participants were virtually transported to each room from two directions as I described in Section 4.2.3, which means they could visit the same room on consecutive trials from the same or a

different direction (Figure 4.4). For example, if they had visited Rm102 from the floor sign side in the preceding trial and visited the same room from the stairs side in the current trial, the initial views were different even though the same room was visited. On the other hand, if they had visited Rm102 from the floor sign side in the preceding trial and visited Rm202 from the floor sign side in the current trial, the initial views were similar even though two rooms were different.

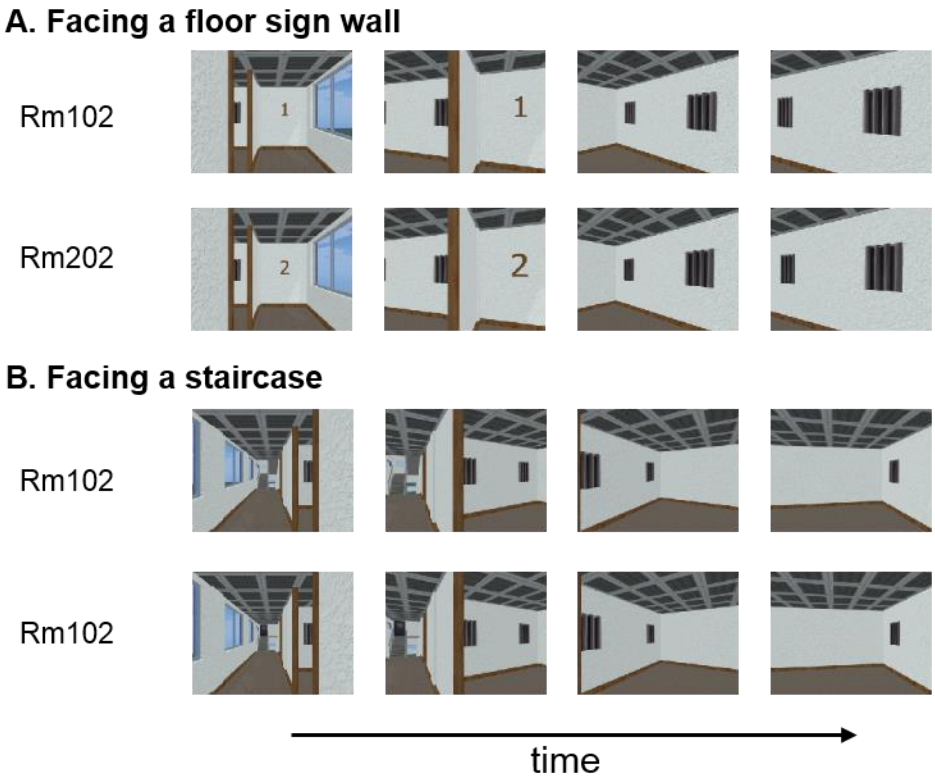


Figure 4.4 Two views associated with each room. Participants could be transported to a room from two opposite ends of the corridor, (A) facing a floor sign wall or (B) facing a staircase. This created two visually distinctive approaches for each room.

I constructed a GLM which modelled each trial based on two factors: whether it was the same or a different room from the previous trial, and whether the starting direction (view) was the same or different direction from the previous trial. This resulted in 4 trial types: “same room, same view”, “same room, different view”, “different room, similar view”, and “different room, different view”. As in the main analysis, only correct trials

were included for the main regressors, and head motion realignment parameters and scanning session-specific constant regressors were included in the GLM.

For each participant, I extracted the mean activity (beta weights) for each trial type in the room encoding regions identified in the “same room < different room” contrast described earlier. I conducted a repeated measures ANOVA and post-hoc paired t-tests to compare the mean beta weights between the “same room, same view”, “same room, different view”, and “different room” (collapsed over similar and different view). If only the view was encoded, then the “same room, same view” would have a reduced fMRI signal compared to “different room”, but “same room, different view” would not be associated with a reduced fMRI signal compared to the “different room” condition. If abstract room information was encoded, the “same room, different view” condition would also be associated with reduced fMRI signal compared to the “different room” due to repetition of the room. I was also able to compare “same room, same view” and “same room, different view” to test view dependency when the room was repeated.

4.3 Results

4.3.1 Behavioural results

Pre-scan egocentric judgment task

In order to examine the influence of vertical and horizontal boundaries on the mental representation of 3D space, I compared the accuracy and response time of spatial judgments for 4 conditions: within, vertical, horizontal and diagonal rooms. Participants were faster at judging the location of paintings within the same room compared to paintings in different rooms (Figure 4.5; $F(3,87)=5.4$, $p=0.002$, post-hoc paired t-tests: within vs. vertical, $t(29)=-3.5$, $p=0.001$; within vs. horizontal, $t(29)=-2.7$, $p=0.011$; within vs. diagonal, $t(29)=-2.2$, $p=0.034$). There was no significant difference in response time

between the vertical, horizontal and diagonal rooms. This result suggests the importance of a physical boundary, but this was not influenced by whether the boundary was vertical or horizontal. Accuracy did not differ significantly between the four conditions (F(3,87)=1.2, p=0.3; mean overall accuracy 80%, SD 12%).

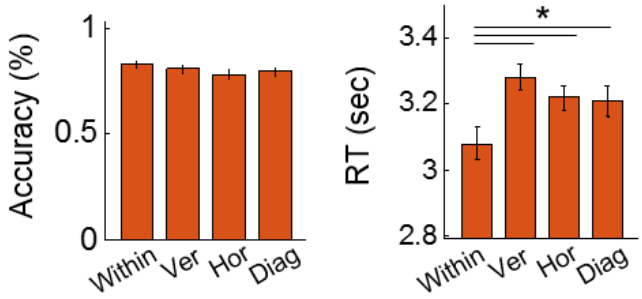


Figure 4.5 Accuracy and response time (RT) during the pre-scan egocentric direction judgments task. Accuracy was not different between the conditions. However, participants were significantly faster for the within-room (Within) condition. There were no differences between vertical (Ver), horizontal (Hor) or diagonal (Diag) rooms. Error bars are SEM adjusted for a within-subjects design (Morey 2008). *p<0.05.

Object-location memory task during scanning

Overall, participants performed well on the object-location memory task (mean accuracy 93%, SD 5.5%). Participants were more accurate and faster at judging whether a painting was in the correct location if they had visited the same room in the preceding trial (Figure 4.6; Accuracy: F(3,87)=4.2, p=0.008; post-hoc paired t-tests: same vs. vertical, t(29)=3.4, p=0.002; same vs. horizontal, t(29)=2.3, p=0.032; same vs. diagonal, t(29)=2.3, p=0.028; response time: F(3,87)=8.3, p<0.001; same vs. vertical, t(29)=-4.0, p<0.001; same vs. horizontal, t(29)=-2.8, p=0.009; same vs. diagonal, t(29)=-4.1, p<0.001), and neither accuracy nor response time differed between the vertical, horizontal or diagonal rooms. This result, along with the pre-scan memory task, suggests the mental representation of 3D space was segmented into each room, regardless of vertical floor.

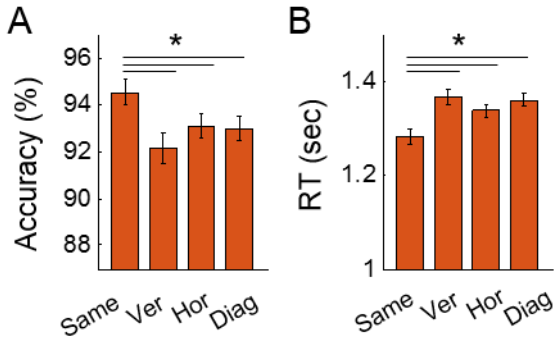


Figure 4.6 The behavioural priming effect of room during the scanning task. (A) Accuracy was significantly higher for the same room condition compared to all other rooms. There was no significant difference between the other room types. (B) Response time (RT) was significantly shorter for the same room condition compared to all other conditions. There was no significant difference between other room types. These results are in line with a room-based representation of 3D space, regardless of vertical floor. Error bars are SEM adjusted for a within-subjects design (Morey 2008). *p<0.05.

4.3.2 fMRI results

Corner information

The "same corner < different corner" contrast revealed left anterior lateral hippocampus (Figure 4.7, peak MNI coordinate [-33, -19, -16], t(29)=5.31, p=0.001, small volume corrected for a bilateral hippocampal mask), suggesting that this region encodes at which corner a participant is located within a room. No other brain region showed a significant corner repetition suppression effect at the whole brain corrected level.

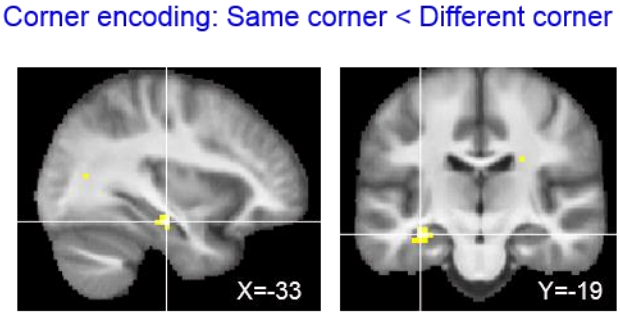


Figure 4.7 Corner encoding region. The whole brain contrast "same corner < different corner" revealed only the left anterior hippocampus (peak MNI = [-33,-19,-16], t(29)=5.31, p=0.001, small volume corrected). The thresholded map is overlaid on a group average structural MRI scan (p<0.001, uncorrected for display purposes). No other brain region survived correction for multiple comparisons.

I further examined this spatial encoding in the left anterior lateral hippocampus by extracting the mean activity (beta) for each condition. I investigated the fMRI signal when exactly the same location was visited ("same corner, same room", e.g. Figure 4.3B, Rm101-D → Rm101-D) and when the same corner, but a different room was visited ("same corner, different room", e.g. Rm101-B → Rm201-B) and compared them to the "different corner" condition (e.g. Rm201-B → Rm102-D). If the entire building is represented in a single volumetric space without a hierarchy, then each of the locations would be uniquely encoded, so repetition suppression is expected only for the "same corner, same room" condition. My finding speaks against the single volumetric representation hypothesis because both "same corner, same room" and "same corner, different room" conditions evoked significant repetition suppression effects compared to the "different corner" condition (one-sided paired t-tests: "same corner, same room" < "different corner", $t(29)=-4.4$, $p<0.001$; "same corner, different room" < "different corner", $t(29)=-4.2$, $p<0.001$). This implies that the anterior hippocampus contains local corner information that is generalised across different rooms, supporting an efficient hierarchical representation of 3D space.

On a related note, one might ask whether the hippocampus showed sensitivity to the heading direction instead of location. Participants faced opposite walls when they were at corner A (or B) and when they were at corner C (or D) (Figure 4.3A). However, they faced the same direction when they were at location A and B (or C and D) and further analysis revealed that there was no difference in the anterior hippocampus when participants visited a corner on the same wall or the opposite wall. Thus, I can conclude that the hippocampus encoded corner information rather than heading direction.

Room information

The "same room < different room" contrast revealed bilateral retrosplenial cortex (right retrosplenial peak [9, -52, 11], $t(29)=8.55$, $p<0.001$; left retrosplenial peak [-9, -58, 14],

t(29)=7.91, p<0.001, small volume corrected for a bilateral retrosplenial mask), right parahippocampal cortex (peak [27, -37, -16], t(29)=7.21, p<0.001), and the posterior part of the hippocampus (right hippocampus peak [27, -28, -10], t(29)=6.12, p<0.001; left hippocampus peak [-27, -34, -7], t(29)=4.32, p=0.014, small volume corrected for a bilateral hippocampus mask) (Figure 4.8A). This suggests that these regions encode in which room a participant was located in the building. It is notable that the room information was detectable in the posterior portion of hippocampus, compared to corner information which was detectable in the anterior hippocampus.

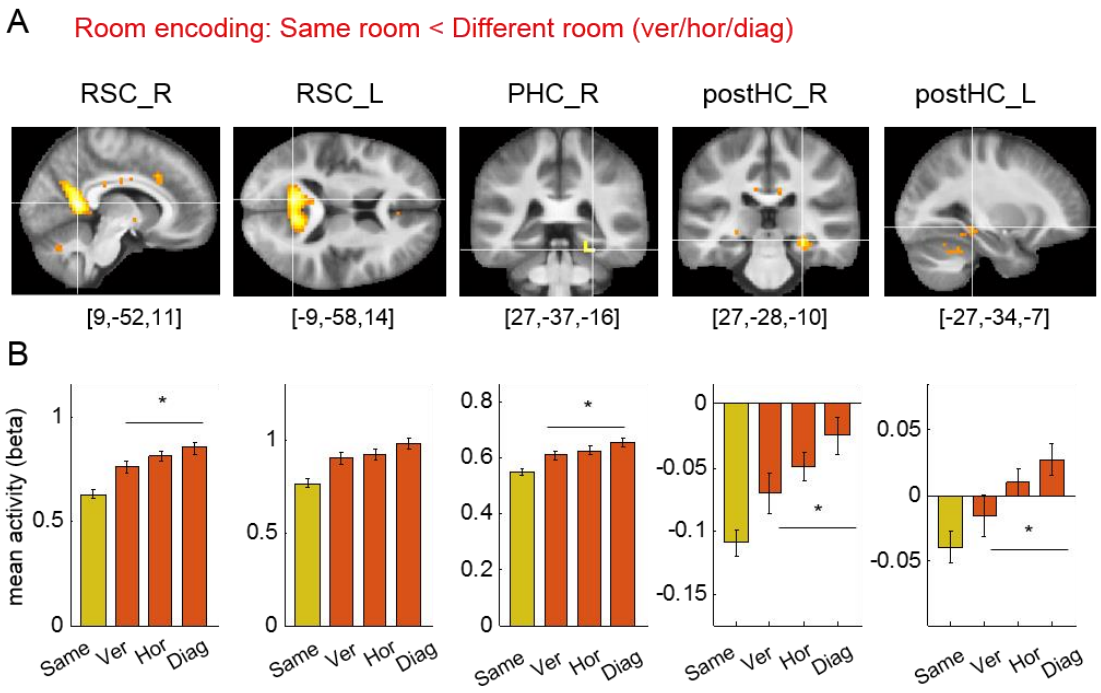


Figure 4.8 Room encoding regions. (A) The whole brain contrast "same room < different room" revealed bilateral retrosplenial cortex (RSC), right parahippocampal cortex (PHC) and bilateral posterior hippocampus (postHC). Given a priori interest in RSC and postHC, their clusters are shown with a small volume corrected threshold level (T(29)>3.67, T(29)>3.75), while the PHC cluster is shown with a whole-brain corrected threshold (T(29)>6.008). The peak MNI coordinate is shown below each cluster. (B) Comparison of mean activity for three different room types (vertical/horizontal/diagonal) at each cluster (5mm sphere at peak voxel). The "same" condition (in yellow) is shown for reference purposes. The response to the diagonal condition was significantly larger than for the vertical condition in all regions except the left RSC. There was no significant difference between the vertical and horizontal conditions. Error bars are SEM adjusted for a within-subjects design (Morey 2008). *p<0.05.

I further examined this spatial encoding in the right and left retrosplenial cortex (RSC_R, RSC_L), right parahippocampal cortex (PHC_R) and right and left posterior

hippocampus (postHC_R, postHC_L) by extracting the mean fMRI activity for the “same corner, same room”, “different corner, same room”, and “different room” conditions. In all regions, I found significant repetition suppression effects for both “same corner, same room” and “different corner, same room” conditions compared to the “different room” (one-sided paired t-tests: “same corner, same room” < “different room”: RSC_R, $t(29)=-6.6$, $p<0.001$; RSC_L, $t(29)=-4.9$, $p<0.001$; PHC_R, $t(29)=-4.9$, $p<0.001$; postHC_R, $t(29)=-4.9$, $p<0.001$; postHC_L, $t(29)=-3.2$, $p=0.002$; “different corner, same room” < “different room”: RSC_R, $t(29)=-4.1$, $p<0.001$; RSC_L, $t(29)=-3.6$, $p<0.001$; PHC_R, $t(29)=-5.0$, $p<0.001$; postHC_R, $t(29)=-3.0$, $p=0.003$; postHC_L, $t(29)=-2.2$, $p=0.02$). These findings suggest the presence of room information that is independent of the local corner.

I then tested for the existence of vertical-horizontal asymmetry in these five room encoding regions - RSC_R, RSC_L, PHC_R, postHC_R, postHC_L - by extracting the mean activity for sub-categories of the different room conditions: vertical room, horizontal room and diagonal room. If vertical information was relatively poorly encoded compared to horizontal information, I would expect that two rooms on top of each other to be more similarly represented in the brain than the two adjacent rooms on the same floor. Consequently, I would expect less fMRI activity for the vertical room condition than the horizontal condition. I also tested whether two rooms in a diagonal relationship were more distinguishable than either the vertically or horizontally adjacent room due to physical or perceptual distance. For this comparison, I used a repeated measures ANOVA with 3 room types as a main factor. In Figure 4.8B, I also plot the same room condition for reference purposes. Since the room encoding region was defined by the “same room < different room” contrast, the “same room” should be associated with reduced activity in all regions.

A significant main effect was found in all regions except for the left retrosplenial cortex (RSC_R, F(2,58)=3.8, p=0.029; RSC_L, F(2,58)=2.4, p=0.10; PHC_R, F(2,58)=3.2, p=0.049; postHC_R, F(2,58)=3.5, p=0.036; postHC_L, F(2,58)=3.6, p=0.032). Post-hoc t-tests showed that this main effect was driven by a small difference between the vertical and diagonal conditions ("ver" versus "diag", RSC_R, t(29)=-2.4, p=0.022; PHC_R, t(29)=-2.5, p=0.017; postHC_R, t(29)=-2.3, p=0.031; postHC_L, t(29)=-2.3, p=0.028). The diagonal condition evoked a larger signal than the vertical condition, implying that two rooms in a diagonal relationship are more differently encoded than two rooms on top of each other. None of the regions showed a significant difference between the vertical and horizontal conditions.

As a side note, the sign of the mean activity (beta) was negative in the hippocampus, implying that the activity was lower during the stimulus presentation period (the virtual navigation and subsequent object-location memory test) compared to the fixation cross inter-trial interval. In the literature, the hippocampus is often reported to show negative beta values during stimulus presentation or task periods (Bakker et al. 2008; Evensmoen et al. 2015; Hodgetts et al. 2015; Brodt et al. 2016). I believe that the absolute beta value of a single condition has little meaning in this study as the implicit baseline (inter-trial interval) was not a meaningful experimental condition. The current study explicitly focussed on comparisons between the main experimental conditions such as the "same room" versus the "horizontal room". The comparisons showed the predicted pattern of repetition suppression, with the fMRI signal associated with the "same" condition reduced compared to the different room conditions.

4.3.3 Supplementary analysis: room versus view encoding

In order to know whether the retrosplenial cortex, parahippocampal cortex and posterior hippocampus encoded view information associated with each room and/or

more abstract spatial knowledge about the room, I conducted a supplementary analysis that separated the same room condition into sub-categories of same view and different view conditions. I then compared them to the different room condition (see Section 4.2.7). I observed repetition suppression effects even when participants visited the same room but approached it from a different view (Figure 4.9; one-sided paired t-tests: “same room, different view” < “different room”, RSC_R, $t(29)=-2.1$, $p=0.021$; RSC_L, $t(29)=-2.4$, $p=0.011$; PHC_R, $t(29)=-1.9$, $p=0.034$; postHC_R, $t(29)=-1.8$, $p=0.041$; postHC_L, $t(29)=-1.9$, $p=0.034$). This suggests that these regions contained abstract room information that was not limited to the exact view. However, there was also evidence for view encoding in some regions. For example, visiting the same room from the same view evoked significantly less activity compared to visiting the same room from different view in the right retrosplenial cortex, parahippocampal cortex and posterior hippocampus (one-sided paired t-tests: “same room, same view” < “same room, different view”, RSC_R, $t(29)=-2.5$, $p=0.010$; PHC_R, $t(29)=-3.4$, $p=0.001$; postHC_R, $t(29)=-1.8$, $p=0.041$). In contrast, the left retrosplenial cortex and left posterior hippocampus did not show any significant differences between the same view and different view ($p>0.1$). In summary, left retrosplenial cortex and left posterior hippocampus showed relatively pure room encoding that was independent of view. Other regions showed additional view dependency, and this was particularly strong in right parahippocampal cortex.

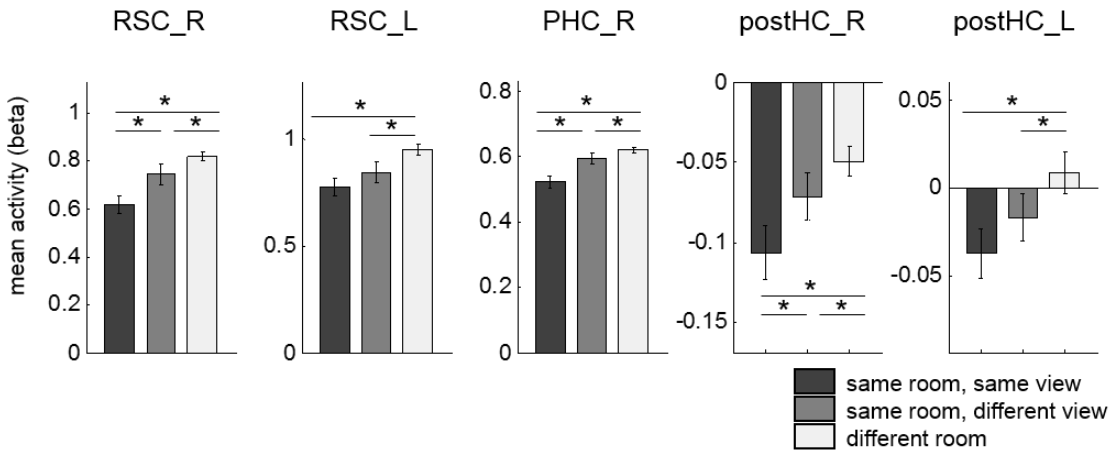


Figure 4.9 Room and view encoding. Even when participants visited the same room from an approach with a different view, the fMRI signal was smaller than when visiting a different room ("same room, different view" < "different room"), suggesting the existence of abstract room information that is not explained by a particular view. However, RSC_R, PHC_R and postHC_R showed additional view dependency ("same room, same view" < "same room, different view"). Error bars are SEM adjusted for a within-subjects design (Morey 2008). *p<0.05.

4.4 Discussion

In this study I investigated how a multi-compartment 3D space (a multi-level gallery building) was represented in the human brain using behavioural testing and fMRI repetition suppression analyses. Behaviourally, I observed faster within-room egocentric spatial judgments and a priming effect of visiting the same room in an object-location memory test, suggesting a segmented mental representation of space. At the neural level, I found evidence of hierarchical encoding of this 3D spatial information, with the left anterior hippocampus containing local corner information within a room, whereas retrosplenial cortex, parahippocampal cortex and posterior hippocampus contained information about the rooms within the building. Furthermore, both behavioural and fMRI data were concordant with unbiased encoding of vertical and horizontal information.

I consider first the behavioural findings. There is an extensive psychological literature suggesting that space is encoded in multiple "sub-maps" instead of a flat single map. Accuracy and/or reaction time costs for between-region spatial judgments (McNamara

et al. 1989; Montello and Pick 1993; Han and Becker 2014), and context swap errors, where only the local coordinate is correctly retrieved (Marchette et al. 2017), are evidence for multiple or recurring sub-maps. Here, I observed faster response for within-room direction judgments and a behavioural priming effect of visiting the same room during a spatial memory task. These findings are therefore consistent with the idea of a segmented representation of space.

Importantly, the current study examined regionalisation in 3D space and compared the influence of vertical and horizontal boundaries for the first time. Some previous studies have suggested a bias in dividing space in the horizontal plane (Jovalekic et al. 2011; Thibault et al. 2013; Flores-Abreu et al. 2014). The horizontal planar encoding hypothesis predicts an additional behavioural cost for spatial judgments across floors and priming effects for the rooms within a same floor. However, I did not find any significant difference in performance for spatial judgments across vertical and horizontal boundaries, or priming effects for rooms on the same floor. Although the absence of significant difference does not necessarily mean equivalence, the most parsimonious interpretation would be that each room within this 3D space was similarly distinguishable.

This fits with the symmetric encoding of 3D location information in a semi-volumetric space tested in Experiment 1, and isotropic place fields found in bats (Yartsev and Ulanovsky 2013). One concern might be that the small number of rooms in the virtual building in the current study allowed participants to encode each room categorically without being truly integrated in a 3D spatial context. However, in order to be successful at the egocentric judgments task across rooms (mean accuracy was 80%), the participants must have had an accurate representation of the 3D building. Testing an environment with more floors and rooms in the future could facilitate the search for any additional hierarchies within 3D spatial representations. For example, rooms might

be further grouped into the horizontal plane or a vertical column in a more complex environment. It might also help to reveal subtle differences, if they exist, between vertical and horizontal planes.

At the neural level, I found that the fMRI response in the left anterior hippocampus was associated with local corner information that was generalised across multiple rooms. This fits well with previous findings that hippocampal place cells in rodents fire at similar locations within each segment of a multi-compartment environment (Derdikman et al. 2009; Spiers et al. 2015). This common neural code enables efficient encoding of information. For example, the 16 locations in the virtual building could be encoded using only 8 unique codes (4 for distinguishing the corners of rooms and 4 for distinguishing the rooms themselves) given its regular substructures. This room-independent representation in the anterior lateral hippocampus can also be seen as a 'schematic' representation of space (Marchette et al. 2017) where the regular structure of the environment is extracted. Furthermore, there is evidence that the ability of the hippocampus to extract regularity in the world is not limited to the spatial domain. A previous fMRI study found that temporal order information in the hippocampus generalised across different sequences (Hsieh et al. 2014). Statistical learning of temporal community structure has also been associated with the hippocampus (Schapiro et al. 2016) and, interestingly, localised to the anterior portion. Rodent electrophysiology and modelling work also suggests that ventral hippocampus (analogous to the human anterior hippocampus) is well suited to generalising across space and memory compared to dorsal hippocampus (analogous to the human posterior hippocampus) (Keinath et al. 2014).

In addition to generalised within-room information, it is also important to know a room's location to identify one's exact position within a building. I found that multiple brain regions represented room information, with the retrosplenial cortex exhibiting the most

reliable room repetition effect. At first this finding might seem surprising, given that head direction information has been consistently associated with the retrosplenial cortex in humans and rodents (Baumann and Mattingley 2010; Marchette et al. 2014; Shine et al. 2016; Jacob et al. 2017). In this experiment, participants faced paintings on opposite walls within a room. Therefore, if retrosplenial cortex encoded the participant's facing direction, local corner encoding would be expected instead of room encoding. However, numerous findings suggest that retrosplenial cortex encodes more than head direction; processing of multiple spatial features such as location, view, velocity and distance have been linked with this region (Cho and Sharp 2001; Sulpizio et al. 2014; Alexander and Nitz 2015; Chrastil et al. 2015). Moreover, retrosplenial cortex has been found to be involved in both a location and an orientation retrieval task when participants viewed static pictures of an environment during fMRI (Epstein et al. 2007).

Given the rich repertoire of spatial, visual and motor information the RSC processes, it is perhaps not surprising that some studies observed local head direction signals and others found global head direction information in this region (Marchette et al. 2014; Shine et al. 2016). This might also be influenced by functional differences within the retrosplenial cortex, or indeed laterality effects. In our experiment, the right retorpslenial cortex showed stronger repetition suppression when participants visited the same room from same view compared to when they visited the same room from different view, whereas the left retrosplenial cortex response was only influenced by the repetition of the room.

Retrosplenial cortex might have a role in integrating local representations within a global environment. A recent theory about the neural encoding of large-scale 3D space proposed that 3D space is represented by multiple 2D fragments, and retrosplenial cortex was proposed to be a candidate area for stitching these together (Jeffery et al. 2015). In the current experiment, room information can be broadly viewed as the

orienting cue within a building that allows integration of the fragmented space (each room). For localisation and orientation of local representations within a larger spatial context, landmark information is crucial and retrosplenial cortex supports the learning of and processing of stable landmarks (Auger et al. 2012, 2015). Furthermore, head direction signals in retrosplenial cortex are dominated by local landmarks (Jacob et al. 2017).

The second region that represented room information was the parahippocampal cortex. It also showed a strong view dependency in addition to room information. This contrasts with the left retrosplenial cortex which only showed a room repetition effect. Together these findings are consistent with the proposed complementary roles of the parahippocampal cortex and retrosplenial cortex in scene perception, whereby the former seems to respond in a view-dependent manner whereas the retrosplenial cortex represents integrative and more abstract scene information. For example, it has been shown that when participants saw identical or slightly different snapshot views from one panoramic scene, retrosplenial cortex showed fMRI repetition effects for both identical and different views, but parahippocampal cortex only exhibited repetition suppression for the identical view (Park and Chun 2009). In addition, multivoxel patterns in retrosplenial cortex have been observed to be consistent across different views from each location, whereas this was not the case for the parahippocampal cortex (Vass and Epstein 2013).

Along with retrosplenial cortex and parahippocampal cortex, the final area to represent room information was the posterior hippocampus. The similarity in spatial encoding between these regions might be predicted from their close functional and anatomical connectivity (Kobayashi and Amaral 2003; Kahn et al. 2008; Blessing et al. 2016). It is notable that in Experiment 1, I also found that posterior hippocampus and retrosplenial cortex encoded the same type of spatial information (vertical direction) while anterior

hippocampus encoded a different type of spatial information (3D location). In that study, different vertical directions resulted in more distinguishable views, although direction information observed in the multivoxel patterns remained significant after controlling for low level visual similarities. The current result does not fit precisely with accounts that associate the posterior hippocampus with a fine-grained spatial map (Poppenk et al. 2013; Evensmoen et al. 2015). In fact, my findings could be interpreted as evidence in the opposite direction, namely that coarser-grained representations of the whole building engage the posterior hippocampus. Nevertheless, overall, the anterior and posterior hippocampal findings provide further evidence of functional differentiation down the long axis of the hippocampus (Baumann and Mattingley 2013; Poppenk et al. 2013; Strange et al. 2014; Zeidman and Maguire 2016).

Finally, as with the behavioural data, I examined the fMRI data for possible differences between the horizontal and vertical planes. I did not find significant differences in fMRI amplitude between the vertical and horizontal conditions in the brain structures that contained room information. This neural finding is consistent with the behavioural results of similar accuracy and response time for spatial judgments across vertical and horizontal rooms, and similar priming effects for each room. These results fit well with an isotropic representation of 3D space, similar to Experiment 1.

Again, as with the behavioural data, one concern might be that each room is represented in retrosplenial cortex, parahippocampal cortex and posterior hippocampus in a categorical, semantic manner without consideration of their physical 3D location in building. However, as I discussed earlier, egocentric spatial judgments in the pre-scan task prevented participants from separately encoding each room without the 3D spatial context. Furthermore, I found that visiting a diagonal room evoked a larger fMRI signal than visiting a vertical room, and this finding cannot be explained if each room was encoded in a flat manner without a spatial organisation. This implies

that the neural representation of two rooms in a diagonal relationship were more distinguishable than two rooms on top of each other. This might be due to the change in two axes for the diagonal room (vertical and horizontal) compared to a change along only one axis for the vertical room, or simply because of a longer distance between two rooms in diagonal relationship. Distance encoding has been previously reported in parahippocampal cortex and RSC (Marchette et al. 2014; Sulpizio et al. 2014).

To disambiguate these possibilities, a larger environment consisting of multiple vertical and horizontal sections should be tested. For example, if the physical distance between the rooms is the main factor for neural dissimilarity, two rooms on the same floor that were separated by 5 other rooms (e.g. Rm101 and Rm106) would be more distinguishable than two rooms that are both vertically and horizontally adjacent (e.g. Rm101 and Rm 202). If the change in both vertical and horizontal axes always has a greater effect than the change in one axis, the diagonal rooms would be more distinguishable than horizontally or vertically aligned rooms regardless of distance. Use of a larger environment would also widen the scope for detecting subtle differences, if any, in the vertical and horizontal axes.

In addition to absolute physical distance, path or navigation distance is also a consideration. For example, a typical multi-level building like the one used in the current study has limited access points for movement across the floors. People cannot directly move up to the room above through the ceiling; rather they have to use stairs or elevators which are often sparsely located in the building. Thus, two rooms on top of each other are further apart in terms of actual navigation than two rooms side-by-side on the same floor, even when absolute distances are identical or the vertical rooms have even shorter physical distance than the horizontal rooms. Representation of space in the hippocampus is not only influenced by absolute distance but also by path distance (Howard et al. 2014), and it has also been suggested that topology instead of

physical geometry is encoded in hippocampal place cells (Dabaghian et al. 2014). It would be intriguing to systematically investigate the effect of physical and path distance, and the potential interaction with vertical/horizontal boundaries, in future studies.

In summary, this second experiment provided novel evidence showing that a multi-compartment 3D space was represented in a hierarchical manner in the human brain, where within-room corner information was encoded by the anterior lateral hippocampus and room (within the building) information was encoded by retrosplenial cortex, parahippocampal cortex and posterior hippocampus. Moreover, similar to the findings in the previous experiment where a 3D lattice structure was used, the behavioural and neural findings showed equivalence of encoding for vertical and horizontal information, suggesting an isotropic representation of 3D space even in the context of multiple spatial compartments. In the next experiment, I will test how a completely continuous 3D space is encoded in the brain.

Chapter 5 Experiment 3: Direction encoding in a volumetric space

5.1 Introduction

In the previous two experiments, participants' movements were restricted to either narrow tracks of the 3D lattice structure (Experiment 1) or the floors of a multi-level building (Experiment 2). Although humans are normally bound to such surfaces due to gravity, they sometimes explore open and volumetric 3D spaces – for example, when they dive in deep oceans, control aircraft and manoeuvre in the microgravity of a spaceship. For navigation in volumetric 3D space, which allows continuous rotations along all three axes, knowing one's 3D orientation is crucial. In bats, a 3D compass system – head direction cells tuned to either azimuth, pitch, roll or the conjunctive components of 3D directions – was found in the presubiculum. However, the neural mechanisms of 3D head direction encoding still remain largely unknown in non-flying mammals because most previous studies have been conducted in a surface based environment (flat or tilted, as described in Section 1.3) rather than a full volumetric space. Technical difficulties associated with recording neural activity in non-flying animals in volumetric space (e.g. the requirement for underwater recording devices or microgravity simulators with a centrifuge) might be one reason for the dearth of knowledge.

This is where studying the human brain with VR techniques becomes advantageous. With recent advances in affordable and easy-to-use VR technologies, I could build a virtual zero gravity spaceship environment where participants felt like they were floating or flying. I then tested for the existence of 3D direction signals in the human brain while participants were moving in a wide range of 3D directions. This is in contrast to Experiment 1, which only examined head direction signals on a 29° tilted track.

Also in comparison to the previous two experiments, I introduced behavioural training with a head-mounted display ("VR goggles") here in Experiment 3. The VR goggles provided a much more immersive sensation of 3D space than a conventional screen because they tapped into multisensory (visual, vestibular and proprioceptive) inputs, which are crucial for the head direction system. It was previously suggested that exposure to both visual and vestibular stimuli during a pre-scan period with VR goggles might lead to a recapitulation of body-based information during later fMRI scanning, where only visual input is available due to head immobilisation (Shine et al. 2016). It was further noted that pre-exposure to vestibular cues could be particularly important for detecting head direction signal in the thalamus (Shine et al. 2016).

In this chapter, I tested whether vertical (pitch) and horizontal (azimuth) direction information was encoded in the brain using fMRI multivoxel pattern similarity analysis. I focused on the brain regions that are known to contain head direction cells – retrosplenial cortex, thalamus, subiculum and entorhinal cortex, but I also searched for direction encoding beyond the ROIs using a whole-brain searchlight analysis. In addition, I examined whether participants were equally good at encoding vertical and horizontal direction using behavioural tests. Neural correlates of individual differences in direction judgment performance was also investigated.

5.2 Methods

5.2.1 Participants

Thirty healthy adults took part in the experiment (16 females; mean age = 25.9 ± 4.8 years; range 19-36 years; all right-handed). All had normal or corrected-to-normal vision and gave informed written consent to participation in accordance with the local research ethics committee. To minimise the risk of VR-related feelings of nausea,

during this experiment, I primarily recruited participants who had previous experience of using VR goggles without discomfort.

5.2.2 The virtual environment

The virtual environment was composed of two distinctive rectangular compartments, called here room A and room B for convenience, which were linked by a corridor (Figure 5.1A). Participants were instructed that they were inside a virtual zero gravity "spaceship" where they could move up, down, forwards and backwards freely. The walls, floors and ceilings had different textures which provided orientation cues. A snapshot of the virtual environment as seen from a participant's perspective during scanning is shown in Figure 5.1B-E. The virtual environment was implemented using Unity 5.4 (Unity Technologies, CA, United States) with textures and sci-fi objects downloaded from the Unity Asset Store. The virtual environment can be viewed here: www.fil.ion.ucl.ac.uk/Maguire/spaceship3D.

The virtual spaceship was rendered on two different mediums for pre-scanning tasks and scanning tasks respectively: VR goggles (Samsung Gear VR, model: SM-R322 with a Samsung Galaxy S6 phone) and a standard computer screen (20.1 inch LCD monitor, Dell 2007FP with Dell Optiplex 980 computer).

The VR goggles provided participants with a fully immersive sensation of 3D space via their built-in head motion tracking system, stereoscopic vision and wide field-of-view (96°). A rotation movement in the VR display was made by a participant's physical head rotation, and a translational movement was made by a button press on the Bluetooth controller (SteelSeries Stratus XL, Denmark). For example, a participant could move up towards the ceiling in the virtual spaceship by physically looking up and pressing the forward button on the controller. To rotate to the right, they physically

rotated their head to the right or rotated their whole body when the required rotation was beyond the range of neck rotation. For ease of rotation, participants were seated on a swivel chair throughout.

During fMRI scanning, participants watched a video that was rendered on a standard computer screen (aspect ratio = 4:3). The video was a first-person perspective that gave the participants the feeling of moving in the virtual spaceship. The stimuli were projected on the screen using a projector at the back of the MRI scanner bore (Epson EH-TW5900 projector), and participants saw the screen through a mirror attached to the head coil. The screen covered a field of view of ~19° horizontally and ~14° vertically.

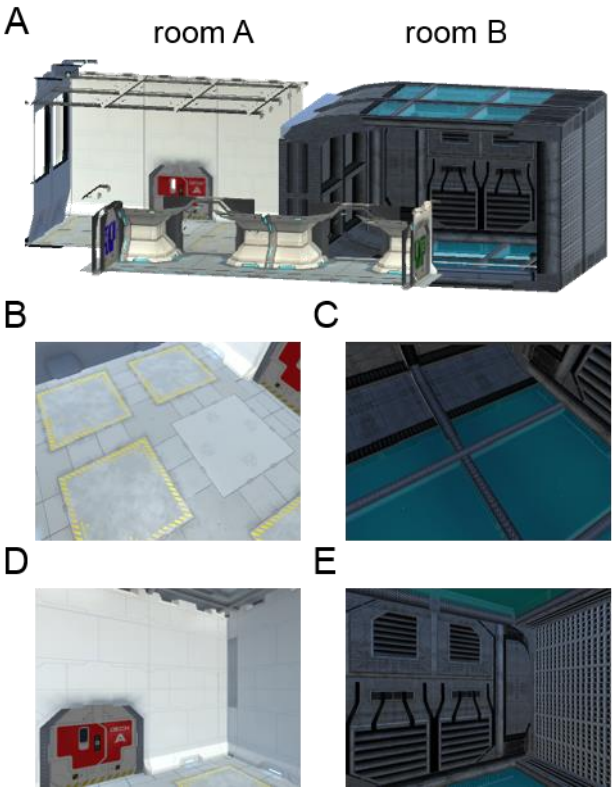


Figure 5.1 The virtual environment. (A) An overview of the virtual spaceship composed of two rooms linked by a corridor. Some walls are shown as transparent here for display purposes. (B-E) Example views from a participant's perspective during scanning. (B) and (C) are views when a participant is facing down in room A and room B, respectively. (D) and (E) are views when a participant is facing straight ahead in room A and room B, respectively.

5.2.3 Procedure

All participants completed the experimental tasks in the following order: familiarisation and pointing task with VR goggles prior to scanning, direction judgment task during scanning (which was preceded by a short practice of the scanner task) and post-scan debriefing.

Pre-scan: familiarisation

Participants first familiarised themselves with the VR goggles and the controller during a simple “ball collection” task (duration = 5 minutes). Multiple balls were scattered in the spaceship and participants moved to the balls one-by-one. When they arrived at each ball, they received auditory feedback (a ‘ping’ sound). The primary purpose of this task was to familiarise participants with controlling their movements in the virtual environment via head/body rotations and button presses on the controller. In addition, participants were asked to pay attention to the overall layout of environment for later tasks. This ball collection task also ensured that the participants visited every part of the virtual environment.

Pre-scan: pointing task

After the initial familiarisation period, participants performed a spatial memory task which required a good sense of direction in the virtual 3D spaceship (duration = 15 ± 2 minutes, Figure 5.2). While wearing the VR goggles, at the beginning of each trial, participants were placed in one of the two rooms in the spaceship. There was one floating ball in the room and participants had to memorise the location of the ball. During this encoding phase (duration = 18 sec), participants could move freely and they were instructed to look at the ball from various directions and distances in order to learn the precise location of the ball. The ball then became invisible and a participant was transported to a random location. Participants were then required to look towards

the remembered location of the ball and press a button when they had made their decision, after which feedback was provided in the form of the angular deviation from the true angle (Figure 5.2B). Throughout the task (encoding and testing), a small red crosshair was shown to aid the orientation (Figure 5.2B).

In the majority of trials ("within-room", n = 16), testing took place in the same room where the ball was located during encoding. There were six additional trials where testing occurred in the other room; for example, participants encoded the ball's location in room A but they were placed in room B during the test phase, requiring them to point to the ball behind the wall. These "across-room" trials were included in order to encourage participants to build an integrated map of the whole spaceship that was not limited to a local room. An integrated mental representation was important for the later fMRI analyses because I searched for direction information that was generalised across the two rooms.

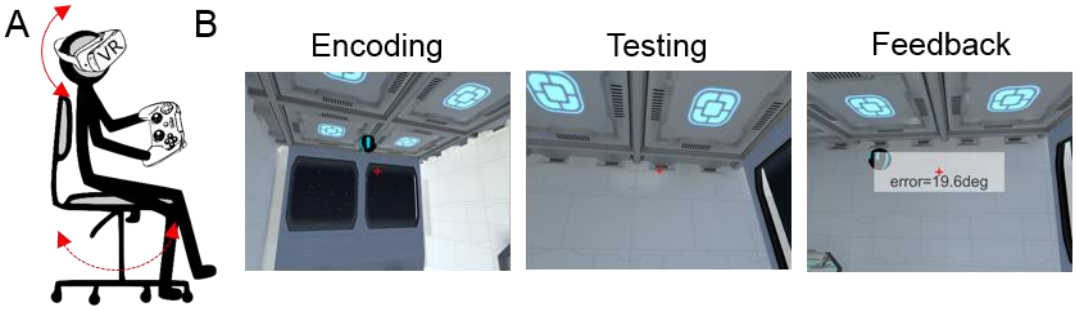


Figure 5.2 The pre-scan pointing task. (A) Wearing VR goggles, participants moved inside the virtual spaceship by rotating their head or body and pressing the forward button on the controller. (B) During the encoding phase, participants memorised the precise location of a ball by looking at it from various directions and distances. During the test phase, they were positioned at a random location and were asked to look towards the remembered location of the balls. Once they had made their decision and pressed the button, feedback was provided in the form of their absolute angular error.

Scanning: direction judgment task

During scanning, participants watched a video rendered on a standard display and performed a direction judgment task. The video provided participants with the feeling that they were flying in a controlled 3D trajectory within the spaceship (Figure 5.3A).

The pre-programmed video allowed tight control of location, direction and timing for all participants. The trajectory consisted of multiple short linear movements (each of 3 sec) followed by rotation (2/2.6 sec). Ideally, I would have sampled all possible head directions in 3D space (from -180° to 180° horizontally and from -90° to 90° vertically), but I restricted the range of linear movement directions in order to acquire reliable measurements of the neural responses to each direction within the limited scanning time. I sampled five levels of horizontal azimuth and five levels of vertical pitch from -60° to 60° with 30° steps, resulting in 25 unique 3D directions (Figure 5.3B,C).

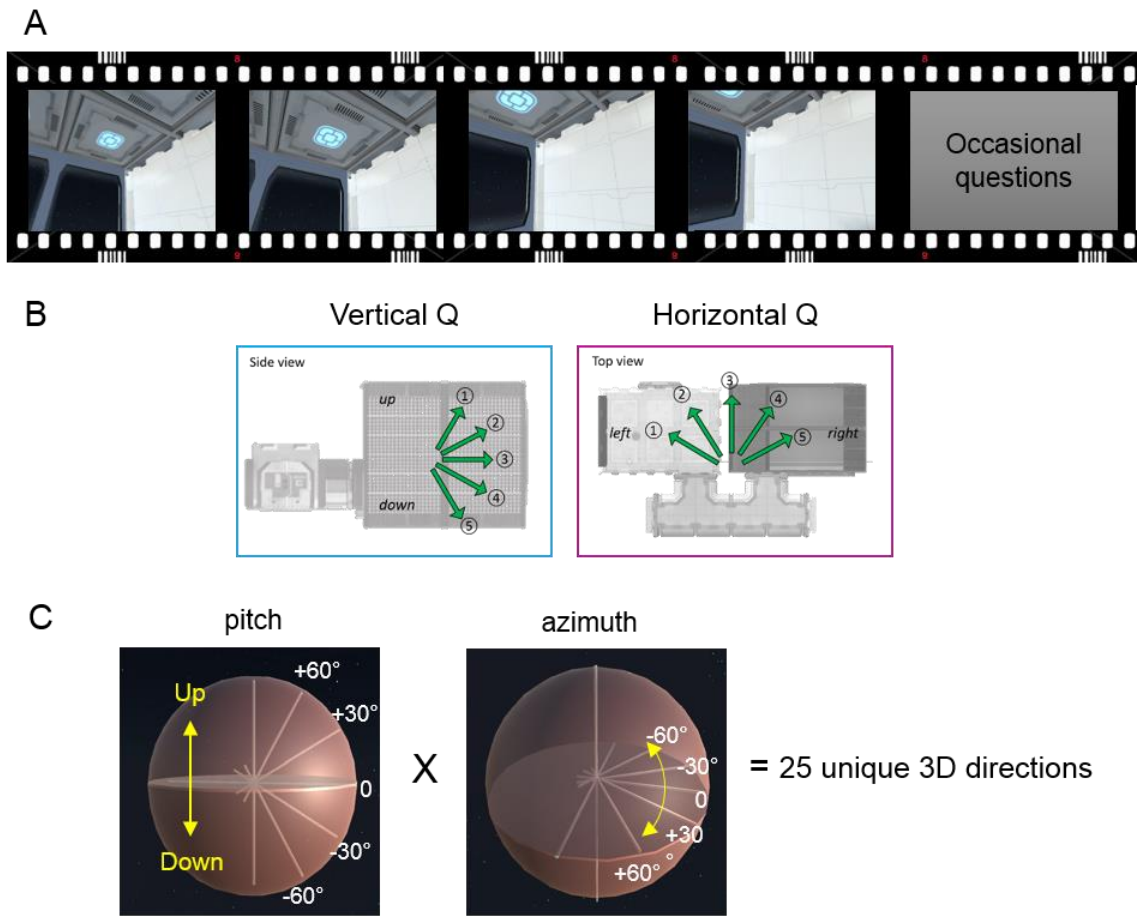


Figure 5.3 The direction judgment task during scanning. (A) Participants watched a video that provided the sensation that they were moving inside a virtual spaceship. (B) Occasionally, participants were asked to indicate either the vertical or horizontal direction of their last movement. (C) During scanning, five levels of pitch and azimuth were sampled, resulting in 25 unique 3D head directions.

A smooth trajectory was used without abrupt rotations (e.g. if a participant's previous direction was 0°, the participant would be facing 0 ± 30° after a turn). A constant linear

and angular velocity was applied in order to control the velocity, which can modulate the firing rate of head direction cells (Stackman and Taube 1998). If a participant reached the boundary of the spaceship, a blank screen appeared for two seconds and then the next trajectory started from the other end of the spaceship. On a quarter of the trials, a question screen appeared immediately after a linear movement and participants indicated the direction of their last movement by pressing a button pad (a 5-alternative forced choice question with a time limit of 5 sec, mean response time = 1.7 ± 0.4 sec, Figure 5.3B). When they made an incorrect choice, the correct direction was then shown on the screen. This direction judgment task ensured participants kept track of their movements during scanning. Since vertical or horizontal direction questions were randomly presented, participants were required to know their 3D direction throughout. The two rooms of the spaceship were visited alternatively for each of four scanning sessions. Half of the participants started in room A and half started in room B. Each scanning session lasted ~11 minutes with a short break between the sessions, making a total functional scanning time of 50 minutes.

Of note, roll rotations are also possible in 3D space (e.g. where the right ear is down or the left ear is down, Figure 1.7B). However, roll rotation is less important for navigation because the azimuth and pitch determine the forward movement vector (Figure 1.1B) and the roll is simply a rotation around this vector. It has been proposed that animals tend to avoid roll rotations for the efficient computation of 3D direction, and that might be the reason for the paucity of cells tuned to roll in bats (Finkelstein et al. 2015, 2016). The head motion of participants recorded during the pre-scan VR experiment indeed confirmed that the roll was close to zero. Thus, I focused on the azimuth and pitch and not roll in this experiment.

Post-scan debriefing session

After scanning, participants were asked to rate how much they felt immersed in the virtual environment, separately for the pre-scan VR goggle tasks and the scanning task with the conventional display. The multiple choice options were similar to those used in Experiment 1 which were: “I felt like I was really in the spaceship”; “I occasionally thought about the environment as being computer-generated, but overall the environment was convincing and I felt I was moving around in the spaceship”; “I was often distracted by the feeling that I was not in a real environment”. They were also asked whether the vertical or horizontal question was easier during the scanning task.

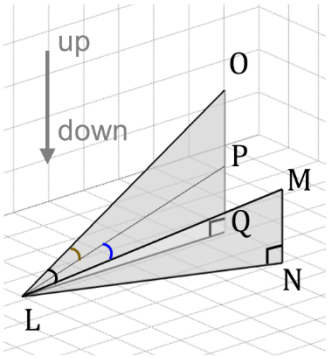
5.2.4 Behavioural analyses

Pre-scan: a pointing task

I first measured the absolute 3D angular error for the within-room trials and the across-room trials. I then tested whether participants were equally good at processing vertical and horizontal information by decomposing the absolute 3D error of each trial into vertical and horizontal components. The geometry of the angular decomposition is shown in Figure 5.4. Let us assume that a participant was at location L and the true location of the ball was M (true direction: L→M). If the participant looked towards O (the actual direction: L→O), the absolute angular error was ∠MLO. I defined the vertical error as the pitch difference between the true direction and the actual direction, which was identical to ∠OLP in Figure 5.4. The horizontal error was defined as the angle between the true direction (L→M) and the looked-at direction where the pitch was matched (L→P), which was ∠PLM. The mean vertical and horizontal error was computed for each subject using the within-room trials, then a paired t-test was used to compare the mean vertical and horizontal error at the group level. I did not include the across-room trials because recalling the location of the ball in a different room involves different cognitive process from recalling within the same room, and the number of

across-room trials was much smaller than within-room trials. As I described earlier, the main purpose of the across-room trials was to encourage participants to build an integrated representation of the whole spaceship. However, even when the across-room trials were included, the result remained similar.

After having observed a significant difference between the vertical and horizontal error, I wondered whether this difference was due to the fact that participants looked up or down with a steep angle to indicate the ball's location, which is infrequently performed in our daily lives. To test this, I divided the trials into quadrants based on the pitch of the true direction. The mean and standard deviation for pitch for each quadrant were as follows: 1.0 ± 1.1°, 9.3 ± 1.5°, 30 ± 2°, 63 ± 4.5°. I then compared the vertical and horizontal error in these subsets of trials which ranged from almost flat (1.1°) to very steep (63°) conditions. I also divided the trials into those when participants performed the pointing task in room A and room B to check whether the difference in the vertical and horizontal error was unique to a particular environment.



$$true\ direction = \overrightarrow{LM}$$
$$response\ direction = \overrightarrow{LO}$$
$$absolute\ error = \angle OLM$$
$$vertical\ error = \angle OLP$$
$$horizontal\ error = \angle PLM$$

$$* \angle PLQ = \angle MLN$$

Figure 5.4 Decomposition of the angular error in the pointing task. When a participant who was located at L looked towards O instead of M (the true location of the ball), the absolute error was ∠MLO. The vertical error was defined as ∠OLP and the horizontal error was defined as ∠PLM.

Scanning: direction judgment task

I first measured overall accuracy (chance = 20%) to confirm whether participants knew their 3D direction in the virtual environment. I then tested whether participants were

better at knowing their vertical or horizontal direction. In comparing vertical and horizontal performance, it was more informative to consider how much a participant's response direction deviated from the true direction and not just whether they made a correct or wrong judgement. For example, when the true direction was 1 ("steep up", Figure 5.3B), a participant could have selected either 2 ("shallow up") or 4 ("shallow down") and these errors were quantitatively different. To quantify the angular sensitivity, I defined the angular error of each trial by assigning 0° when participants chose the correct response; 30° when participants chose the adjacent direction such as 2 for 1, 60° when participants chose the direction 2 steps away from the correct direction such as 3 for 1, and so on. The mean angular error and response time was computed for vertical and horizontal questions respectively in each participant (excluding trials where participants did not respond within the time limit of 5 sec, which occurred very rarely - less than 1% of trials) and paired t-tests were used to compare the vertical and horizontal angular error and response time at the group level.

Post-scan debriefing session

I counted the number of responses for each option in the multiple choice debriefing (i.e. how much they felt immersed in the virtual environment and whether the vertical or horizontal question was easier, or whether they were similar).

5.2.5 Scanning and pre-processing

Functional MRI scans (voxel size = 3 x 3 x 3 mm, 44 slices, TR = 3.08 sec) and a structural MRI scan (1 x 1 x 1 mm) were acquired using a 3T Siemens Trio scanner and preprocessed using SPM12. The scanning parameters and preprocessing are described in detail in Section 2.5.4. To summarise, after removing the first 5 volumes to allow for T1 equilibration effects, functional images were realigned to the first volume of each session and geometric distortion was corrected by the SPM unwarp function

using the fieldmaps. Each participant's anatomical image was then coregistered to the distortion corrected mean functional images. Functional images were normalised to MNI space, and were left unsmoothed for multivoxel pattern analysis to preserve the fine scale activity patterns.

5.2.6 fMRI analysis: ROI

Anatomical ROIs

The thalamus ROI was extracted from the AAL atlas (Tzourio-Mazoyer et al. 2002) (Figure 5.5). Although head direction cells have been mainly found in the presubiculum in animals, here I used a broader subiculum mask containing pre/parasubiculum because it was not feasible here to distinguish these fine structures with the standard resolution of the fMRI images. This broad subiculum ROI was delineated on the group averaged structural MRI scan following the protocol in Dalton et al. (2017) (Figure 5.5). I used the retrosplenial cortex and entorhinal cortex which I created in Experiment 1 (see Section 3.2.6 for details). The number of functional voxels (3 x 3 x 3 mm) within each ROI (L = left, R = right) were as follows: thalamus_L, 302; thalamus_R, 286; retrosplenial_L, 158; retrosplenial_R, 135; entorhinal_L, 47; entorhinal_R, 49; subiculum_L, 34; subiculum_R, 34.

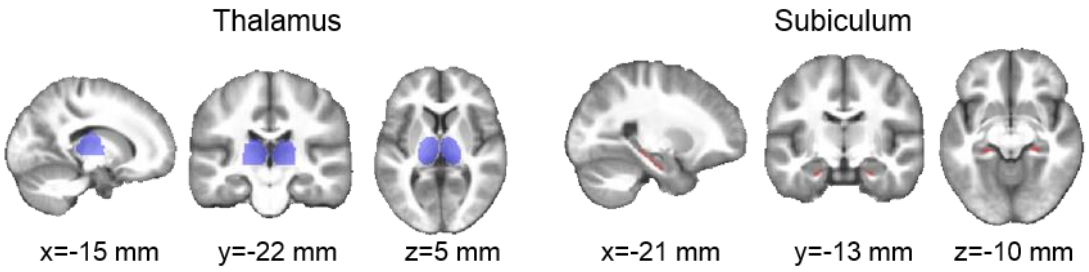


Figure 5.5 The thalamus (blue) and subiculum (red) ROIs are shown on the group averaged structural MRI scan.

Representational similarity analysis – ROI

To examine whether each ROI contained vertical or horizontal direction information, or both, I used a multivoxel pattern analysis similar to that used in previous studies (e.g. Carlin et al. 2011; Vass and Epstein 2013). This analysis compared the neural similarity measures to model similarity values predicted from different encoding hypotheses. As a first step in the analysis, I estimated the neural responses to each 3D head direction using a GLM. The design matrix contained 25 main regressors which were boxcar functions that modelled the period when participants moved straight in one of 25 directions (5 levels for vertical pitch x 5 levels for horizontal azimuth) (Figure 5.3C), convolved with the SPM canonical hemodynamic response function. In addition, the occasional questions and blank screen periods (when participants came to the border of the spaceship) were separately modelled in the GLM as regressors of no interest. Six head realignment parameters were also included as nuisance regressors. The GLMs were applied for each scanning session in each participant.

I then computed the neural representational similarities between each direction using Pearson's correlation between the multivoxel T-values within the ROIs that were estimated in the preceding GLM. Crucially, representational similarity was calculated between neural responses to the 3D directions when a participant was in different rooms of the virtual spaceship. This ensured that neural similarity was calculated between independent scanning sessions (because each room was alternatively visited in separate scanning sessions). More importantly, this across-room similarity analysis allowed me to detect relatively pure spatial direction information that was independent of view, which is naturally linked to head direction.

Figure 5.1B-E shows example views when participants moved in two different directions in the two rooms. For instance, when I calculated the neural similarity between the "down-left" direction and "flat-right" direction, the correlation between

“down-left” in room A (Figure 5.1B) and “flat-right” in room B (Figure 5.1E) and the correlation between “down-left” in room B (Figure 5.1C) and “flat-right” in room A (Figure 5.1D) were averaged. Therefore, the higher neural similarity between pairs of directions was not attributable to the higher visual similarity between the views associated with the directions within the same room. In summary, I calculated a symmetric 25 x 25 pairwise representational similarity matrix for each participant. I converted the similarity value (Pearson’s r) into a dissimilarity value by inverting it (1-r) for ease of later analysis.

Finally, these neural dissimilarity measures were compared to the vertical and horizontal directional encoding models using multiple regression. I used encoding models in which neural dissimilarity is linearly dependent on the difference in pitch or azimuth between two directions (Figure 5.6). For example, a vertical encoding model predicts that neural similarity between two directions that have the same pitch will be the highest, while neural similarity between two directions where pitch is -60° and 60° respectively will be the lowest, regardless of azimuth. I also included a visual texture similarity model to control for low-level visual similarity across the rooms. Therefore, pitch distance, azimuth distance, visual similarity and a constant term were included in the multiple regression model. I computed the visual texture similarity using Renninger and Malik's (2004) model as I described in Section 3.2.8. This visual control model was also used in previous studies that investigated direction encoding (Vass and Epstein 2013; Sulpizio et al. 2014).

The regression coefficients (beta) of each participant were fed into a group level analysis to test whether the neural response in the selected ROIs was explained by the vertical or horizontal encoding model. I tested whether the regression coefficient was significantly greater than zero using a t-test. I also performed paired t-tests to compare

the betas of the vertical and horizontal models to ascertain whether the neural response was more sensitive to one model or the other.

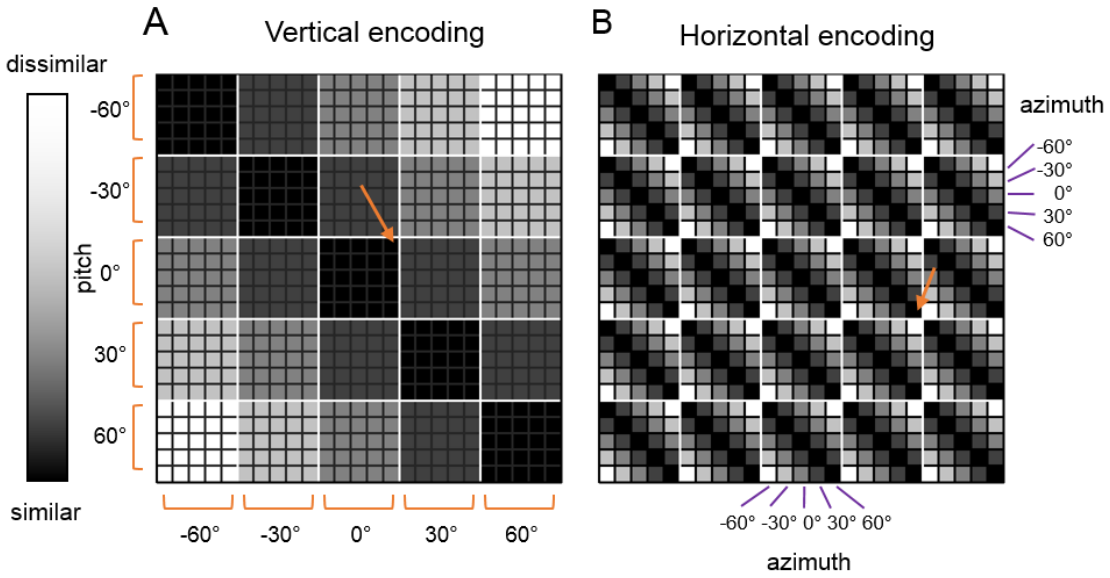


Figure 5.6 The model representational similarity matrix for the fMRI analysis. The representational similarity matrix (25 x 25) contained pairwise similarity values between each of the 25 unique directions. (A) If the vertical direction was encoded, the neural similarity between the directions that share a common vertical tilt, pitch would be high (dark colours), e.g. between (pitch, azimuth) = (0°,-60°) and (0°,60°) as indicated by the orange arrow. Similarity falls as the difference in pitch between two directions increases. (B) If the horizontal direction is encoded, the neural similarity between the directions that share a common horizontal angle, azimuth would be high (dark colours), e.g. between (pitch, azimuth) = (0°,60°) and (30°,60°), as indicated by the orange arrow. Similarity falls as the difference in azimuth between two directions increases.

5.2.7 Neural correlates of individual differences

I also tested whether there was a relationship between the direction information represented in the multivoxel pattern in the ROIs and behavioural performance during the scanning direction judgment task. For the behavioural performance measure, I used the mean angular error pooled across the vertical and horizontal direction questions, given that the vertical and horizontal errors were highly correlated (Pearson's $r = 0.81$, $p < 0.001$). I defined the direction information in individuals as the regression coefficient for the vertical and horizontal direction model in my ROIs. Pearson's correlation coefficient was used for the significance test.

5.2.8 fMRI analysis – searchlight

While my main interest was in testing for the existence of vertical and horizontal direction information in my pre-specified ROIs, I also conducted a whole-brain searchlight analysis (Kriegeskorte et al. 2006) to test whether there were any other brain regions sensitive to vertical and horizontal direction. Moreover, the searchlight analysis complemented findings from the ROI analysis particularly in the thalamus, by providing additional anatomical localisation, given that the thalamus is a heterogeneous structure containing multiple functionally distinct nuclei. For localisation of thalamic structures, I relied on the WFUpickAtlas software (Lancaster et al. 1997, 2000; Maldjian et al. 2003) and a human thalamus atlas (Morel 2007). I performed the same representation similarity analysis using the multivoxel T-values within small spherical ROIs (radius 6 mm) centred on each voxel across the whole brain. This generated regression coefficient maps for vertical and horizontal encoding models for each participant. These maps were fed into the group-level analysis (one-sample t-test) in SPM. I report voxel-wise p-values corrected for my anatomical ROIs. For the rest of the brain, I report voxels that survived whole-brain multiple comparison correction (family-wise error rate of 0.05).

5.3 Results

5.3.1 Behavioural results – pointing task

The pre-scan pointing task involved participants wearing the VR goggles and looking towards the remembered position of balls while they were positioned at random locations. The group mean angular error was 21 ± 9° for within-room trials. Figure 5.2B shows an example view when a participant made a ~20° error, and demonstrates that the participant's pointing direction (the centre of the screen, a red crosshair) was reasonably close to the target ball. The error for across-room trials was slightly larger

(28 ± 20°). This is unsurprising, because participants had to orient themselves to the target ball behind the wall. Given this overall good level of performance, I am confident that participants went into the subsequent scanning experiment with a reasonable sense of orientation in the 3D virtual environment.

Furthermore, I found that the mean vertical error was significantly smaller than the horizontal error when angular errors were decomposed into the vertical and horizontal components ($t(29)=-3.4$, $p=0.002$, Figure 5.7A). This suggests that participants were more accurate at recalling the ball's location along the vertical dimension. This effect was significant even when room A and room B trials were analysed separately (room A, $t(29)=-3.1$, $p=0.005$; room B, $t(29)=-3.3$, $p=0.003$). Furthermore, the difference cannot be attributed to looking up or down with an unusually steep angle. When I divided the trials into quadrants based on the vertical pitch of the true direction, the vertical error was significantly smaller than the horizontal error for the first three quadrants where the mean vertical tilt was only 1.0°, 9.3°, or 30° (first quadrant, $t(29)=-3.5$, $p=0.001$; second quadrant, $t(29)=-4.1$, $p<0.001$; third quadrant, $t(29)=-2.6$, $p=0.01$; fourth quadrant, $t(29)=0.6$, $p=0.6$).

5.3.2 Behavioural results – direction judgment test

During scanning, participants were moved in a preprogrammed 3D trajectory and were occasionally asked about their movement direction, either vertically or horizontally. The mean accuracy (74 ± 16%) was well above chance level (20%), suggesting that participants were able to keep track of their movement direction. I found that participants made significantly smaller errors for the vertical questions compared to the horizontal questions ($t(29)=-2.4$, $p=0.02$, Figure 5.7B). In contrast, I observed a small, but significant, difference in response time. Participants were slightly faster at the

horizontal questions (vertical = 1.79 ± 0.36 sec, horizontal = 1.67 ± 0.38 sec, t(29)=2.6, p=0.015).

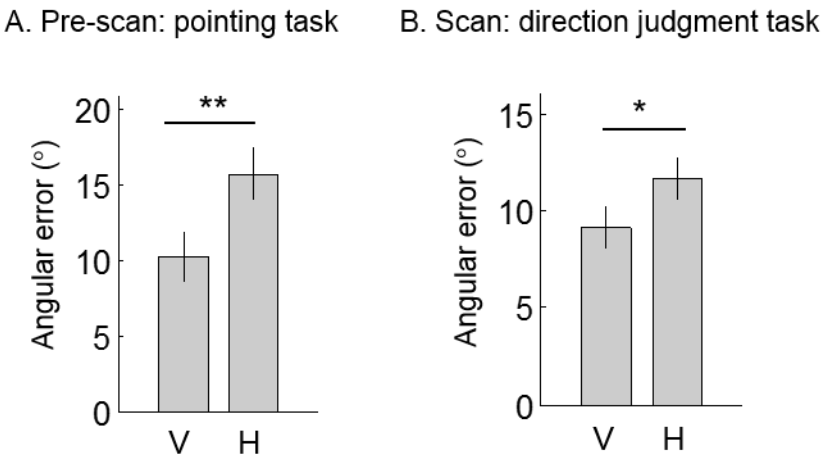


Figure 5.7 Behavioural results. (A) In the pre-scan pointing task, the vertical component of the angular error was significantly smaller than the horizontal component. (B) During the scanning direction judgement task, participants made smaller errors for the vertical question than the horizontal questions. Error bars are SEM adjusted for a within-subjects design (Morey 2008). ** p<0.01, * p<0.05.

5.3.3 Post-scan debriefing session

The rating data showed that participants felt immersed in the virtual environment, with the vast majority choosing either “I felt like I was really in the spaceship” (57% for the pre-scan VR goggle task, 10% for the scanning task) or “I occasionally thought about the environment as being on a computer screen, but overall the environment was convincing and I felt I was moving around in the spaceship” (43% for the pre-scan VR goggle task, 80% for the scanning task). This result implies that my virtual environment effectively conveyed a sense of being in 3D space, and particularly strongly during the pre-scan session with the VR goggles.

In relation to the subjective difficulty of the vertical and horizontal questions during scanning, 50% of the participants reported that the vertical questions were easier while only 10% of the participants found the horizontal questions easier. This subjective report was in line with the objective measures reported above (i.e. the vertical error

was smaller). The rest of the participants (40%) reported that the difficulty was similar for vertical and horizontal questions.

5.3.4 fMRI results – ROIs

I investigated whether vertical and/or horizontal direction information was encoded in the multivoxel patterns in the pre-defined ROIs (that are known to contain head direction cells in 2D). The right retrosplenial cortex showed both vertical and horizontal direction information (vertical, $t(29)=3.69$, $p=0.001$; horizontal, $t(29)=2.05$, $p=0.050$, Figure 5.8A), but this region was significantly more sensitive to vertical direction (paired t-test, $t(29)=2.61$, $p=0.014$). In contrast, the left thalamus showed only horizontal direction encoding ($t(29)=2.81$, $p=0.009$, Figure 5.8B), and horizontal encoding was significantly stronger than vertical encoding (paired t-test, $t(29)=-2.36$, $p=0.025$). The right thalamus and left subiculum also showed horizontal direction information (thalamus, $t(29)=2.27$, $p=0.031$; subiculum, $t(29)=2.63$, $p=0.013$, Figure 5.8C,D), but direct comparison between vertical and horizontal sensitivity was not significant. Bilateral entorhinal cortex, right subiculum and left retrosplenial cortex did not show any significant evidence of vertical or horizontal direction encoding.

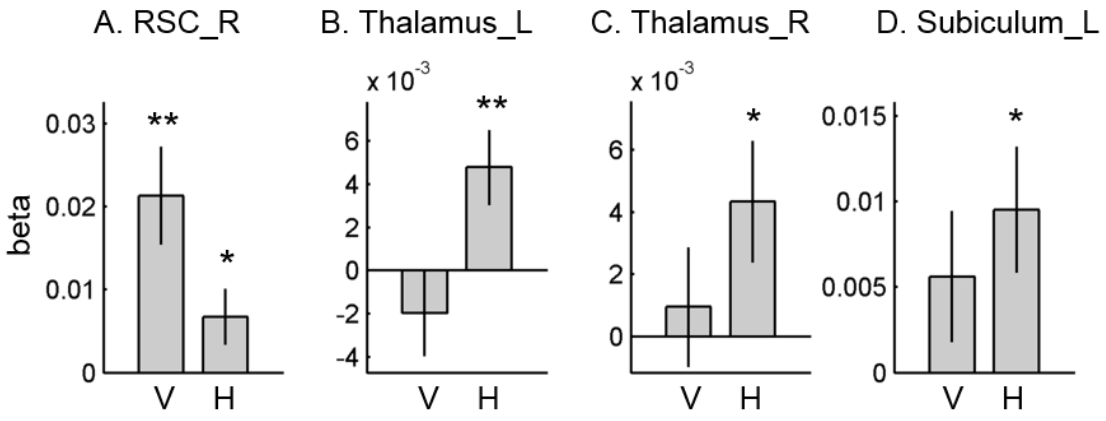


Figure 5.8 Multivoxel pattern analysis in the ROIs. Each ROI is overlaid on the group averaged structural MR image on the top row. (A) Right retrosplenial cortex (RSC) showed both vertical and horizontal direction encoding, but it was more sensitive to the vertical direction. Bilateral thalamus (B, C) and left subiculum (D) showed only horizontal direction encoding. V, vertical; H, horizontal; R, right; L, left. Error bars are SEM. ** $p<0.01$, * $p<0.05$.

5.3.5 Individual differences

The above analysis revealed evidence of vertical and horizontal direction information in the retrosplenial cortex, thalamus and subiculum at the group level. I then tested whether direction information in these regions could explain the individual differences in behavioural performance during the direction judgment test. I found that vertical direction information in the right retrosplenial cortex was significantly correlated with angular error ($r = -0.45$, $n = 30$, $p = 0.01$, Figure 5.9). This means that participants whose right retrosplenial cortex showed more vertical direction information were more accurate at making direction judgments. Horizontal direction information in the right retrosplenial cortex, bilateral thalamus and left subiculum was not correlated with behaviour ($p>0.05$).

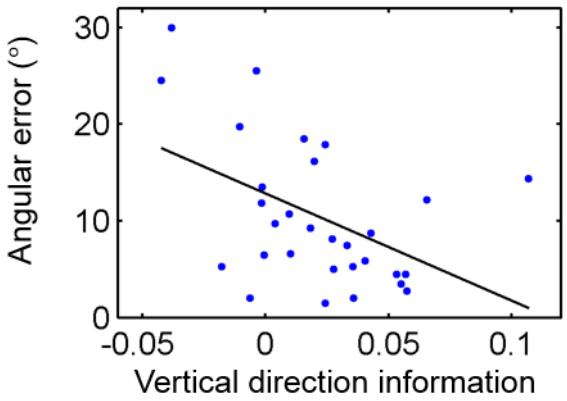


Figure 5.9 Neural correlates of individual differences. Participants whose right retrosplenial cortex exhibited more vertical direction information were better at the direction judgment task (i.e. had a smaller angular error). $r = -0.45$, $n = 30$, $p = 0.01$.

5.3.6 fMRI results – searchlight

A whole-brain searchlight analysis for vertical direction encoding identified bilateral retrosplenial cortex (right, peak at [9, -58, 8], $t(29)=5.62$, $p=0.001$; left, [-9, -46 ,2], $t(29)=5.04$, $p=0.003$, small volume corrected for bilateral retrosplenial masks, Figure 5.10A), similar to the finding from the ROI analysis. Clusters in lingual gyrus (peak, [-

12, -61, 2], t(29)=7.29, p=0.002; [3, -61, 8], t(29)=6.87, p=0.005) and cuneus (peak, [6, -82, 17], t(29)=7.24, p=0.002) also showed vertical direction information.

Horizontal direction information was observed in the anterior part of the left thalamus (peak at [-9, -10, 11], t(29)=4.73, p=0.016, small volume corrected for bilateral thalamus masks, Figure 5.10B). The peak coordinate is most likely located in the ventral anterior nucleus, but I caveat this localisation by noting that the spatial resolution of fMRI scans used here (3 mm) was not fine enough to identify small thalamic nuclei with confidence. Furthermore, neural responses in the neighbouring thalamic nuclei could have contributed to this finding due to the nature of multivoxel pattern analysis (6 mm radius). I also observed a voxel in the left subiculum which showed horizontal direction information, as in the earlier ROI analysis ([-27, -25, -16], t(29)=3.58, p=0.04, small volume corrected for the bilateral subiculum mask). At the whole-brain corrected level, horizontal direction information was also observed in the central sulcus ([-33, -22, 50], t(29)=8.63, p<0.001), supplementary motor cortex ([-6, 5, 53], t(29)=6.10, p=0.04) and visual cortex ([-9, -82, -10], t(29)=6.22, p=0.03; [-9, -79, 5], t(29)=6.04, p=0.047; [-6,-73,-7], t(29)=6.24, p=0.03).

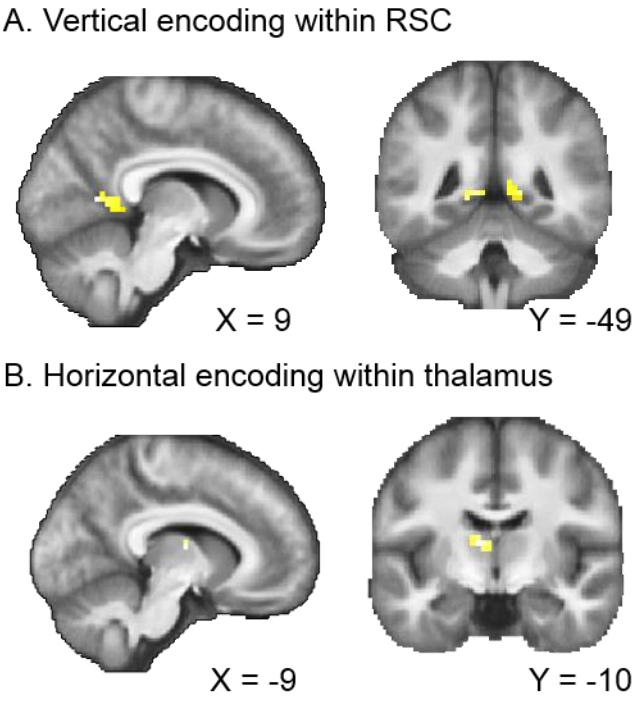


Figure 5.10 Searchlight results. (A) Vertical direction information within the bilateral retrosplenial cortex (RSC) mask. (B) Horizontal direction information within the bilateral thalamus mask, p<0.001 uncorrected for display purposes. See the main text for the other regions that survived whole-brain multiple comparison correction.

5.4 Discussion

In this experiment I investigated how 3D head direction was encoded in the human brain when participants moved in a volumetric space. Using a VR environment and fMRI multivoxel pattern similarity analysis, I found that the thalamus and subiculum were sensitive to the horizontal component of 3D head direction. By contrast, vertical heading information was dominant in retrosplenial cortex, and vertical direction information in retrosplenial cortex was significantly correlated with behavioural performance during a direction judgment task. Participants were more accurate along the vertical dimension for both pre-scan and scan tasks.

The anterior thalamic nuclei (ATN) are important subcortical structures for spatial navigation and memory (Jankowski et al. 2013). Within the hierarchy of the head direction cell network, the ATN receive vestibular inputs via the lateral mammillary

nuclei and project to higher cortical areas including retrosplenial cortex and dorsal presubiculum (Taube 2007). Most head direction cells in the ATN have been recorded when rodents moved on a 2D plane. The current study therefore significantly extends our understanding of the head direction system by providing the first evidence that the thalamus (especially the anterior portion) encodes horizontal heading even when participants moved in a volumetric 3D space.

The lack of vertical direction information in the thalamus resembles the early finding of head direction cells in the lateral mammillary nuclei, which were insensitive to the vertical head tilt of rats (Stackman and Taube 1998), although we should be mindful of the difference in structures (thalamus versus mammillary nuclei) and environments (3D spaceship versus 2D plane), and the limitations of the recording apparatus used in this early rat study. The vertical insensitivity of the thalamus might also be related to previous findings that showed head direction cells in the rat ATN maintained the preferred direction on the vertical wall as if the wall was an extension of the floor, and the head direction cells only cared about the rotation along the body axis, not the rotation of the body axis relative to the vertical gravity axis (Calton and Taube 2005; Taube, Wang, et al. 2013).

Why the head direction system in the thalamus was not sensitive to vertical pitch is an interesting question that requires further investigation. One possible explanation is that the vestibular system, which is responsible for angular integration and updating of the responses of head direction cells in the thalamus, might be less sensitive to vertical rotation because humans are surface-based animals and we infrequently rotate vertically. Although participants' heads were immobilised during scanning, vestibular and proprioceptive inputs they experienced during the pre-scan task with the VR goggles might have been reinstated by visual cues during scanning and contributed to head direction encoding as suggested by a previous study (Shine et al. 2016).

Furthermore, optic flow during scanning alone could have stimulated the vestibular nuclei (Glasauer 2005), and head direction cells in the thalamus of rats were modulated by pure optic flow without visual landmarks (Arleo et al. 2013). Vertical and horizontal optokinetic responses are known to activate both common and unique vestibular nuclei (Bense et al. 2006). It is also possible that vertical information might be more evident in the thalamus if spatial encoding is studied in a real environment instead of a virtual environment. Recently, cells tuned to gravity (vertical tilt) were found in the macaque monkey anterior thalamus (Laurens et al. 2016).

The next finding concerns the subiculum. The presubiculum is reciprocally connected to the anterior thalamus, and a lesion in the thalamus disrupts head direction cells in the presubiculum (Goodridge and Taube 1997). To date, the presubiculum is the only brain structure where head direction cells have been recorded in animals exploring a volumetric space (Finkelstein et al. 2015). In this bat study, cells that were sensitive to either azimuth (horizontal) or pitch (vertical) as well as conjunctive cells were found in the presubiculum. In the present study, I found only horizontal direction information in the human subiculum. This might be attributable to a difference in species (bat versus human) or to methodological differences. Unlike invasive recordings, fMRI measures aggregated neural responses. Therefore, the current study might have captured the dominance of cells tuned to azimuth compared to the cells tuned to pitch that was observed in the bat study (52% and 21% respectively). A future fMRI study with higher spatial resolution might be able to detect less abundant pitch and conjunctive head direction cells in the human presubiculum, if indeed a similar anatomical gradient of azimuth, pitch and conjunctive cells also exists in the human brain (Finkelstein et al. 2015) .

Unlike the thalamus or subiculum, the right retrosplenial cortex exhibited vertical direction information, although horizontal information was also present in this region.

Therefore, in principle it seems that retrosplenial cortex could serve as a 3D compass on its own. My finding of a significant correlation between vertical direction information in the retrosplenial cortex and behavioural accuracy might reflect the functional relevance of retrosplenial cortex for processing 3D direction information (although it is unclear why only vertical direction information and not horizontal direction information in this region correlated with individual differences). The dominance of vertical information in the retrosplenial cortex is concordant with my previous finding of vertical direction encoding when participants moved on a 3D junglegym in Experiment 1. One explanation could be that visual cues might be more salient for the vertical axis compared to the horizontal axis (the influence of view is discussed further below). Within the head direction system, retrosplenial cortex is directly connected to early visual cortex (Kobayashi and Amaral 2003) and head direction cells in retrosplenial cortex are dominated by local visual landmarks (Jacob et al. 2017). Of note, the presubiculum is also known to have direct connections with secondary visual cortex in rodents (Vogt and Miller 1983), but whether there are direct connections between the presubiculum and early visual cortex in primates is unknown.

Behaviourally, participants were more accurate at judging vertical direction (though slightly slower in response) during scanning. Half of the participants reported that the vertical question felt easier and some subjects anecdotally reported that it was due to the views of the ceiling and floor, even though I designed the side walls to also provide clear polarisation cues for the horizontal direction. Views are naturally dependent on head direction, and the horizontal component of head direction has less influence on views as the vertical tilt increases in 3D space. For example, let us assume that the azimuth of a participant is 0° and 180° when the vertical tilt is 0°. In this case, the participant faces towards global East and West and the views can be very different due to distinct landmarks. In contrast, when the vertical tilt is 90°, the participant looks straight up in the sky, regardless of whether they face East or West. The views of the

sky will be rotated 180° in retina space, but this change in view is less dramatic than the view when looking straight East or West. Although I tried to orthogonalise the view and head direction by measuring the neural similarity between pairs of directional responses across different rooms in my virtual environment (as I explained in Section 5.2.6), and I also added a low-level visual texture similarity regressor for extra control, there still remains a possibility that the views were more similar when the vertical tilts were similar compared to when the horizontal direction was similar. This could reflect the nature of the relationship between head direction and view in 3D space, rather than being a particular feature of my virtual environment.

In the pre-scan pointing task where participants were required to look at the remembered location of the ball, I also found smaller vertical errors compared to horizontal errors. There are multiple interpretations for this finding. First, the vertical axis might be inherently more salient than other axes because of gravity providing a reference direction (Barnett-Cowan and Bülthoff 2013). Participants physically tilted their head up and down when wearing the VR goggles in this task. Therefore, it would be interesting to test how much contribution physical gravity made to the encoding and recall of an object in 3D space by conducting the same experiment with a standard flat display and comparing performance between the different types of equipment. Second, as I discussed above, participants might have felt that there were more visual cues for the vertical dimension (ceiling and floor) than the horizontal dimension, even though I designed the environment to contain ample cues along all dimensions. Third, the height of the spaceship was shorter than one of the axis on the floor and this might have facilitated the vertical encoding of the location of the ball. One way to rule out the influences of visual appearance and geometry of an environment would be to conduct the experiment with two environments of which one is a 90° rotated version of the other so that the left-right axis becomes the up-down axis. Fourth, the larger horizontal error might be attributable to poor depth perception. It has been suggested that the depth

axis (which corresponds to one of the horizontal axes when participants are in an upright position, as in the current experiment) is uniquely represented, and not in fact the vertical axis (Orban 2013).

Related to the vertical-horizontal asymmetry, one interesting question is the potential influence of an explicit cognitive task on the neural representation of head direction. In the current experiment, I occasionally asked participants to indicate their vertical or horizontal direction between movements. This task could be answered rapidly and easily, thus minimising interruption to the task and the need for additional scanning time, while ensuring that participants paid attention to their 3D movement direction. However, the explicit and separate questions for vertical and horizontal directions might have contributed to the encoding of vertical and horizontal information in different brain regions. Vertical and horizontal information might be more homogenously represented in these brain regions if participants move freely in 3D space without explicitly paying attention to the vertical and horizontal components of direction. Experimenters could then avoid using the terms “vertical” and “horizontal” during the experiment, and participants could be asked to directly indicate their 3D direction (although I note that it is almost impossible to indicate precisely and rapidly one’s 3D direction without dividing it to vertical and horizontal components). Alternatively, cognitive tasks that test an explicit awareness of movement direction could be removed, given that head direction cells are often recorded in rodents when animals forage in an environment without active navigation or a spatial memory test. In contrast, more spatially demanding tasks, such as 3D path integration with multiple pitch, roll and yaw rotations (Vidal et al. 2004), might result in stronger head direction signals both vertically and horizontally.

In conclusion, this experiment provided evidence showing that the thalamus, subiculum and retrosplenial cortex (the ‘classic’ head direction system) also encode horizontal or vertical heading in 3D space. The entorhinal cortex which was among my ROIs did not

show significant direction encoding. However, I found a different type of spatial representation in the entorhinal cortex, which I describe in the next chapter.

Chapter 6 Experiment 4: A grid code in volumetric space

6.1 Introduction

In the previous chapter, I presented evidence for vertical and horizontal direction encoding in the classical head direction cell network – the thalamus, retrosplenial cortex and subiculum – using an fMRI virtual navigation paradigm in a volumetric space. This 3D head direction information is expected to provide crucial inputs to place cells and grid cells – the main cell types supporting a 3D map of the environment in the brain. The grid code is particularly intriguing because it enables efficient encoding of a large space using relatively few cells, compared to when each individual cell fires at unique locations, as is the case with the hippocampal place code. The need for an efficient representation can be even greater for a volumetric 3D space than a 2D space because the former scales much faster than 2D space (a cubic function versus a square function).

As I described in Section 1.4, theoretical studies have predicted grid cells to follow 3D arrangements such as a face-centred cubic (FCC) or hexagonal close packing (HCP) in a volumetric space (Mathis et al. 2015; Stella and Treves 2015). However, empirical studies are scarce, and knowledge about the activity of grid cells in 3D space remains elusive (except for some preliminary evidence of 3D grid cells reported in bats, Ginosar et al. (2016)). To fill this gap in our understanding of 3D spatial representations, in this experiment, I developed analysis methods to test for a grid code in a volumetric 3D space using human fMRI.

The main question I asked was whether it is possible to detect a grid code using fMRI, given that this technique does not have cellular resolution. Fortunately, previous fMRI studies have successfully investigated grid cells in 2D space using a known property of

grid cells whereby their activity is modulated by movement direction (Doeller et al. 2010). Here I proposed that the same property could be used to probe grid codes in 3D.

In this chapter, I first describe the proposed firing patterns of 3D grid cells, namely FCC and HCP, with graphical user interface (GUI) software I developed to visualise 3D grid structures and direction-modulated grid signals. Next, I explain the principles of detecting 3D grid cells using fMRI. A detailed description of the analysis is then presented together with theoretical and methodological considerations that are either unique to 3D grid codes or relevant to both 2D and 3D contexts. Finally, I tested the feasibility of this new analysis method using the fMRI data obtained in my virtual spaceship experiment which I described in the previous chapter.

6.2 Methods: grid analysis

6.2.1 Expected response profiles of grid cells in a volumetric 3D space

Optimal encoding of 3D volumetric space with grid cells has been extensively discussed in previous theoretical and modelling studies (Mathis et al. 2015; Stella and Treves 2015), and here I briefly describe two potential arrangements of grid fields: FCC and HCP. Both FCC and HCP arrangements are analogous to the spatial arrangement of tightly stacked spheres inside a box with a minimum gap (e.g. oranges in a crate). These 3D arrangements can be viewed as 2D hexagonal lattices (Figure 6.1A left panel) stacked on top of each other with a translational shift between the layers. FCC is composed of three repeating layers (blue, yellow and red spheres, see Figure 6.1B left panel) and HCP is composed of two repeating layers (blue and red spheres, see Figure 6.1C left panel). The grid axis, which is the key property used for my analysis, is the

direction linking one grid field to its neighbouring grid fields (the green lines linking a centre black sphere to neighbouring coloured spheres in Figure 6.1 middle panel).

When considering one centre sphere and its neighbouring 12 spheres, the positions of 3 spheres differ between the FCC and HCP arrangements (the yellow spheres in Figure 6.1B,C middle panel). This difference leads to different direction-modulated fMRI signals between the FCC and HCP, which is described in the next section. I developed interactive web-based software where users can zoom, pan, rotate and cross-sect a 3D grid cell's structure (the software, including a manual, can be accessed here: www.fil.ion.ucl.ac.uk/Maguire/grid3D_gui). Example screenshots are shown in Figure 6.2. The software was implemented using Unity 5.4 (Unity Technologies, CA, United States).

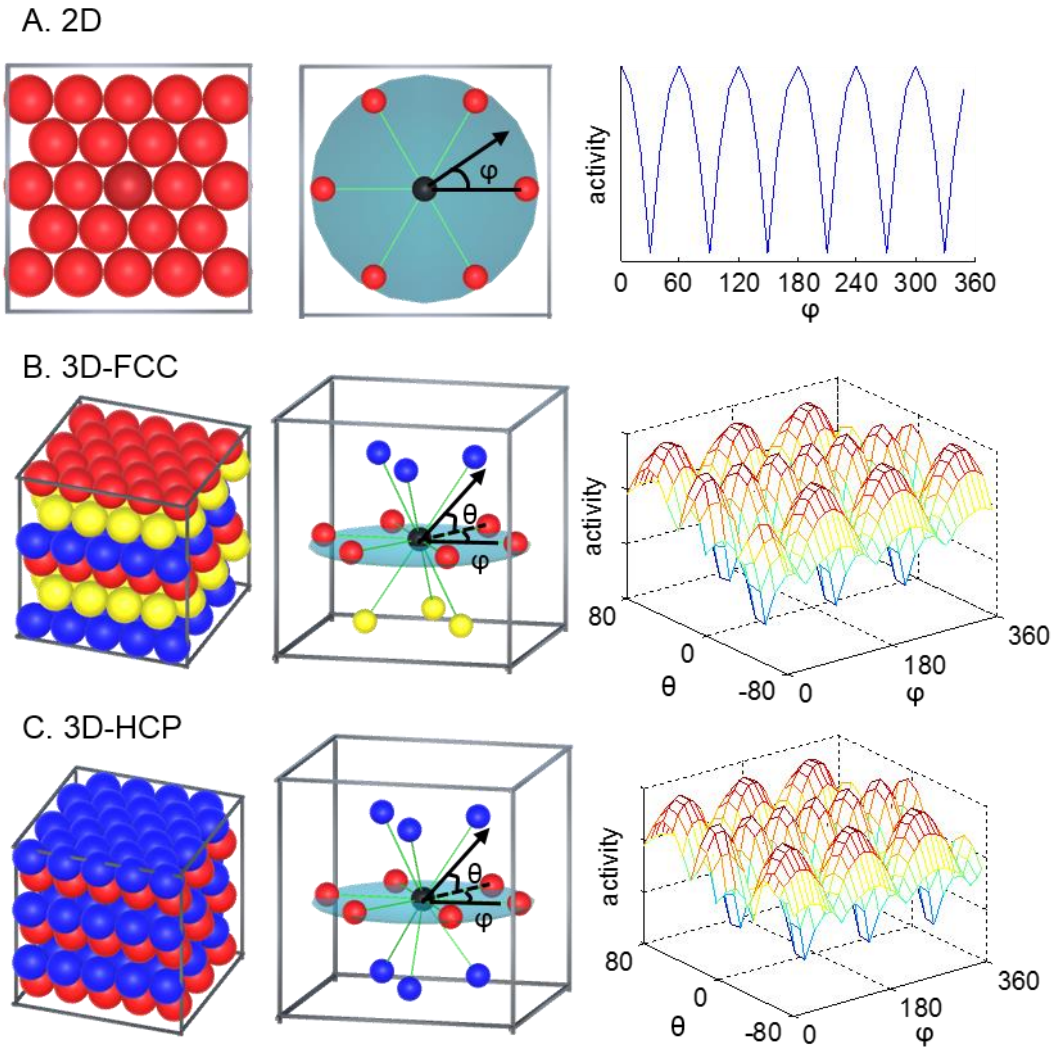


Figure 6.1 Grid cells in 2D and 3D. The left panels show the receptive fields (grid fields) of grid cells in different arrangements – (A) hexagonal in 2D, (B) FCC in 3D, (C) HCP in 3D. The middle panels show the unit cell and grid axis of each arrangement. In 2D, one grid field (black sphere) is surrounded by 6 fields (red balls). In 3D, one grid field (black sphere) is surrounded by 12 grid fields (red, blue and yellow spheres in FCC and red and blue spheres in HCP). The grid axis (green lines) is the direction linking one grid field to its neighbouring grid fields. A grid cell's activity is expected to be modulated by the animal's movement direction (black arrow) relative to the grid axis (the vertical (θ) and horizontal (φ) angles). The right panel shows the simulated grid cell's activity as a function of the movement direction relative to the grid axis. Due to the regularity of the grid axis (e.g. 60° periodicity in 2D), grid activity also shows some periodic patterns.

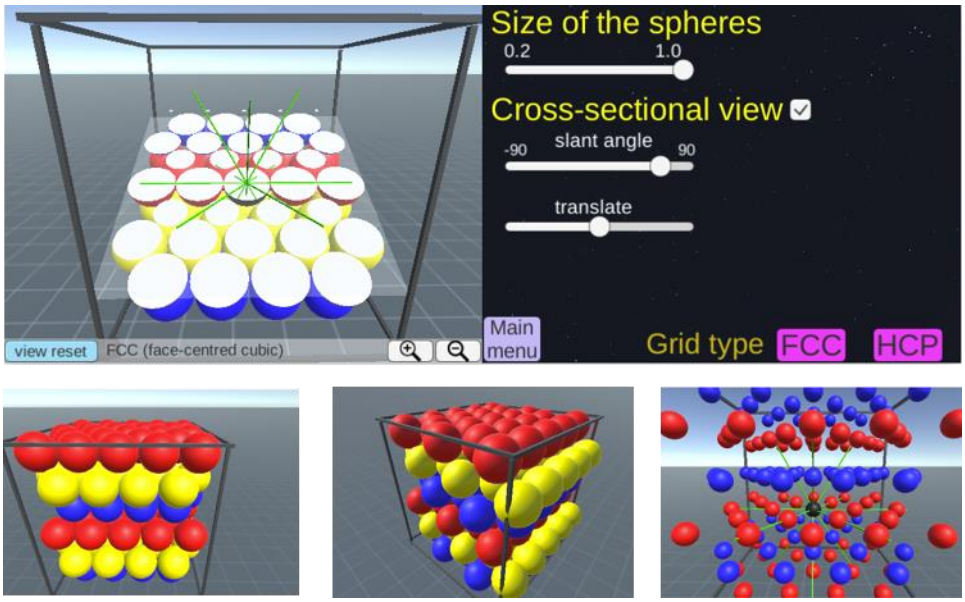


Figure 6.2 Screenshots of my 3D grid cell visualisation software. The software, including a manual, can be accessed here: www.fil.ion.ucl.ac.uk/Maguire/grid3D_gui. Users can change the viewpoint, the size of the spheres and select cross-sectional views of grid fields. Users can also switch between the FCC and HCP arrangements.

6.2.2 Principles for detecting 3D grid cells using fMRI

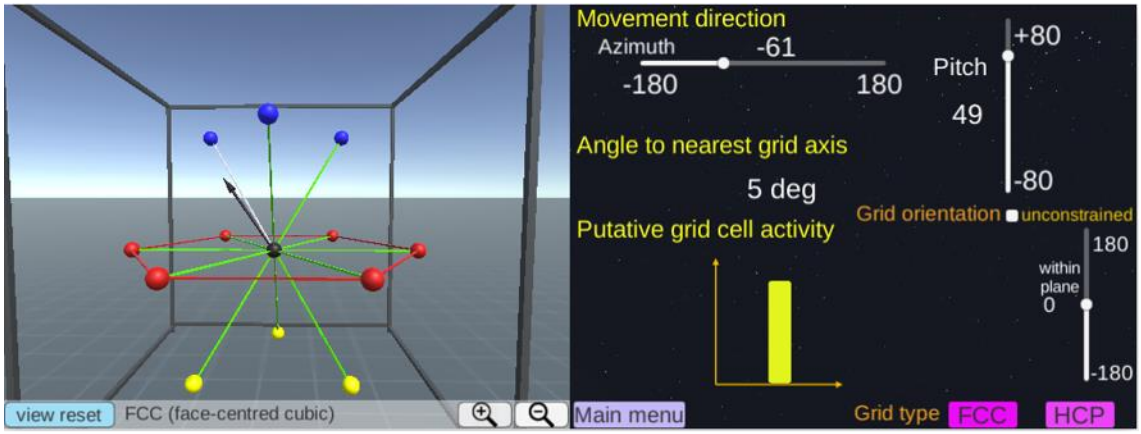
fMRI measures the gross activity of thousands of neurons via complex neural-hemodynamic coupling. When the thousands of grid cells that fire at different locations are summed up, the gross activity is no longer expected to respond to fixed periodic locations in the environment. However, there is another important property of grid cells that enables their detection at a macroscopic level like fMRI. The activity of grid cells is known to be modulated by the alignment between the movement direction of an animal and the grid axis (Doeller et al. 2010). This means that a grid cell shows greater activity when an animal moves along the grid axis (Figure 6.1A middle and right panel). As the majority of grid cells share a common grid axis, the summed response of thousands of grid cells can be systematically modulated by the movement direction of a participant. In previous studies that investigated grid cells in 2D, fMRI activity was modelled as a cosine function of movement direction relative to the grid axis with a period of 60° to

account for hexagonal symmetry (Doeller et al. 2010; Constantinescu et al. 2016; Horner et al. 2016).

I assume that the same principle of direction-modulation will hold in 3D, so that fMRI activity can be modelled as the degree of alignment between 3D movement direction and the grid axis. I defined the alignment score as a cosine of the angle between the 3D movement direction and the nearest grid axis. Thus, a larger signal is expected when the angle is smaller (i.e. the movement is aligned to the grid axis). The nearest grid axis forms the minimum angle with the direction vector. The grid cell's expected response when a participant is moving in a particular 3D direction, defined by azimuth (horizontal angle) and pitch (vertical angle), can be visualised using my interactive software. Figure 6.3A,B show examples of when movement direction is aligned to the grid axis (large signal) and when movement direction is misaligned (small signal), respectively.

According to this model, FCC and HCP grid cells will show complex response patterns as a function of the vertical and horizontal components of a participant's movement direction relative to the grid axis (Figure 6.1B,C right panels). The response patterns of FCC and HCP are largely similar except for the difference in vertical symmetry. As the HCP arrangement is symmetric across the horizontal plane (e.g. the blue spheres on a layer above the red spheres are located at the identical position as the blue spheres on the layer below, Figure 6.1C middle panel), two movement directions that only differ in the sign of vertical pitch (e.g. (azimuth, pitch) = (45°, 30°) and (45°, -30°)) result in the same grid activity (symmetric across $\theta = 0$ plane in Figure 6.1C right panel). In contrast, the FCC arrangement is symmetric across the origin (e.g. the blue and yellow spheres are facing each other, Figure 6.1B middle panel) and therefore the two directions opposing across the origin will result in the same grid response (Figure 6.1B right panel).

A. Movement direction: (azimuth, pitch) = (-61°, 49°)



B. Movement direction: (azimuth, pitch) = (-61°, 24°)

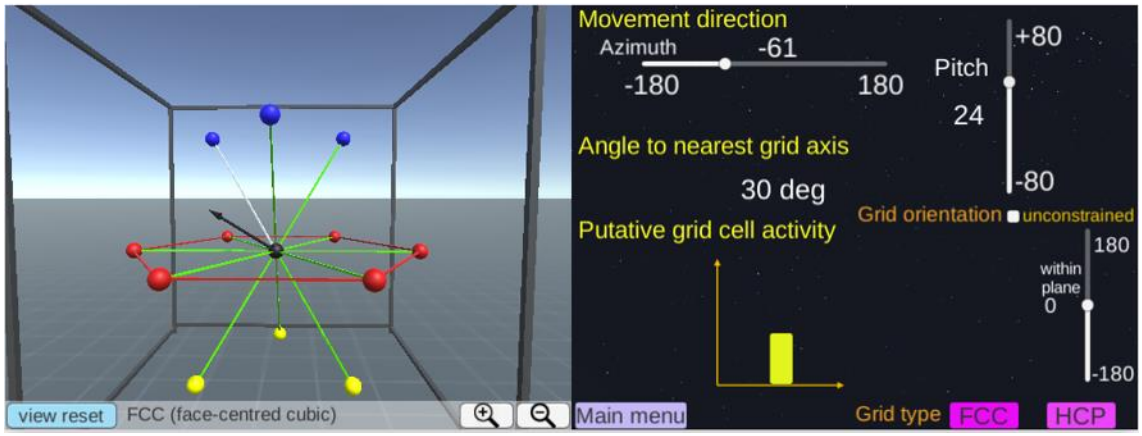


Figure 6.3 A grid cell's activity is modulated by movement direction relative to the grid axis. (A) A participant's 3D movement direction (the black arrow, left panel) is close to the grid axis (the white line, left panel) with 5° deviation. Thus, the activity of grid cells is expected to be high (the yellow bar graph, right panel). (B) A participant's movement direction (the black arrow, left panel) is far away from the grid axis (the white line, left panel) with 30° deviation. Consequently, less activity is expected (the yellow bar graph, right panel). Users can change the movement direction using the sliders on the right panel. The nearest grid axis turns from green to white as users change the movement direction. Users can also switch between the FCC and HCP models and change the viewpoint. See the online manual for a detailed explanation.

6.2.3 The orientation of the grid axis relative to the environment

Crucially, the activity of a grid cell, or a grid voxel in fMRI, depends upon both the movement direction of a participant (which is known to experimenters) and the orientation of a grid axis relative to the 3D environment (which is unknown to experimenters). Figure 6.4 describes two hypothetical cases where a participant moves in the same direction but the grid axis is oriented differently. In Figure 6.4A, a

participant's movement direction (the black arrow) is relatively closely aligned to the grid axis with an angular deviation of 14°, resulting in high activity (the yellow bar graph). In Figure 6.4B, due to different orientation of the grid axis, the same movement direction is further away from the grid axis with an angle of 35°, resulting in low activity. The orientation of the grid axis should be numerically estimated by iteratively fitting the experimental data - a process I describe in a later section.

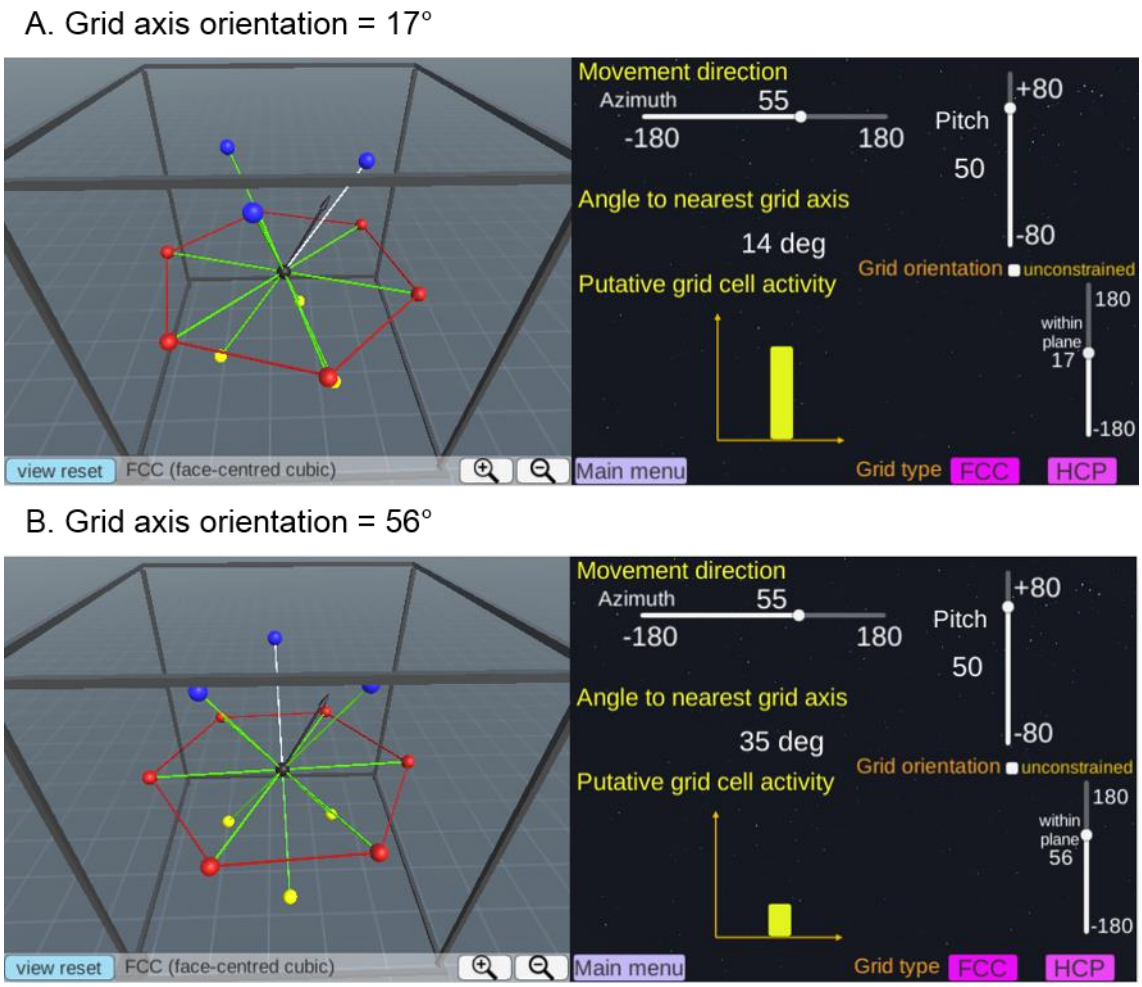


Figure 6.4 The orientation of the grid axis relative to the environment. Movement direction (the black arrow, left panel) is identical in A and B (azimuth = 55°, pitch = 50° in this example). However, grid cells are aligned differently relative to the environment (the rectangular frame, left panels in A and B), meaning that the grid axes are rotated from each other (left panel). Thus, the grid alignment scores measured as the angle between movement direction and the nearest grid axis differ (14° versus 35°), resulting in different amounts of grid activity (the yellow bar graphs, right panels in A and B). Users can change the orientation of the grid axis using the sliders on the right panel.

Unlike in 2D, where the orientation of the grid axis can be specified by one polar angle from a reference direction (e.g. 20° from the north-south axis), the grid axis in 3D can be, in principle, rotated along any three arbitrary axes and the order of applying each rotation also matters (known as the non-commutative property of 3D rotation). Although users can explore these 3D rotation options in my software (Figure 6.5), I restricted the rotation of the 3D grid axis to only one axis so that six hexagonal grid fields (the red spheres in Figure 6.4) remained parallel to the ground of the environment when I analysed my fMRI data. This restriction in grid orientation can be justified by the fact that grid cells on a 2D horizontal surface show the corresponding hexagonal grid fields. This restriction is also required for when researchers want to compare putative grid orientation across multiple voxels or multiple participants using a standard circular statistic like a von Mises distribution. This restriction also simplifies the modelling process and reduces the computational cost and the risk of overfitting GLMs hundreds of times.

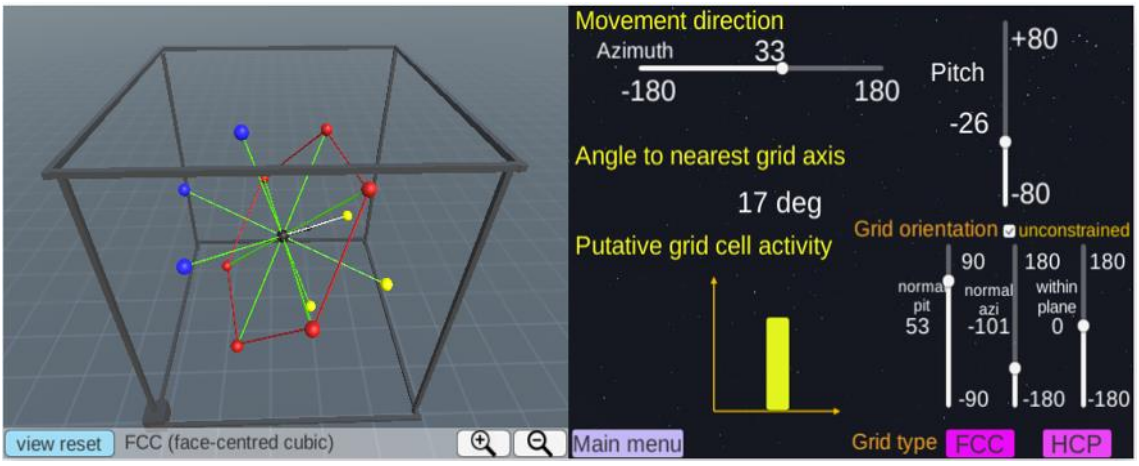


Figure 6.5 When a grid axis is rotated freely along all three axes. A hexagonal grid pattern on the horizontal plane would not be observed in this case.

6.2.4 The relationship between the grid alignment score and fMRI activity

A grid voxel's activity is expected to be modulated by the degree of alignment between movement direction and the grid axis, and I defined the grid alignment score as the cosine of the angle between movement direction and the nearest grid axis. This is similar to the previous grid analysis in 2D which used parametric regressors of cosine and sine functions (Doeller et al. 2010; Horner et al. 2016). However, the precise form of the direction-modulated firing rate of grid cells is not known in either 2D or 3D. There is also additional complexity in measuring the grid cell's signal via neural-hemodynamic coupling. Therefore, it is also possible to model the grid voxel's activity with a non-sinusoidal function, like a linear or binary function. The exact relationship between direction-modulated grid activity and the fMRI response should be examined in future studies.

6.2.5 Estimating 3D grid orientation in fMRI data

In this section, I describe how to estimate 3D grid orientation from an fMRI time series. As explained earlier, the grid alignment score can be calculated as the cosine of the angle between the participant's movement direction (known to experimenters) and the nearest grid axis. The nearest grid axis is determined by the orientation of the grid axis relative to the environment, which is unknown to the experimenters (Figure 6.4). In 2D, the grid activity can be modelled as a simple cosine function of the movement direction (φ) and the orientation of grid axis (ω):

$$Y = \cos(60° \cdot \varphi - \omega)$$

Using the compound angle formula, the orientation (ω) can be estimated analytically by fitting cosine and sine functions in a GLM:

$$\cos(60° \cdot \varphi - \omega) = \cos(60° \cdot \varphi) \cdot \cos(\omega) - \sin(60° \cdot \varphi) \cdot \sin(\omega)$$

However, a simple analytical function of azimuth (φ), pitch (θ) and the grid axis orientation, which can have up to 3 degrees of freedom (3 parameters), describing the complex response pattern of grid activity in 3D is unknown (Figure 6.1B,C right panel). Therefore, I suggest a simple numerical method to estimate grid orientation as follows.

I first assume that grid orientation is aligned at 0° from a reference direction (e.g. parallel to the side wall of the environment) and then calculate the grid alignment score. This grid alignment vector is then convolved with the hemodynamic response function (the SPM canonical hemodynamic response function). The resulting vector serves as a hypothetical grid voxel signal. I create a GLM which contains this predictive 3D grid signal and nuisance regressors that include six head motion realignment parameters and experiment-specific conditions like the occasional question and response periods. The fMRI time series (after standard preprocessing) in each voxel and in each scanning session is then fitted with the GLM, and the outcomes - beta (regression coefficients) and mean square residual - are saved for each voxel.

I then repeat the whole procedure with newly calculated grid alignment scores with different assumptions, namely that the grid orientation is aligned at 15°, 30°, 45°, …, 120° relative to the environment (note it is sufficient to sample the grid orientation up to 120°, as the geometry of the 3D lattice structure of both FCC and HCP is symmetric for 120° rotations on a plane). For each voxel and each scanning session, I select the orientation of the grid axis that gives the best fit by comparing the mean square residual of these multiple GLMs. A GLM with the smallest mean square residual and a positive regression coefficient for the grid signal regressor is selected. The reason I select the GLM with a positive regression coefficient is to avoid the inverted relationship between the hypothetical grid cell's signal and the fMRI response (e.g. when movement is more aligned to the grid axis, the fMRI signal is lower). In rare cases (<10% of voxels in my empirical data – see later sections) where all grid

orientation models yield a negative regression coefficient, I simply select the GLM with the smallest mean square residual. To summarise, this iterative fitting process identifies which grid orientation best describes the fMRI signal in each voxel and in each scanning session.

6.2.6 Testing for a grid signal in the fMRI data

I then test whether each voxel shows a consistent 3D grid signal across different scanning sessions by quantifying the regression coefficient of the grid signal model. For instance, if the fMRI data in one scanning session (e.g. session 1) is best fitted with a grid model that aligns at 15°, I measure the grid score as the beta of the same grid orientation (15°) model in the another scanning session (e.g. session 3). The beta values of voxels in the ROI are averaged for each participant, and a t-test is used to test whether the beta is positive at the group level, as this hypothesis is inherently one-sided. This approach is similar to previous 2D grid analyses where the grid orientation was estimated from one half of the dataset and tested on the other half of the dataset, and the regression coefficient was tested against zero at the group level (e.g. Doeller et al. (2010); a standard group level inference for fMRI experiments).

However, there is a difference between my study and some of the previous studies in terms of grid orientation averaging. In Doeller at al. (2010) and Horner et al. (2016), the estimated grid orientation of each voxel within the ROI was averaged, and this averaged grid orientation model was tested in the other half of the data. Here, I estimate and test the grid orientation model within each voxel, then I later summarise the grid score of voxels within the ROI. Neighbouring grid cells share a common grid orientation which is the essential property of grid cells that allows for the detection of the direction-modulated signal at the fMRI voxel level, and earlier fMRI studies assumed one unique grid orientation for the entire entorhinal cortex. However, there is

also evidence of multiple grid modules within the entorhinal cortex that have different grid orientations and scales (Stensola et al. 2012), and I believe that estimating and testing grid orientation at the voxel level, instead of the whole ROI, can maximise the sensitivity of the analysis. This voxel-by-voxel estimation and test approach was used in more recent 2D grid cell study (Nau et al. 2018).

6.3 Methods: empirical data

In this section, I report on how I tested the feasibility of the 3D grid analysis that I proposed above, by applying the analysis method to empirical data.

6.3.1 Participants, task and fMRI data

The data from Experiment 3 were used in the analysis, given that the question I was addressing here was orthogonal to that considered in that experiment. To reprise briefly, 30 participants explored a virtual zero gravity spaceship while wearing VR goggles prior to scanning. During scanning, they were passively moved along a preprogrammed 3D trajectory inside the virtual spaceship (Figure 5.3). The 3D trajectory allowed tight control and even sampling of each 3D direction and timing. I used a 3D trajectory which sampled a limited range of movement directions, -60° to 60° both vertically or horizontally, to measure a reliable neural signal within a limited scanning time. Importantly, participants were required to keep track of their 3D movements during scanning and, as described in the previous chapter (Section 5.4.2), accuracy for the direction judgment question was well above chance. Participants visited the two compartments of the spaceship alternatively for each of 4 scanning sessions (11 minutes each). Functional images (3 x 3 x 3 mm) were realigned, normalised, and smoothed with a 6mm kernel.

6.3.2 ROI selection

Taking my lead from the extensive animal literature on grid cells, I focused on my analysis on the entorhinal cortex. Left and right entorhinal masks that I described previously were used (Section 3.2.6). Of note, entorhinal cortex has been further divided into posterior medial and anterior lateral parts in one previous fMRI study (Bellmund et al. 2016), based on the finding in rodents that grid cells are typically reported in the medial entorhinal cortex. However, my study used standard resolution fMRI, and further segmentation of this kind was not feasible. Functional specialisation within the entorhinal cortex is an interesting topic that needs to be further addressed in future studies with high-resolution scanning sequences.

Another important point to consider is that entorhinal cortex is notoriously difficult to image because of fMRI susceptibility artefacts in this vicinity. Although sequence development continues in this regard, entorhinal cortex still has inherently low raw BOLD signal compared to other cortical regions. Crucially, standard fMRI analysis softwares like SPM exclude voxels of low signal by default. The “global masking threshold” parameter in the first-level model specification in SPM determines which voxels are to be included in the analysis based on the raw intensity, and voxels in the entorhinal cortex can often be excluded. It can also result in a different number of voxels in the entorhinal ROI for each participant. In my studies, I defined the entorhinal ROI anatomically, without excluding any voxels based on raw signal intensity for several reasons. First, an exclusion criterion based on the mean BOLD intensity is arbitrary. Depending on the version of the software, the same voxels can be included or excluded from the analysis. Second, raw BOLD intensity alone does not predict whether a voxel shows functional modulation. For instance, whereas the raw signal intensity of cerebrospinal fluid is higher than most other cortical areas, we rarely

observe meaningful signal in the cerebrospinal fluid in typical cognitive experimental paradigms.

6.3.3 Main grid analyses

I applied the 3D grid analysis that I described in Section 6.2 to the preprocessed fMRI data. In essence, I estimated the orientation of the 3D grid (separately for FCC and HCP grid types) in each voxel within the entorhinal cortex by iteratively fitting the fMRI time series in each scanning session to the predicted grid alignment score defined as the cosine of movement direction and the nearest grid axis. When I fitted the grid model signals to my fMRI data using GLMs, the experiment-specific condition (turn, question and blank screen periods) were also included as regressors of no interests (see Section 5.2.3 for detailed experimental procedures). The grid model was then tested on data from another scanning session. Because my virtual spaceship had two compartments, I trained and tested the grid cell models within each compartment, and averaged the regression coefficient of the two compartments. This regression coefficient was tested against zero at the group level (excluding outliers - participants with more than a standard deviation of 3 in my empirical data) using a one-sided t-test.

I then tested whether the estimated grid orientation was clustered across participants. If the grid axis was anchored to visual features in the environment such as landmarks or the boundary, every participant would exhibit a similar grid orientation, given they were in the same environment. I calculated the average grid orientation across all voxels within the ROIs for each subject and each compartment using a circular mean. I then applied a Rayleigh test for non-uniformity of circular data. The circular mean and non-uniformity test was computed using CircStat2012a toolbox (Berens 2009).

6.3.4 Control analyses – direction or view encoding

My 3D grid analysis (as well as the 2D grid analyses in the published literature) relied on the dependency of the neural signal on movement direction, and one concern was whether a neural signal that is responsive to one particular direction (or the view associated with a direction) could be weakly correlated with a grid cell model and so identified as a grid cell. This was why I used a direction (or view) encoding model as a control analysis. I created a direction-sensitive model signal which was sensitive to one of nine 3D directions that were visited by participants in the virtual environment. The nine directions were regularly sampled both horizontally and vertically: (azimuth, pitch) = (-60°, -60°), (-60°, 0°), (-60°, 60°), (0°, -60°), (0°, 0°), (0°, 60°), (60°, -60°), (60°, 0°), (60°, 60°).

Following Bellmund et al. (2016), I assumed that each direction-sensitive neural response had a margin of 30°. This meant that neurons or voxels that responded strongly to (0°, 0°) direction would also respond strongly to (±30°, ±30°), and would respond weakly to the rest of the movement directions. I convolved the binary direction response vector with the hemodynamic response function. I created a GLM similar to the grid model described in the previous section but now the grid signal was replaced by the direction encoding signal. Again, the best direction-encoding model was selected for each voxel from one scanning session and then tested on the other scanning session. If voxels in my ROIs (left and right entorhinal cortices) responded to unique directions, I would see a significantly positive regression coefficient for a direction model at the group level.

6.3.5 Control analyses – other grid models

In 2D, a non-hexagonal grid model, such as a 4-fold symmetry, was tested as a control model (Doeller et al. 2010). Similarly, I tested whether fMRI signal in the entorhinal

cortex was explained by a square lattice model (Figure 6.6A). A square lattice model has a lower packing density than the FCC and HCP models, and so the square lattice model is not an optimal way of encoding 3D space. Just as in my testing of the FCC and HCP models, I used the same assumption that the activity of a grid voxel was modulated by the alignment score (cosine of angle) between movement direction and the grid axis orientation (Figure 6.6B,C).

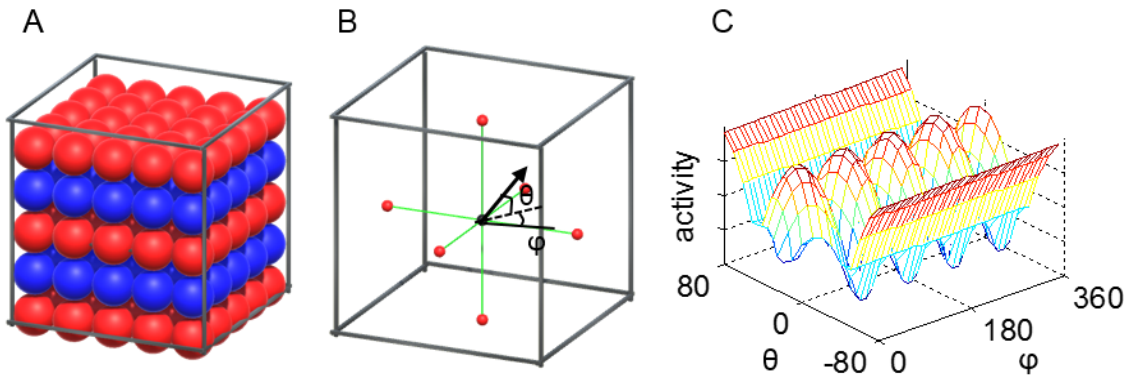


Figure 6.6 A square lattice model in 3D. (A) The receptive fields of hypothetical grid cells which follow a square lattice arrangement. (B) A grid cell's activity is expected to be modulated by the animal's movement direction (the black arrow) relative to the orthogonal grid axis (the green lines); θ, vertical movement angle relative to the grid axis; ϕ, horizontal movement angle relative to the grid axis. (C) A simulated grid cell's activity as a function of movement direction. It displays 90° periodicity when vertical pitch (θ) is close to zero.

I also tested a hexagonal grid model which only cares about horizontal movement direction regardless of pitch. This model is related to the previous observation in rats that receptive field of grid cells were vertically elongated on the wall and in the spiral staircase apparatus (Hayman et al. 2011) (Figure 6.7A). It predicts that grid activity is only modulated by the horizontal angle between movement direction and grid axis (Figure 6.7B,C). For example, grid activity would be high if a participant moves in 0°, 60°, 120°, 180°, 240°, 300° direction (azimuth-wise) independent of whether they move up or down.

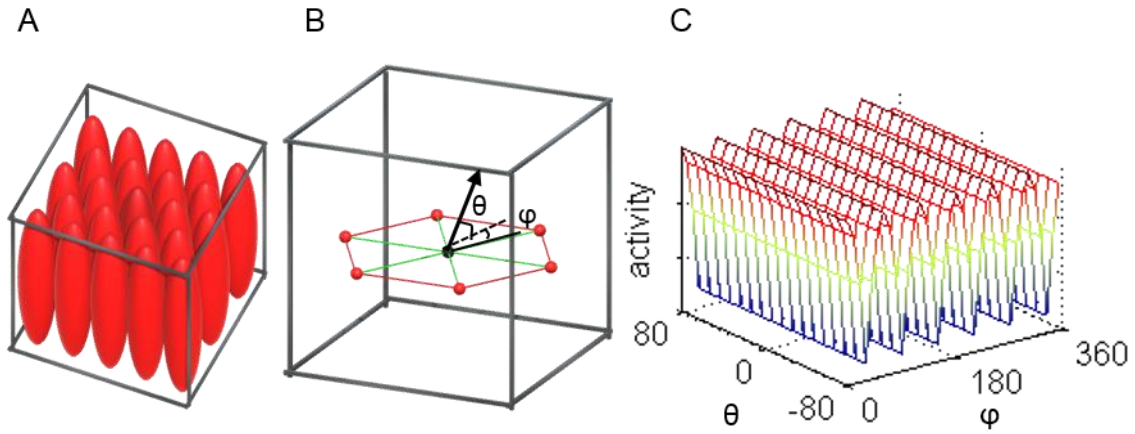


Figure 6.7 An azimuth-only grid model in 3D. (A) The receptive fields of hypothetical grid cells which show hexagonal periodicity only along the horizontal axes and not along the vertical axis. (B) A grid cell's activity is expected to be modulated by the animal's horizontal movement direction (the black arrow) relative to the grid axis (the green lines); θ, vertical movement angle relative to the grid axis; ϕ, horizontal movement angle relative to the grid axis. (C) A simulated grid cell's activity as a function of movement direction. It displays 60° periodicity along the horizontal dimension (ϕ), independent of whether an animal is moving up or down (θ).

6.3.6 Control analyses – other regions

I also tested the main grid models (FCC and HCP) in other brain regions to reassure that the grid-like representation I observed in the entorhinal cortex was not merely an artefact that was present across the whole brain. The control ROIs were primary auditory, visual and motor cortices. I defined the peak coordinate for each ROI using the fMRI meta-analysis toolbox Neurosynth (Yarkoni et al. 2011) (primary auditory cortex [-44, -24, 8]; primary visual cortex [-8, -86, 0]; primary motor cortex [-40, -16, 60]), and then created 7mm spherical ROIs (51 voxels) centred at these peaks. This radius was chosen to match the number of voxels in entorhinal cortex mask (left entorhinal cortex = 47, right entorhinal cortex = 49).

6.4 Results

6.4.1 Main grid analyses

I tested whether fMRI signal in the left and right entorhinal cortices (Figure 6.8A) was modulated by participants' 3D direction, as predicted by the two optimal grid models in 3D - FCC and HCP. The FCC grid model was significant in bilateral entorhinal cortex (left, $t(28)=2.9$, $p=0.003$; right, $t(28)=1.8$, $p=0.04$, one-sided, Figure 6.8B). The HCP grid model did not significantly explain the response of either left or right entorhinal cortex (left, $t(28)=1.4$, $p=0.08$; right, $t(29)=0.4$, $p=0.4$, Figure 6.8C).

I then examined whether the putative FCC grid orientation was clustered across participants using Rayleigh's test for non-uniformity. Figure 6.8D shows that grid orientation was not significantly clustered in the left entorhinal cortex (room A, $Z = 2.2$, $p = 0.1$; room B, $Z = 1.1$, $p = 0.3$), or the right entorhinal cortex (room A, $Z = 1.1$, $p = 0.1$; room B, $Z = 0.2$, $p = 0.8$). This means that the grid axis was not anchored to, or driven by, particular features of the environment, consistent with previous studies in circular environments, even though my environment was rectangular (Doeller et al. 2010; Nau et al. 2018).

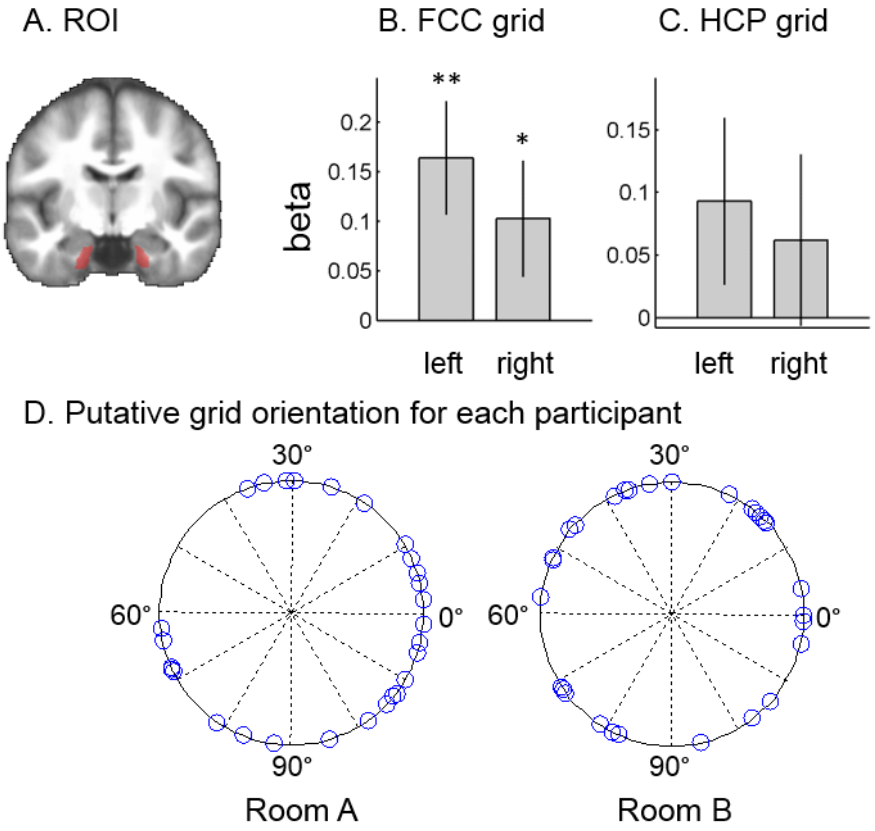


Figure 6.8 Grid cell-like representations in the entorhinal cortex. (A) The anatomical ROIs, left and right entorhinal cortices, are shown on the group-averaged structural MRI scan. (B) The mean beta of the FCC grid cell model was significantly positive in the entorhinal cortex. (C) The HCP model was not significant in either left or right entorhinal cortex. (D) Putative FCC grid orientation was not clustered across participants. Each blue circle represents the grid orientation of individual participants in the left entorhinal cortex. ** p<0.01, * p<0.05

6.4.2 Control analyses

To exclude the possibility of a neural signal sensitive to one particular direction (or associated view) being identified as a grid voxel, I tested a unique direction encoding model as a control. The direction encoding model was not significant in either entorhinal cortex (left, t(29)=-0.8, p=0.8; right, t(29)=-1.1, p=0.9), suggesting that the FCC grid-like signal found in the entorhinal cortex was not driven by one particular direction.

I also tested a square lattice model and a hexagonal grid model, where vertical pitch was ignored. Neither of these models significantly explained the fMRI signal in the

entorhinal cortex (square lattice model: left, t(29)=-0.9, p=0.8; right, t(29)=-0.3, p=0.6; azimuth-only model: left, t(29)=1.6, p=0.06; right, t(29)=0.7, p=0.2).

Finally, I examined the FCC grid model in the size-matched primary sensorimotor areas, and no region showed a significant grid code (primary auditory cortex, t(29)=0.5, p=0.3; primary visual cortex, t(29)=1.7, p=0.053; primary motor cortex, t(29)=-2.3, p=1.0). This further suggests that my finding of an FCC grid-like signal in the entorhinal cortex was not a spurious effect that was detectable anywhere in the brain.

6.5 Discussion

In this chapter, I presented a novel analysis method to investigate grid cells in 3D non-invasively in humans. I also developed associated software to help researchers visualise grid cells in 3D and predict their responses. Using these methods, I observed an fMRI signal in the entorhinal cortex which was in line with one particular 3D grid cell model – an FCC lattice model.

The main advance of this study was the probing of putative grid cells in 3D by predicting the fMRI signal as a function of 3D movement direction and the grid axis. The principle of measuring direction-modulated grid signals has been widely used in 2D (Doeller et al. 2010; Kunz et al. 2015; Bellmund et al. 2016; Constantinescu et al. 2016). Here, I extended, for the first time, this principle into 3D volumetric space, thereby opening up the possibility of empirically studying grid cells in high-dimensional space. Predicting the activity of a 3D grid cell was inevitably more complicated than a 2D grid cell due to the added dimension, as I described in the Methods section. Consequently, I estimated the 3D grid orientation by iteratively fitting the neural data with signals predicted by different possible grid orientations. I successfully demonstrated the feasibility of this analysis approach by finding that my data were

concordant specifically with the FCC model in the entorhinal cortex, the candidate brain structure for 3D grid encoding.

This result is suggestive of 3D grid cells in the human entorhinal cortex that are responsive when people explore a volumetric space. However, I remain cautious until this finding is further corroborated by future studies that directly measure cellular responses. It is not yet proven that the direction modulation principle also holds in 3D, because there is as yet no clear evidence of grid cells showing a regular 3D lattice pattern (note that a recent study in bats found multiple firing fields that were not perfectly regular, Ginosar et al. (2016)). Invasive recordings from the human brain in a similar 3D virtual reality experiment to the one I used here may be able to provide more direct evidence of grid cells showing 3D receptive fields. Once there is a fuller understanding of the cellular physiology of grid cells, it will be possible to determine the optimal fMRI analysis protocol that takes into consideration the multiple factors I have described here. This includes whether the orientation of the 3D grid axis is parallel to the ground or not, the precise model between the grid alignment and the fMRI signal (e.g. cosine, linear, binary) and the distribution of the grid orientation across different voxels within the entorhinal cortex.

The distinction between the FCC and HCP models also deserves further investigation. Although I showed that only the FCC model significantly explained my empirical data, it does not provide unequivocal evidence of the superiority of one model over the other, because it does not involve direct model comparison. Neither the regression coefficient nor the residuals between the two models were significantly different (data not shown). To have a formal comparison between the models, I believe that more data and more sampling directions (beyond the -60° to 60° used in the current experiment) are needed in future experiments. The FCC and HCP models have identical encoding efficiency (Mathis et al., 2015). Stella and Treves (2015) found that simulated grid fields did not

converge to either FCC or HCP alone, and the outcome had a small deviation from both models. It may be that factors such as the anisotropy of the environment (all animals are influenced by gravity, thus the vertical axis is distinguished from the other two axes) might influence the optimal encoding of space. It is currently unknown how this might be related to either the FCC or HCP models. Furthermore, grid cells are known to be influenced by the geometry of environment (Krupic et al. 2014). For instance, a previous study showed that grid orientation was anchored to the boundary of the environment (Stensola et al. 2015; Julian et al. 2018) although such clustered orientation was not observed in my virtual spaceship with a rectangular shape. FCC and HCP models should be tested in environments of different shapes (e.g. spherical or asymmetric shapes) in the future.

In the process of extending grid analysis from 2D to 3D, I felt there was a pressing need for 3D visualisation software, because it was not trivial to imagine the lattice structures of 3D grid cells and the relationship between the grid axis and 3D movement vectors. Therefore, I developed simple interactive software that can help researchers to visualise and understand the direction-modulated signal of 3D grid cells. This web-based software is easy to use without requiring an additional third-party CAD (computer-aided design) program. Of note, this software was developed for visualisation purposes, and it does not have a data analysis function. There is already grid analysis software for 2D grid codes (Stangl et al. 2017), and researchers who are familiar with 2D grid analysis can readily implement a 3D analysis with their own protocols, once they grasp the geometry of 3D grid cells.

In summary, I believe that my experimental paradigm, analysis method and software serve as a useful initial stepping-stone for studying grid cells in realistic 3D worlds. Grid cells have been reported to encode not only physical space, but also more abstract knowledge (Constantinescu et al. 2016; Aronov et al. 2017; Julian et al. 2018; Nau et

al. 2018), and I hope my approach might, in due course, also promote interrogation of more abstract high-dimensional cognitive processes.

Chapter 7 General discussion

7.1 Overview

The aim of this thesis was to extend our current understanding of the neural representation of space into 3D about which little was known, particularly in humans. To achieve this goal, I conducted a series of behavioural and fMRI experiments using various types of custom-built 3D virtual environments. Before I discuss the overarching findings of this endeavour in relation to 3D spatial encoding hypotheses, and before proposing future directions for this research, I will first recapitulate the main findings from my four experiments.

In Experiment 1, I examined whether vertical and horizontal spatial information (place and direction) were equally well encoded in a semi-volumetric lattice environment where participants could move on straight, tilted-up or tilted-down pathways. Behaviourally, participants had similarly accurate memory for vertical and horizontal locations (except for a small response time difference). Concordant with the behavioural findings, fMRI multivoxel pattern analysis revealed that the right anterior hippocampus contained place information that was similarly sensitive to both horizontal and vertical axes, supporting an isotropic 3D place encoding hypothesis. In contrast, participants indicated their heading direction more quickly and more accurately when they were in a tilted direction, and they also overestimated the vertical slope. Retrosplenial cortex and posterior hippocampus were the brain areas that I found to be sensitive to vertical direction.

In Experiment 2, I assessed how a multi-compartment space (a building with multiple levels and rooms) was represented in the brain using spatial memory tests and fMRI repetition suppression analyses. Behaviourally, participants were faster at within-room

egocentric spatial judgments and showed priming effects of visiting the same room, suggesting a compartmentalised representation of space. There was no evidence of additional compartmentalisation into either vertical columns or horizontal rows. At the neural level, the anterior hippocampus showed local information within a room, whereas retrosplenial cortex, parahippocampal cortex and posterior hippocampus showed room information within the wider building, supporting a hierarchical map of this 3D space. Again, vertically adjacent rooms and horizontally adjacent rooms were similarly distinguishable.

In Experiment 3, I tested how 3D direction information was encoded in a volumetric space where participants could move freely along all 3 axes. Behaviourally, participants were more accurate at indicating their vertical direction than their horizontal direction. Using fMRI multivoxel pattern analysis, I found that the thalamus, particularly the anterior portion, and the subiculum encoded the horizontal component of 3D direction, whereas the retrosplenial cortex predominantly encoded the vertical direction.

In Experiment 4, I developed an fMRI analysis method to study, non-invasively, the response of putative human grid cells in a volumetric 3D space. I found that fMRI signal in the entorhinal cortex was best explained by a FCC grid model which has previously been suggested as the optimal encoding model of 3D space.

In the following sections, I discuss my findings in relation to different 3D spatial encoding hypotheses.

7.2 A generalised spatial code

One extreme hypothesis predicts that a 3D spatial code is a simple extension of a 2D code. In other words, the brain uses an elegant spatial encoding scheme that can be

generalised across multiple dimensions, as if the periodic finding of grid cell in 1D environments is a slice through a 2D lattice (Yoon et al. 2016). Consequently, a 3D map should be isotropic (e.g. all 3 axes are equal). For example, place cells, which have circular receptive fields in 2D, are expected to show spherical receptive fields in 3D instead of ellipsoidal or irregular shapes. Place cell recorded in flying bats indeed showed isotropic 3D receptive fields (Yartsev and Ulanovsky 2013). Head direction cells, which are tuned to a specific azimuth in 2D (i.e. a 2D vector), are expected to be tuned to specific combinations of azimuth and pitch in 3D (i.e. a 3D vector). Such conjunctive head direction cells were observed in the bat presubiculum (Finkelstein et al. 2015). Grid cells that show hexagonal lattice firing patterns on a 2D plane are expected to show 3D lattice firing patterns in 3D (Horiuchi and Moss 2015; Mathis et al. 2015; Stella and Treves 2015). At the behavioural level, animals are expected to perceive and remember their location equally well along all 3 axes and freely explore the space without a preference for movement in one particular dimension.

My findings in a 3D lattice environment (Experiment 1) – of similarly accurate memory for vertical and horizontal locations and 3D place information detected in the human anterior hippocampus – supports the generalised, isotropic encoding hypothesis (note, I will discuss the limitations of fMRI spatial resolution in Section 7.6.3). In the multi-level building environment (Experiment 2), I again found that two rooms on top of each other were as distinguishable as two side-by-side rooms in terms of representations in the posterior hippocampus, retrosplenial cortex and parahippocampal cortex. The finding of a 3D FCC grid-like representation in the entorhinal cortex while participants moved in a virtual spaceship (Experiment 4), also supports the notion that the brain uses an efficient representation of space not only in 2D but also in 3D.

However, other findings raise the question about whether humans have a perfectly isotropic map of 3D space and whether every spatial code is fundamentally 3D,

particularly the direction code. In Experiment 1, the presence of vertical distractors slowed down responses even though it did not affect accuracy. If participants were perfectly confident about their 3D location without vertical-horizontal biases, both accuracy and response time should have been comparable. Multiple previous behavioural studies in multi-level buildings have found navigational strategy and performance differences that favoured either vertical or horizontal segmentation (see Section 1.6.2).

Of particular note, vertical and horizontal direction encoding did not take place in the same brain structure with the same precision. In Experiment 1, retrosplenial cortex and posterior hippocampus contained vertical direction information only. In Experiment 3, the retrosplenial cortex again showed predominantly vertical direction encoding whereas the thalamus and subiculum showed horizontal direction encoding.

These findings beg the question as to whether there are more realistic alternatives to a generalised spatial encoding hypothesis that could explain my findings.

7.3 Separate neural substrates for vertical and horizontal direction encoding

Head direction cells that are tuned to a specific 3D direction might not be necessary for a 3D spatial map. Instead, groups of head direction cells (in the same or different brain regions) that are tuned to either azimuth or pitch could be sufficient. Even in bats who naturally explore a volumetric 3D space, the majority of head direction cells were tuned to either azimuth or pitch alone, rather than 3D direction (Finkelstein et al. 2015). If vertical and horizontal angular velocity signals are fed into separate sets of head direction cells, some head direction cells would respond to pitch and the others would respond to azimuth alone. Cells that were sensitive to angular head velocity on a horizontal plane were found in the dorsal tegmental nucleus, lateral mammillary

nucleus and other vestibular nuclei (see Taube 2007 for review). Cells that are sensitive to vertical angular head velocity have not yet been studied, but a separate population of cells encoding vertical angular velocity is likely to exist because the vestibular system in the inner ear also consists of multiple organs that are sensitive to either vertical or horizontal rotation (see Section 1.3). Consequently, the separate streams of vertical and horizontal direction information might be integrated at the level of hippocampal place cells or entorhinal grid cells. In principle, attractor neural network models that explain linear path integration for 2D place cells (McNaughton et al. 2006) can be extended to 3D by adding separate intermediate layers that contain either vertical or horizontal direction signals, although wiring of this recurrent network could be challenging (Horiuchi and Moss 2015). Separate encoding of pitch and azimuth can save on neuronal resources, because if each cell is tuned to every possible combination of azimuth and pitch, the number of combinations would be enormous.

Regarding neural substrates of head direction encoding, the influence of landmarks and views should be also considered. Angular path integration is prone to error accumulation and head direction cells should be anchored to landmarks for reliable direction encoding. As I discussed earlier (Section 5.4), views can be more dependent on vertical head direction in the natural environment, as in Experiments 1 and 3. Thus, a head direction system that is more influenced by visual cues, like the retrosplenial cortex due to its close anatomical connections with visual cortex (Kobayashi and Amaral 2003; Jacob et al. 2017), might show higher sensitivity to vertical heading differences. Future studies could examine vertical and horizontal head direction information that is independent of views by using less naturalistic setups where view and head direction are orthogonalised, e.g. static images or imagination without visual cues (Vass and Epstein 2013; Marchette et al. 2014; Sulpizio et al. 2014; Shine et al. 2016).

7.4 Vertical-horizontal asymmetry and its implications for a 3D map

In the previous section, I presented the possibility of separate neural substrates for vertical and horizontal direction information. I now discuss a different issue which is vertical-horizontal symmetry. Vertical and horizontal direction encoding can be equally or unequally precise and this might affect the isotropy of a 3D map.

As I reviewed in Section 1.6, numerous behavioural studies have suggested that humans, rodents and birds are sensitive to their body orientation relative to the gravity axis (the vertical axis) and use the slope as a navigational cue (e.g. Grobéty and Schenk 1992; Steck et al. 2003; Nardi and Bingman 2009). For example, human participants were more accurate on a navigation task in the sloped condition even when the slope was only 4° (Steck et al. 2003). The energy cost associated with vertical movement has been proposed as an explanation for participants' tendency to overestimate slopes (e.g. perceiving a 30° hill as 50°) (Proffitt et al. 1995; Creem-Regehr et al. 2004; Shaffer and Flint 2011).

The questions then arise as to whether, and how, people could have an accurate map of 3D space if their perception of 3D geometry is distorted due to overestimation of vertical angles or height (Brandt et al. 2015). To consider these issues, we should first distinguish the explicit (verbal or visual) estimation of a slope on the one hand and knowledge about one's actual 3D orientation on the other. When people estimate a slope, they typically stand upright on a horizontal plane at a distance from the slope. Incorrect estimation of a slope does not prevent people from correctly behaving in 3D environments, e.g. people walk normally on a 10° hill without wrongly adjusting their ankles for a 20° hill. On a related note, Proffitt et al. (1995) showed that participants made fairly accurate judgments of a slope when they were asked to adjust a tilt board to match the slope of a hill using their hands. In my Experiment 1, although participants

incorrectly overestimated the angle of the slope, the slope actually helped them to know which direction they were heading. In Experiment 3, participants successfully kept track of their 3D direction, and the angular errors were actually smaller for the vertical direction questions. The vertical angular error was also smaller in the pre-scan pointing task in the same experiment (although there could be several explanations for this behaviour, including the asymmetric visual cues, see Section 5.4). To summarise, at least in my experiments, participants had good knowledge about their 3D orientation, particularly the pitch (but see Vidal et al. (2004) and Gramann et al. (2012) who found poor path integration performance for rotation on a vertical plane).

If we calculate our location purely by integrating velocity (direction and speed) over time, behavioural and neural differences in vertical and horizontal direction encoding would be followed by similar differences in location encoding. In other words, more accurate encoding of vertical movement direction would lead to more precise vertical location estimation. However, findings of direction and location encoding were not in parallel in my Experiment 1. Although, behavioural and neural results implied the vertical direction was more salient, participants were equally good at knowing their vertical and horizontal location (in terms of accuracy) and indeed seemed less confident about their vertical location (slower response time), while fMRI signal in the anterior hippocampus was in line with an isotropic place encoding hypothesis. There are at least two explanations for this discrepancy between the direction and location encoding.

First, the 3D map might not be perfectly isotropic. As I discussed earlier (Section 3.4, 4.4), there might have been a small difference in the precision of place encoding that was not detectable in the discretized environment. For example, if 0.8 vertical unit difference was equivalent to 1.0 horizontal unit difference, this small difference was unlikely to be detected in the discretized environment. This explanation also applies to

the finding in the multi-level building in Experiment 2 which had only two rooms and two floors. The best experimental setup to investigate the relationship between direction and place encoding is a continuous and volumetric 3D environment like the one used in Experiment 3. However, in Experiment 3, my priority was to understand the 3D direction code and I did not include behavioural tasks to measure the accuracy or precision of place encoding.

Second, path integration from the head direction system is not the only source of knowing one's location. Unless participants navigate blindfolded in a sensory-deprived environment (zero auditory, tactile, odour cues), visual landmarks play an important role. In all of my experiments, I included several visual cues to aid spatial cognition in the virtual 3D environments (e.g. a green door in Experiment 1, a floor sign in Experiment 2, ceilings, floors, windows in Experiment 3). Therefore, even if direction information is asymmetric for the vertical and horizontal dimensions, additional sensory cues might compensate for this and enable an isotropic place map.

7.5 Flexible representations of 3D space

Throughout my thesis, one question was how 3D spatial information is represented in the human brain, and the isotropy of a 3D map was one aspect of that question. Instead of drawing a simple conclusion about the mechanism of 3D spatial encoding, such as "the brain uses an isotropic 3D volumetric map", "a 3D map is fundamentally anisotropic due to gravity", I want to emphasise that 3D spatial encoding is a generic problem where multiple factors should be considered. I believe it is more likely that we use a flexible representation of 3D space according to the nature of environment and our behaviour within it, rather than using a fixed, stereotyped representation.

When we navigate in a volumetric space like the virtual zero gravity spaceship, a neural compass that can track continuously changing 3D directions (Experiment 3) and a grid cell system that can efficiently represent a 3D space, like the FCC lattice (Experiment 4), would be appropriate. In contrast, 3D head direction cells or grid cells would be mostly redundant for navigation in a multi-level building where we mainly move on a horizontal plane and only occasionally move across planes. In the multi-compartments environment, a hierarchical representation would be optimal and I found behavioural and neural evidence that exactly supports that in Experiment 2 (e.g. a priming effect of visiting the same room, dissociation of local and room information in the brain). In the 3D lattice environment where movement was constrained to narrow tracks (Experiment 1), a continuous tracking of 360° azimuth and 180° pitch might not be necessary.

Whether the map is isotropic or not will also depend on the behavioural demands and the shape of environment. In all of my experiments, participants were explicitly asked to know their precise 3D location or direction. The behavioural and neural results implied that both vertical and horizontal information were well represented in the brain, although different neural structures were involved in this information processing. I would predict that vertical-horizontal asymmetry might be more prominent if an explicit task was absent. For example, people normally stop thinking about which floor they are on once they arrive at that floor. If a place code is measured in this situation, vertical information would appear to be missing or reduced compared to horizontal location information, as in place cells and grid cells recorded in rodents that showed reduced information for the vertical axis (Hayman et al. 2011). In contrast, if participants receive an incentive for knowing their vertical location, or if there are more salient landmarks in the vertical dimension, then I would expect the neural representation to be sharpened for the vertical axis. Reward and attention are known to modulate the activity of place cells (Markus et al. 1995; Hölscher et al. 2003).

7.6 Outstanding issues and future directions

In this thesis, I explored multiple questions about how we represent 3D space, which have been rarely investigated to date, but this thesis is only a starting point. There are limitations to my experiments, and numerous other issues that still need to be addressed, as I will discuss in this section.

7.6.1 From virtual reality to real world navigation

“Is navigation in virtual reality with fMRI really navigation?” (Taube, Valerio, et al. 2013) is a common criticism applied to fMRI studies of spatial cognition in both 2D and 3D. Active movements, vestibular and proprioceptive inputs are absent when a participant’s head is immobilised inside a narrow MRI scanner bore. However, numerous behavioural and neural studies have suggested that active movement is important for navigation and reliable neural representations of space. For example, even for a simple scene perception task, participants were better when they actively changed their viewpoint compared to when they remained static and the scene (some real objects on a table) was rotated (Wang and Simons 1999). In Klatzky et al. (1998), participants correctly updated their heading and pointed to the origin in a blindfolded walking condition, but not in the passive optic flow (random lines) condition (although note that in a naturalistic virtual environment, the vast majority of participants correctly updated their heading using visual motion alone (Riecke et al. 2012)). At the neural level, place cells showed less spatial-specificity (i.e. large place fields, low peak firing rates, noisier firing patterns) when rats were passively moved or remained static (Terrazas et al. 2005). Spatial selectivity of place cells was dramatically reduced in body-fixed rats in a 2D virtual arena (Aghajan et al. 2014). On the other hand, in another VR setup where rats could walk freely and rotate, place cells, head direction cells and grid cells were all

observed, although the grid fields were expanded compared to real world navigation (Aronov and Tank 2014). Passive transport is known to impair path integration in head direction cells (Stackman et al. 2003).

The abovementioned findings suggest that head-restrained virtual navigation is not the optimal method to study the neural correlate of spatial cognition, but virtual navigation in fMRI is the main workhorse when studying spatial cognition in the human brain (including this thesis) for following reasons. First, fMRI (along with MEG) is the only non-invasive method that can reliably measure neural activity in deep brain structures like the hippocampus and thalamus, and the head must be immobilised for fMRI. Electroencephalography (EEG) and functional near infrared spectroscopy (fNIRS) are non-invasive and portable imaging methods, allowing measurement of freely moving human participants, but these methods do not have good spatial resolution and coverage (i.e. electric resistivity of skull and non-transparent tissue limit the imaging depth), and EEG is particularly susceptible to muscle artifacts. Second, virtual navigation paradigms provide the means for both rigorous and flexible experimental control over what participants see and experience, which is much more difficult to achieve in the real world. For example, the bespoke 3D environments I created for the experiments in this thesis were created with great care for every detail (colour, texture, size, etc.), and one of the environments - the virtual zero gravity spaceship - could not be easily realised in the real world.

However, in the future, we might be able to mitigate the current technical limitations of studying spatial navigation in the human brain with a recently developed mobile MEG system. MEG measures small magnetic fields induced by neural activity (i.e. it has excellent temporal resolution compared to fMRI which measures slower hemodynamic responses) and homogenous magnetic permeability of the brain enables clear source localisation (i.e. reasonable spatial resolution) (Baillet 2017). Although it has long been

thought that MEG cannot measure deep structures due to signal attenuation, this notion has been challenged by recent advances in precise modelling and accurate sensor positioning techniques (Meyer et al. 2017). Most importantly, unlike the traditional MEG scanner which is as bulky and restrictive as MRI scanners due to the need for cryogenically cooled magnetic sensors, the newly developed MEG system with optically-pumped magnetometers (OPMs) is lightweight and portable. A recent study showed that neural activity can be measured when participants made natural movements such as drinking, stretching and playing ping pong while seated (Boto et al. 2018). This opens an exciting possibility to study the human brain in more naturalistic setups. Using a wearable MEG system, we could study the neural correlates of 3D physical head rotation and location. It would be possible to achieve the optimal balance between rigorous and flexible experimental control and multisensory navigation by combining the wearable MEG system with immersive VR techniques (e.g. a virtual treadmill, a motion tracking system). Furthermore, OPM-MEG will permit a different perspective on navigation compared to that of fMRI, which I discuss in the next section.

7.6.2 Temporal dynamics of 3D spatial memory and navigation

In this thesis, I mainly focused on which brain areas encode 3D spatial information and whether vertical and horizontal information are equally well encoded in the spatially distributed pattern of neural activity using fMRI. However, neural code is not only spatial (e.g. which neurons are activated), but also temporal (e.g. the timing and temporal order of neural activation). For instance, we might recall the location of an object in 3D space by serially accessing vertical and horizontal information. Or, vertical and horizontal information might be processed simultaneously (parallel processing). Moreover, different types of information might be encoded in different frequency domains. Neural oscillations is an important topic in cognitive neuroscience. In spatial neuroscience, a prominent low-frequency oscillation (theta) in the hippocampus has

been extensively studied (see review, Buzsáki (2005)). For example, the timing of a place cell's firing is linked to the animal's location within a place field and the background theta oscillation (O'Keefe and Recce 1993). Multiple behavioural correlates of hippocampal theta have been found including speed, view and distance (Watrous et al. 2011; Vass et al. 2016). It was not possible to study this important temporal code in this thesis due to the poor temporal resolution of fMRI. Future studies using MEG or invasive electrophysiology will allow us to develop a fuller understanding of spatial representations in 3D by incorporating temporal information.

7.6.3 From neuroimaging to invasive cell recordings

As I explained in Chapter 2 (Section 2.5.3), fMRI measures spatiotemporally filtered neural activity, and this often gives rise to the criticism that fMRI cannot validly study cellular representations like place cells, head direction cells and grid cells. Taube et al. (2013) argued that it might be intrinsically impossible to measure head direction signals using the BOLD signal. They suggest that this is because the head direction signal is always "on" regardless of a participant's active navigation state, and that numerous head direction cells tuned to different directions covering the entire 360° (even larger when the vertical dimension is added) is mixed within a single fMRI voxel. Indeed, this is a question I asked myself throughout my PhD. However, I believe that fMRI research can provide valuable insights and evidence, albeit indirect, of spatial representation when well-designed experiments and analysis techniques such as repetition suppression and multivariate analysis are used (Section 2.8, 2.9).

Although Taube et al. (2013) suggested that direction (or place) tuning is mixed within a voxel, previous electrophysiological studies have typically measured less than a hundred neurons simultaneously (often much fewer), and a complete distribution of tuning functions of millions of neurons within the entire hippocampus, thalamus or

entorhinal cortex is as yet unknown. Successful detection of 3D place and direction information in the hippocampus, retrosplenial and thalamus in this thesis, and in the previous fMRI studies reviewed in Section 1.7.1, imply that place cells and head direction cells are not perfectly evenly distributed at a macroscopic level; rather there is some unevenness (e.g. voxel 1 contains slightly more neurons tuned to +60° than -30°) that is detectable by fMRI.

However, there is still an inherent limitation of fMRI studies which is that multiple neuronal encoding scenarios can give rise to the same fMRI signal. For example, I found the fMRI multivoxel pattern activity in the anterior hippocampus contained 3D location information that was similarly sensitive to the vertical and horizontal axes in Experiment 1. This result does not guarantee that individual place cells have a complete 3D receptive field (see the schematic representation in Figure 7.1A). Even when one population of place cells encodes only vertical height (e.g. “y=1” and “y=2” cells in Figure 7.1B) and the other population encodes only horizontal location (e.g. “x=1” and “x=2” cells in Figure 7.1B), the fMRI analysis will detect both vertical and horizontal location information when these cells are mixed within a 3 mm isotropic voxel. Therefore, while I can claim that the anterior hippocampus contains 3D location information, but I cannot claim that place cells in the anterior hippocampus have isotropic 3D receptive fields. Similarly, a vertical head direction signal is less likely to be detected by fMRI if the majority of head direction cells are tuned to azimuth and only a minority of the cells are tuned to pitch (Finkelstein et al. 2015) as I already discussed in Experiment 3 (Section 5.4). Thus, the current fMRI findings should be complemented by invasive cellular electrophysiology studies in the future.

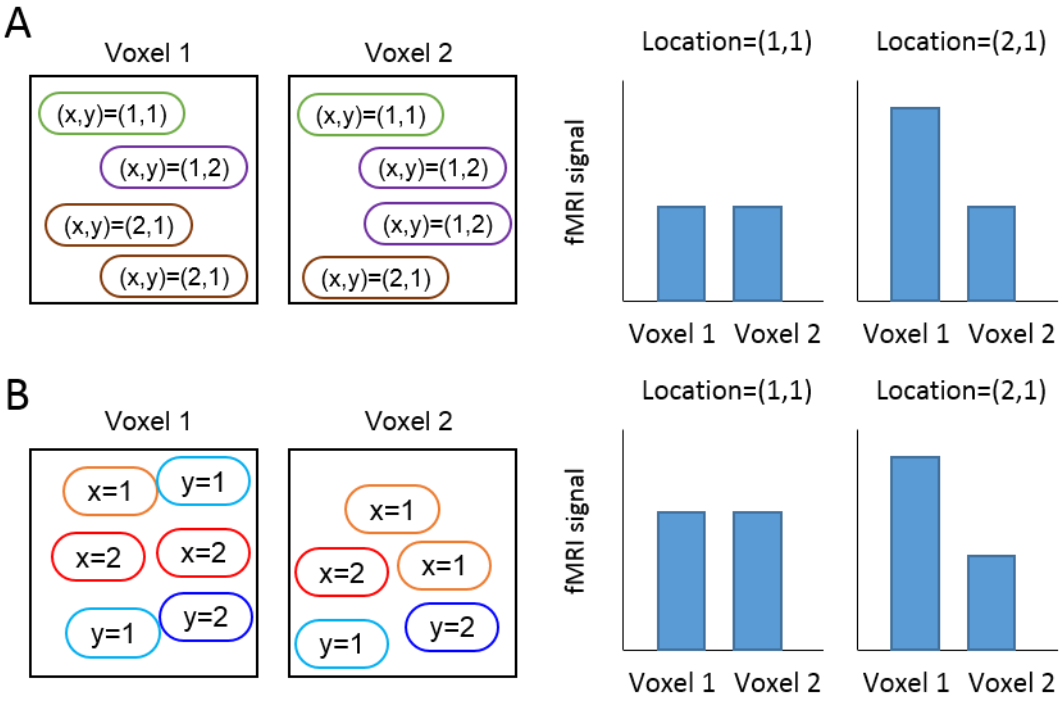


Figure 7.1 Different neuronal encoding hypotheses leading to similar fMRI results. (A) Each cell is tuned to a specific 3D location, e.g. (x,y)=(1,1), (1,2), (2,1). For simplicity, assume that x represents the horizontal coordinate and y represents the vertical coordinate. The cell tuned to (1,1) is equally distributed in voxel 1 and voxel 2, but the cell tuned to (2,1) is more abundant in voxel 1, leading to different fMRI multivoxel patterns for these two locations. (B) Each cell is tuned to either vertical or horizontal coordinates and 3D location information is encoded by a combination of these neurons. For example, when a participant is located at (2,1), the cells tuned to x=2 and the cells tuned to y=1 are activated. Due to an uneven distribution of these cells, location (1,1) and location (2,1) result in different multivoxel fMRI patterns. Note the similarity of fMRI patterns for conjunctive encoding (A) and separate encoding (B) hypotheses.

7.6.4 Individual differences

As neuroscientists, we are often interested in drawing a general conclusion about how the brain works, but we should not forget that none of the seven billion human brains in this world is identical. There are large individual differences in spatial ability, from those who easily get lost even with the help of modern technology (like the author of this thesis) to those who can draw an accurate map after just one brief visit to a location. There are multiple methods to quantify individual differences in navigation and spatial ability including subjective questionnaires (Hegarty et al. 2002) and a more objective measures of spatial performance during tasks. In my thesis, I found a significant relationship between neural activity (vertical direction information in the retrosplenial

cortex) and behavioural performance (angular error during a direction judgment task) only in Experiment 3. The simple scanning tasks and the pre-scan training that resulted in overall high behavioural accuracy in my experiments (because I wanted all participants to have an accurate representation of the virtual 3D space so I could seek unambiguous neural correlates of spatial encoding) could have made it difficult to reveal subtle individual differences in the other experiments. More challenging behavioural tasks that result in a large variance in performance could be used in the future. It would also be possible to study the neural correlates of individual differences in 3D navigation strategies, e.g. vertical first or horizontal first (Hölscher et al. 2006; Büchner et al. 2007) using structural MRI measures such as voxel based morphometry (Ashburner and Friston 2000). It should also be noted that the participants recruited in my experiments (and in many cognitive neuroscience studies) had a very specific demographic, young and highly educated attendees of a UK university. That such participants are not fully representative of the population more generally (who have a broader range of ages, education, socioeconomic states, etc.) should be borne in mind.

Another related issue is that the brain is a highly adaptive and flexible organ shaped by our experiences. One of the most well-known examples of this in spatial neuroscience is in London taxi drivers who have structural changes in their hippocampus after years of spatial knowledge accumulation (Maguire et al. 2000; Woollett and Maguire 2011). An interesting question in 3D spatial representation is whether extensive exposure to 3D navigation would yield different spatial encoding in the brain, such as more volumetric and symmetric representation of 3D space. For example, I found separate encoding of vertical and horizontal direction information in the retrosplenial cortex, thalamus and subiculum in Experiment 3, but professional astronauts or pilots might encode vertical and horizontal direction information equally well in the same brain structure. Studying the brains of specialised populations could also provide insights into the origin of interspecies differences in 3D spatial encoding. As I reviewed in Section

1.6, rodents, birds and fish show different behaviours in 3D, e.g. birds seem more confident of their vertical location memory while rats seem more confident of horizontal location memory (Flores-Abreu et al. 2014), and it is unknown whether the difference is evolutionary or developmental. If an astronaut's brain shows a more similar response profile to a bat's brain than a fellow human brain, it would favour the developmental account. Of course, the difference between the brain of astronauts and undergraduate psychology students might be negligible given that even astronauts or divers spend more time on the ground than in a volumetric space, but this is intriguing hypothesis to test in the future.

7.6.5 Clinical implications

In relation to individual differences, one can also ask whether there are any clinical implications (e.g. aiding diagnosis) in understanding 3D spatial representation in the human brain. Spatial disorientation is a common symptom associated with normal aging and various brain pathologies (Lester et al. 2017). The hippocampus, retrosplenial cortex, entorhinal cortex – the main brain structures studied in this thesis – are known to be compromised early in the course of Alzheimer's disease (Braak and Braak 1991; Jack et al. 1997; Pengas et al. 2012). In addition to general structural and metabolic changes in these regions during Alzheimer's disease, more specific dysfunction was observed in the entorhinal cortex. Kunz et al. (2015) showed that young adults with a genetic risk for Alzheimer's disease (APOE-ε4 carriers) showed reduced grid-like representations, which could potentially serve as a prognostic marker decades before disease onset.

Overall, while it is clear that knowledge about spatial cognition has clinical relevance, the question is what "3D" can add over and above the existing link between spatial cognition in 2D and neurological conditions. I can only speculate about this because I

did not study any patient populations in my experiments and, to the best of my knowledge, this issue has not been addressed yet in the published literature. One potential contribution is the extensive use of the vestibular system for 3D navigation compared to 2D. To maintain balance and orientation in sloped and vertical environments, vestibular inputs are crucial. Vestibular input is closely linked to hippocampal structure and function. For example, patients with partial bilateral vestibulopathy showed reduced hippocampal volume, higher spatial anxiety and a more dominant route-finding strategy (Kremmyda et al. 2016). Therefore, navigation tasks in 3D might be more sensitive in revealing some neurological conditions. Indeed, it has been proposed that 3D spatial memory and navigation should be included in routine neurological examinations (Brandt et al. 2017).

7.6.6 From representation to action in space

One limitation of my experiments was that participants were passively moved while their brain activity was scanned inside the MRI scanner. This passive movement approach (without active wayfinding behaviour) was appropriate for my main research goal - to investigate the neural representation of 3D spatial information - because it allowed optimal sampling of 3D locations and directions. Similarly, even in the animal literature, place cells, head direction cells and grid cells have been mostly recorded when animals simply run in an experimental arena without goal-directed behaviour. However, the neural code for 3D location or direction is just the starting point for understanding our active navigation behaviour. We ultimately want to understand how we know our location and direction from sensorimotor inputs, compute a shortcut to the goal location and update the route online. It is unlikely that active navigation involves just a few structures that contain place cells, head direction cells and grid cells (e.g. hippocampus, entorhinal cortex and retrosplenial cortex – the main ROIs in this thesis). Rather, it likely also engages a parietal network for egocentric spatial processing (e.g.

“the goal is above me”), a striatum network for automatic stimulus-response (e.g. “turn upward when the tree is visible”) and the prefrontal cortex for controlling the hippocampal and striatal systems (see a review on the cognitive architecture of spatial navigation - Chersi and Burgess (2015)).

Regarding navigation in 3D, the computation of energy cost when moving vertically should be considered when we plan a route, unless we are making a spacewalk. Where is this computation implemented in the brain? If it happens outside the medial temporal lobe, how does it influence the spatial code in the hippocampal formation? How do we balance competing goals – minimising energy cost versus minimising travel time, if a shortcut involves a steep vertical movement? These are some of the questions that should be addressed in the future.

7.6.7 Beyond the spatial domain

In this thesis, I investigated how 3D space is encoded in the human brain, particularly in the hippocampal formation. However, recent literature suggests that the hippocampal formation encodes not only physical space but also more abstract cognitive variables, indeed this wider definition of a ‘cognitive map’ was envisaged from its inception (Tolman 1948). For example, a changing auditory tone (from low frequency to high frequency) can be mapped in a 1D auditory space. Aronov et al. (2017) found place cells in the hippocampus that fired at specific auditory frequencies and grid cells in the entorhinal cortex that fired at multiple auditory frequencies. In humans, entorhinal cortex showed a grid-like signal for a 2D visual space (eye gaze location, Julian et al. (2018); Nau et al. (2018)) and other 2D stimulus spaces (e.g. the length of the neck and legs of cartoon stimuli, Constantinescu et al. (2016)). In a social context, for example, each person could be described in abstract 2D space where one axis is “power” and the other axis is “trustworthiness”. This social 2D space was

proposed to be encoded in the hippocampus (Tavares et al. 2015). If the hippocampal formation indeed organises knowledge about any behaviourally-relevant information in a map-like fashion, the map cannot be limited to 2D, because our world is multidimensional and complex.

During my PhD, I aimed to understand how physical space mapping is generalised from 2D to 3D, and the same can be asked of non-physical space mapping. For example, it would be interesting to test whether the FCC grid representation is developed in the entorhinal cortex when participants learn abstract cartoon stimuli with three attributes (e.g. the length of neck, legs and arms). Furthermore, unlike physical 3D space where x, y and z axes are orthogonal, attributes might covary in abstract space. How the brain builds an efficient representation of abstract high dimensional space is a key question in neuroscience.

7.7 Summary and conclusions

The spatial world is 3D and humans and other animals move horizontally but also vertically within it. Most previous studies on spatial encoding have been conducted on a simple 2D plane and there was a dearth of knowledge about how 3D spatial information is encoded in the human brain. To start to fill this gap in our knowledge, I created various 3D virtual environments ranging from a conventional building to a zero gravity spaceship and measured the behavioural and fMRI responses while participants explored these 3D environments in a quasi-naturalistic setup. Behaviourally, participants showed overall high performance for their 3D location and direction judgments, with vertical pitch often regarded as more salient. Navigationally relevant structures – the hippocampus, retrosplenial cortex, entorhinal cortex and thalamus – that contain place cells, head direction cells, grid cells on a 2D plane also

showed the relevant spatial codes in my 3D environments. In a multi-level environment, I observed evidence of a hierarchical representation of space.

Overall, these findings demonstrate the capacity of the human brain to implement a flexible and efficient representation of 3D space. I believe that this thesis can serve as a stepping-stone to understand how we navigate in complex real worlds such as multi-level buildings, undulating terrain and in the open air or under water. I also hope that the investigation of the spatial representation of 3D space going forward will provide new insights into the more general question in neuroscience of how the brain has adapted to solve high-dimensional cognitive problems.

References

Adler DH, Pluta J, Kadivar S, Craige C, Gee JC, Avants BB, Yushkevich PA. 2014. Histology-derived volumetric annotation of the human hippocampal subfields in postmortem MRI. Neuroimage. 84:505–523.

Aghajan ZM, Acharya L, Moore JJ, Cushman JD, Vuong C, Mehta MR. 2014. Impaired spatial selectivity and intact phase precession in two-dimensional virtual reality. Nat Neurosci. 18:121–128.

Aguirre GK. 2007. Continuous carry-over designs for fMRI. Neuroimage. 35:1480–1494.

Aguirre GK, D'Esposito M. 1999. Topographical disorientation: A synthesis and taxonomy. Brain. 122:1613–1628.

Alexander AS, Nitz DA. 2015. Retrosplenial cortex maps the conjunction of internal and external spaces. Nat Neurosci. 18:1143–1151.

Amaral D, Lavenex P. 2006. Hippocampal Neuroanatomy. In: Andersen P, Morris R, Amaral D, Bliss T, O'Keefe J, editors. The Hippocampus Book. Oxford University Press. p. 37–114.

Aoki H, Ohno R, Yamaguchi T. 2005. The effect of the configuration and the interior design of a virtual weightless space station on human spatial orientation. Acta Astronaut. 56:1005–1016.

Arleo A, Dejean C, Allegraud P, Khamassi M, Zugaro MB, Wiener SI. 2013. Optic flow stimuli update anterodorsal thalamus head direction neuronal activity in rats. J Neurosci. 33:16790–16795.

Aronov D, Nevers R, Tank DW. 2017. Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. Nature. 543:719–722.

Aronov D, Tank DW. 2014. Engagement of neural circuits underlying 2D spatial navigation in a rodent virtual reality system. Neuron. 84:442–456.

Ashburner J, Friston KJ. 2000. Voxel-based morphometry-the methods. Neuroimage. 11:805–821.

Attwell D, Buchan AM, Charpak S, Lauritzen M, MacVicar BA, Newman EA. 2010. Glial and neuronal control of brain blood flow. Nature. 468:232–243.

Auger SD, Mullally SL, Maguire EA. 2012. Retrosplenial cortex codes for permanent landmarks. PLoS One. 7:e43620.

Auger SD, Zeidman P, Maguire EA. 2015. A central role for the retrosplenial cortex in de novo environmental learning. Elife. 4:e09031.

Averbeck BB, Latham PE, Pouget A. 2006. Neural correlations, population coding and computation. Nat Rev Neurosci. 7:358–366.

Avery G, Day R. 1969. Basis of the horizontal-vertical illusion. J Exp Psychol. 81:376–380.

Baillet S. 2017. Magnetoencephalography for brain electrophysiology and imaging. Nat Neurosci. 20:327–339.

Bakker A, Kirwan CB, Miller M, Stark CEL. 2008. Pattern separation in the human hippocampal CA3 and dentate gyrus. Science. 319:1640–1642.

Balaguer J, Spiers H, Hassabis D, Summerfield C. 2016. Neural mechanisms of hierarchical planning in a virtual subway network. Neuron. 90:893–903.

Barnett-Cowan M, Bülthoff HH. 2013. Human path navigation in a three-dimensional world. Behav Brain Sci. 36:544–545.

Barron HC, Dolan RJ, Behrens TEJ. 2013. Online evaluation of novel choices by simultaneous representation of multiple memories. Nat Neurosci. 16:1492–1498.

Barron HC, Garvert MM, Behrens TEJ. 2016. Repetition suppression: a means to index neural representations using BOLD? Philos Trans R Soc B Biol Sci. 371:20150355.

Baumann O, Mattingley JB. 2010. Medial parietal cortex encodes perceived heading direction in humans. J Neurosci. 30:12897–12901.

Baumann O, Mattingley JB. 2013. Dissociable representations of environmental size and complexity in the human hippocampus. J Neurosci. 33:10526–10533.

Bellmund JLS, Deuker L, Schröder TN, Doeller CF. 2016. Grid-cell representations in mental simulation. Elife. 5:e17089.

Bense S, Janusch B, Vucurevic G, Bauermann T, Schlindwein P, Brandt T, Stoeter P, Dieterich M. 2006. Brainstem and cerebellar fMRI-activation during horizontal and vertical optokinetic stimulation. Exp Brain Res. 174:312–323.

Berens P. 2009. CircStat : A MATLAB toolbox for circular statistics. J Stat Softw. 31.

Blessing EM, Beissner F, Schumann A, Brünner F, Bär K-J. 2016. A data-driven approach to mapping cortical and subcortical intrinsic functional connectivity along the longitudinal hippocampal axis. Hum Brain Mapp. 37:462–476.

Boccara CN, Sargolini F, Thoresen VH, Solstad T, Witter MP, Moser EI, Moser MB. 2010. Grid cells in pre-and parasubiculum. Nat Neurosci. 13:987–994.

Boto E, Holmes N, Leggett J, Roberts G, Shah V, Meyer SS, Muñoz LD, Mullinger KJ, Tierney TM, Bestmann S, Barnes GR, Bowtell R, Brookes MJ. 2018. Moving magnetoencephalography towards real-world applications with a wearable system. Nature. 555:657–661.

Braak H, Braak E. 1991. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol. 82:239–259.

Brandon MP, Koenig J, Leutgeb JK, Leutgeb S. 2014. New and distinct hippocampal

place codes are generated in a new environment during septal inactivation. Neuron. 82:789–796.

Brandt T, Huber M, Schramm H, Kugler G, Dieterich M, Glasauer S. 2015. “Taller and shorter”: Human 3-D spatial memory distorts familiar multilevel buildings. PLoS One. 10:e0141257.

Brandt T, Zwergal A, Glasauer S. 2017. 3-D spatial memory and navigation: Functions and disorders. Curr Opin Neurol. 30:90–97.

Brodt S, Pöhlchen D, Flanagin VL, Glasauer S, Gais S, Schönauer M. 2016. Rapid and independent memory formation in the parietal cortex. Proc Natl Acad Sci. 113:13251–13256.

Büchner S, Hölscher C, Strube G. 2007. Path choice heuristics for navigation related to mental representations of a building. In: Vosniadou S, Kayser D, Protopapas A, editors. Proceedings of the European Cognitive Science Conference. Delphi, Greece: Taylor and Francis. p. 504–509.

Burgess N, Barry C, Keefe JO. 2007. An oscillatory interference model of grid cell firing. Hippocampus. 17:801–812.

Burt de Perera T, Holbrook RI, Davis V. 2016. The representation of three-dimensional space in fish. Front Behav Neurosci. 10:40.

Buzsáki G. 2005. Theta rhythm of navigation: Link between path integration and landmark navigation, episodic and semantic memory. Hippocampus. 15:827–840.

Calton JL, Taube JS. 2005. Degradation of head direction cell activity during inverted locomotion. J Neurosci. 25:2420–2428.

Carlin JD, Calder AJ, Kriegeskorte N, Nili H, Rowe JB. 2011. A head view-invariant representation of gaze direction in anterior superior temporal sulcus. Curr Biol. 21:1817–1821.

Chadwick MJ, Jolly AEJ, Amos DP, Hassabis D, Spiers HJ. 2015. A goal direction signal in the human entorhinal/subicular region. Curr Biol. 25:87–92.

Chen G, Manson D, Cacucci F, Wills TJ. 2016. Absence of visual input results in the disruption of grid cell firing in the mouse. Curr Biol. 26:2335–2342.

Chersi F, Burgess N. 2015. The cognitive architecture of spatial navigation: Hippocampal and striatal contributions. Neuron. 88:64–77.

Cho J, Sharp PE. 2001. Head direction, place, and movement correlates for cells in the rat retrosplenial cortex. Behav Neurosci. 115:3–25.

Chrastil ER, Sherrill KR, Hasselmo ME, Stern CE. 2015. There and back again: Hippocampus and retrosplenial cortex track homing distance during human path integration. J Neurosci. 35:15442–15452.

Constantinescu AO, O’Reilly JX, Behrens TEJ. 2016. Organizing conceptual

knowledge in humans with a gridlike code. Science. 352:1464–1468.

Creem-Regehr SH, Gooch AA, Sahm CS, Thompson WB. 2004. Perceiving virtual geographical slant: Action influences perception. J Exp Psychol Hum Percept Perform. 30:811–821.

Cullen KE, Taube JS. 2017. Our sense of direction: progress, controversies and challenges. Nat Neurosci. 20:1465–1473.

Dabaghian Y, Brandt VL, Frank LM. 2014. Reconceiving the hippocampal map as a topological template. Elife. 3:e03476.

Dalton MA, Zeidman P, Barry DN, Williams E, Maguire EA. 2017. Segmenting subregions of the human hippocampus on structural magnetic resonance image scans: An illustrated tutorial. Brain Neurosci Adv. 1:239821281770144.

Davis VA, Holbrook RI, Schumacher S, Guilford T, Perera TB de. 2014. Three-dimensional spatial cognition in a benthic fish, corydoras aeneus. Behav Processes. 109:151–156.

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.

Derdikman D, Whitlock JR, Tsao A, Fyhn M, Hafting T, Moser M-B, Moser EI. 2009. Fragmentation of grid cell maps in a multicompartment environment. Nat Neurosci. 12:1325–1332.

Doeller CF, Barry C, Burgess N. 2010. Evidence for grid cells in a human memory network. Nature. 463:657–661.

Doeller CF, King JA, Burgess N. 2008. Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. Proc Natl Acad Sci. 105:5915–5920.

Dordek Y, Soudry D, Meir R, Derdikman D. 2016. Extracting grid cell characteristics from place cell inputs using non-negative principal component analysis. Elife. 5:e10094.

Dyde RT, Jenkin MR, Harris LR. 2006. The subjective visual vertical and the perceptual upright. Exp Brain Res. 173:612–622.

Eichenbaum H, Wiener SI, Shapiro ML, Cohen NJ. 1989. The organization of spatial coding in the hippocampus: a study of neural ensemble activity. J Neurosci. 9:2764–2775.

Ekstrom A. 2010. How and when the fMRI BOLD signal relates to underlying neural activity: The danger in dissociation. Brain Res Rev. 62:233–244.

Ekstrom A, Suthana N, Millett D, Fried I, Bookheimer S. 2009. Correlation between BOLD fMRI and theta-band local field potentials in the human hippocampal area. J

Neurophysiol. 101:2668–2678.

Ekstrom AD, Kahana MJ, Caplan JD, Fields TA, Isham EA, Newman EL, Fried I. 2003. Cellular networks underlying human spatial navigation. Nature. 425:184–187.

Epstein RA, Parker WE, Feiler AM. 2007. Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. J Neurosci. 27:6141–6149.

Evensmoen HR, Ladstein J, Hansen TI, Møller JA, Witter MP, Nadel L, Håberg AK. 2015. From details to large scale: The representation of environmental positions follows a granularity gradient along the human hippocampal and entorhinal anterior-posterior axis. Hippocampus. 25:119–135.

Fiete IR, Burak Y, Brookings T. 2008. What grid cells convey about rat location. J Neurosci. 28:6858–6871.

Finkelstein A, Derdikman D, Rubin A, Foerster JN, Las L, Ulanovsky N. 2015. Three-dimensional head-direction coding in the bat brain. Nature. 517:159–164.

Finkelstein A, Las L, Ulanovsky N, A F, L L, Ulanovsky N. 2016. 3D maps and compasses in the brain. Annu Rev Neurosci. 39:171–196.

Flores-Abreu IN, Hurly TA, Ainge JA, Healy SD. 2014. Three-dimensional space: locomotory style explains memory differences in rats and hummingbirds. Proc R Soc B Biol Sci. 281:20140301.

Friston KJ, Glaser DE, Henson RNA, Kiebel S, Phillips C, Ashburner J. 2002. Classical and Bayesian inference in neuroimaging: Applications. Neuroimage. 16:484–512.

Fyhn M, Molden S, Witter MP, Moser EI, Moser M-B. 2004. Spatial representation in the entorhinal cortex. Science. 305:1258–1264.

Garrido L, Vaziri-Pashkam M, Nakayama K, Wilmer J. 2013. The consequences of subtracting the mean pattern in fMRI multivariate correlation analyses. Front Neurosci. 7:1–4.

Ginosar G, Finkelstein A, Rubin A, Las L, Ulanovsky N. 2016. 3D grid cells and border cells in flying bats. In: Society for Neuroscience.

Giocomo LM, Hasselmo ME. 2008. Time constants of h current in layer II stellate cells differ along the dorsal to ventral axis of medial entorhinal cortex. J Neurosci. 28:9414–9425.

Glasauer S. 2005. Vestibular and motor processing for head direction signals. In: Wiener SI, Taube JS, editors. Head direction cells and the neural mechanisms of spatial orientation. Cambridge, MA: The MIT press. p. 113–135.

Goodale MA, Milner AD. 1992. Separate visual pathways for perception and action. Trends Neurosci. 15:20–25.

Goodridge JP, Taube JS. 1997. Interaction between the postsubiculum and anterior

thalamus in the generation of head direction cell activity. J Neurosci. 17:9315–9330.

Gramann K, Wing S, Jung TP, Viirre E, Riecke BE. 2012. Switching spatial reference frames for yaw and pitch navigation. Spat Cogn Comput. 12:159–194.

Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzchak Y, Malach R. 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron. 24:187–203.

Grobéty M-C, Schenk F. 1992a. Spatial learning in a three-dimensional maze. Anim Behav. 43:1011–1020.

Grobéty M-C, Schenk F. 1992b. The influence of spatial irregularity upon radial-maze performance in the rat. Anim Learn Behav. 20:393–400.

Gross CG, Bender DB, Rocha-Miranda CE. 1969. Visual receptive fields of neurons in inferotemporal cortex of the monkey. Science. 166:1303–1306.

Hafting T, Fyhn M, Molden S, Moser M-B, Moser EI. 2005. Microstructure of a spatial map in the entorhinal cortex. Nature. 436:801–806.

Han X, Becker S. 2014. One spatial map or many? Spatial coding of connected environments. J Exp Psychol Learn Mem Cogn. 40:511–531.

Hartley T, Burgess N, Lever C, Cacucci F, O'Keefe J. 2000. Modeling place fields in terms of the cortical inputs to the hippocampus. Hippocampus. 10:369–379.

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.

Hassabis D, Chu C, Rees G, Weiskopf N, Molyneux PD, Maguire EA. 2009. Decoding neuronal ensembles in the human hippocampus. Curr Biol. 19:546–554.

Hasselmo ME, Brandon MP. 2012. A model combining oscillations and attractor dynamics for generation of grid cell firing. Front Neural Circuits. 6:1–13.

Haxby J V, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science. 293:2425–2430.

Hayman R, Verriotis MA, Jovalekic A, Fenton AA, Jeffery KJ. 2011. Anisotropic encoding of three-dimensional space by place cells and grid cells. Nat Neurosci. 14:1182–1188.

Hayman RMA, Casali G, Wilson JJ, Jeffery KJ. 2015. Grid cells on steeply sloping terrain: evidence for planar rather than volumetric encoding. Front Psychol. 6:1–14.

Hegarty M, Richardson A, Montello D. 2002. Development of a self-report measure of environmental spatial ability. Intelligence. 30:425–447.

Henderson J, Hurly TA, Healy SD. 2001. Rufous hummingbirds' memory for flower location. Anim Behav. 61:981–986.

Henderson J, Hurly TA, Healy SD. 2006. Spatial relational learning in rufous hummingbirds (Selasphorus rufus). Anim Cogn. 9:201–205.

Hirtle SC, Jonides J. 1985. Evidence of hierarchies in cognitive maps. Mem Cognit. 13:208–217.

Hodgetts CJ, Postans M, Shine JP, Jones DK, Lawrence AD, Graham KS. 2015. Dissociable roles of the inferior longitudinal fasciculus and fornix in face and place perception. Elife. 4:1–25.

Holbrook RI, Burt de Perera T. 2009. Separate encoding of vertical and horizontal components of space during orientation in fish. Anim Behav. 78:241–245.

Holbrook RI, Burt de Perera T. 2011. Fish navigation in the vertical dimension: Can fish use hydrostatic pressure to determine depth? Fish Fish. 12:370–379.

Holbrook RI, Burt de Perera T. 2013. Three-dimensional spatial cognition: Freely swimming fish accurately learn and remember metric information in a volume. Anim Behav. 86:1077–1083.

Hölscher C, Jacob W, Mallot HA. 2003. Reward modulates neuronal activity in the hippocampus of the rat. Behav Brain Res. 142:181–191.

Hölscher C, Meilinger T, Vrachliotis G, Brösamle M, Knauff M. 2006. Up the down staircase: Wayfinding strategies in multi-level buildings. J Environ Psychol. 26:284–299.

Horiuchi TK, Moss CF. 2015. Grid cells in 3D: Reconciling data and models. Hippocampus. 1500:1–51.

Horner AJ, Bisby JA, Zotow E, Bush D, Burgess N. 2016. Grid-like processing of imagined navigation. Curr Biol. 26:842–847.

Howard LR, Javadi AH, Yu Y, Mill RD, Morrison LC, Knight R, Loftus MM, Staskute L, Spiers HJ. 2014. The hippocampus and entorhinal cortex encode the path and Euclidean distances to goals during navigation. Curr Biol. 24:1331–1340.

Hsieh L-T, Gruber MJ, Jenkins LJ, Ranganath C. 2014. Hippocampal activity patterns carry information about objects in temporal context. Neuron. 81:1165–1178.

Hsieh L-T, Ranganath C. 2015. Cortical and subcortical contributions to sequence retrieval: Schematic coding of temporal context in the neocortical recollection network. Neuroimage. 121:78–90.

Indovina I, Maffei V, Mazzarella E, Sulpizio V, Galati G, Lacquaniti F. 2016. Path integration in 3D from visual motion cues: A human fMRI study. Neuroimage. 142:512–521.

Jack CR, Petersen RC, Xu YC, Waring SC, O'Brien PC, Tangalos EG, Smith GE, Ivnik

RJ, Kokmen E. 1997. Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. Neurology. 49:786–794.

Jacob P-Y, Casali G, Spieser L, Page H, Overington D, Jeffery K. 2017. An independent, landmark-dominated head-direction signal in dysgranular retrosplenial cortex. Nat Neurosci. 20:173–175.

Jacobs J, Kahana MJ, Ekstrom AD, Mollison M V., Fried I. 2010. A sense of direction in human entorhinal cortex. Proc Natl Acad Sci. 107:6487–6492.

Jacobs J, Weidemann CT, Miller JF, Solway A, Burke JF, Wei X-X, Suthana N, Sperling MR, Sharan AD, Fried I, Kahana MJ. 2013. Direct recordings of grid-like neuronal activity in human spatial navigation. Nat Neurosci. 16:1188–1190.

Jankowski MM, Ronnqvist KC, Tsanov M, Vann SD, Wright NF, Erichsen JT, Aggleton JP, O'Mara SM. 2013. The anterior thalamus provides a subcortical circuit supporting memory and spatial navigation. Front Syst Neurosci. 7:1–12.

Jeffery KJ, Jovalekic A, Verriotis M, Hayman R. 2013. Navigating in a three-dimensional world. Behav Brain Sci. 36:523–543.

Jeffery KJ, Wilson JJ, Casali G, Hayman RM. 2015. Neural encoding of large-scale three-dimensional space—properties and constraints. Front Psychol. 6:1–12.

Jovalekic A, Hayman R, Becares N, Reid H, Thomas G, Wilson J, Jeffery K. 2011. Horizontal biases in rats' use of three-dimensional space. Behav Brain Res. 222:279–288.

Julian JB, Keinath AT, Frazzetta G, Epstein RA. 2018. Human entorhinal cortex represents visual space using a boundary-anchored grid. Nat Neurosci. 21:191–194.

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.

Kahn I, Andrews-Hanna JR, Vincent JL, Snyder AZ, Buckner RL. 2008. Distinct cortical anatomy linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. J Neurophysiol. 100:129–139.

Keinath AT, Wang ME, Wann EG, Yuan RK, Dudman JT, Muzzio IA. 2014. Precise spatial coding is preserved along the longitudinal hippocampal axis. Hippocampus. 24:1533–1548.

Kinsella-Shaw JM, Shaw B, Turvey MT. 1992. Perceiving "Walk-on-able" slopes. Ecol Psychol. 4:223–239.

Kjelstrup KB, Solstad T, Brun VH, Hafting T, Leutgeb S, Witter MP, Moser EI, Moser M-B. 2008. Finite scale of spatial representation in the hippocampus. Science. 321:140–143.

Klatzky RL, Loomis JM, Beall AC, Chance SS, Golledge RG. 1998. Updating an egocentric spatial representationduring real, imagined and virtual locomotion. Psychol Sci. 9:293–298.

Knierim JJ, McNaughton BL. 2001. Hippocampal place-cell firing during movement in three-dimensional space. J Neurophysiol. 85:105–116.

Knierim JJ, McNaughton BL, Poe GR. 2000. Three-dimensional spatial selectivity of hippocampal neurons during space flight. Nat Neurosci. 3:209–210.

Kobayashi Y, Amaral DG. 2003. Macaque monkey retrosplenial cortex: II. Cortical afferents. J Comp Neurol. 466:48–79.

Kremmyda O, Hüfner K, Flanagin VL, Hamilton DA, Linn J, Strupp M, Jahn K, Brandt T. 2016. Beyond dizziness: Virtual navigation, spatial anxiety and hippocampal volume in bilateral vestibulopathy. Front Hum Neurosci. 10:139.

Kriegeskorte N, Goebel R, Bandettini P. 2006. Information-based functional brain mapping. Proc Natl Acad Sci. 103:3863–3868.

Kriegeskorte N, Mur M, Bandettini P. 2008. Representational similarity analysis - connecting the branches of systems neuroscience. Front Syst Neurosci. 2:4.

Krupic J, Bauza M, Burton S, Lever C, O'Keefe J. 2014. How environment geometry affects grid cell symmetry and what we can learn from it. Philos Trans R Soc Lond B Biol Sci. 369:20130188.

Krupic J, Burgess N, O'Keefe J. 2012. Neural Representations of Location Composed of Spatially Periodic Bands. Science. 337:853–857.

Kunz L, Schröder TN, Lee H, Montag C, Lachmann B, Sariyska R, Reuter M, Stirnberg R, Stöcker T, Messing-Floeter PC, Fell J, Doeller CF, Axmacher N. 2015. Reduced grid-cell-like representations in adults at genetic risk for Alzheimer's disease. Science. 350:430–433.

Lancaster JL, Rainey LH, Summerlin JL, Freitas CS, Fox PT, Evans AC, Toga AW, Mazziotta JC. 1997. Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. Hum Brain Mapp. 5:238–242.

Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov P V, Nickerson D, Mikiten S a, Fox PT. 2000. Automated Tailairach Atlas Labels for Functional Brain Mapping. Hum Brain Mapp. 10:120–131.

Laurens J, Kim B, Dickman JD, Angelaki DE. 2016. Gravity orientation tuning in macaque anterior thalamus. Nat Neurosci. 19:1566–1568.

Layton O, O'Connell T, Phillips F. 2009. The traveling salesman problem in the natural environment. J Vis. 9:1145.

Lee ACH, Scahill VL, Graham KS. 2008. Activating the medial temporal lobe during

oddity judgment for faces and scenes. Cereb Cortex. 18:683–696.

Lee SA, Miller JF, Watrous AJ, Sperling MR, Sharan A, Worrell GA, Berry BM, Aronson JP, Davis KA, Gross RE, Lega B, Sheth S, Das SR, Stein JM, Gorniak R, Rizzuto DS, Jacobs J. 2018. Electrophysiological signatures of spatial boundaries in the human subiculum. J Neurosci. 38:3265–3272.

Lester AW, Moffat SD, Wiener JM, Barnes CA, Wolbers T. 2017. The aging navigational system. Neuron. 95:1019–1035.

Lever C, Burton S, Jeewajee A, O'Keefe J, Burgess N. 2009. Boundary vector cells in the subiculum of the hippocampal formation. J Neurosci. 29:9771–9777.

Levine DN, Warach J, Farah M. 1985. Two visual systems in mental imagery: dissociation of "what" and "where" in imagery disorders due to bilateral posterior cerebral lesions. Neurology. 35:1010–1018.

Logothetis NK. 2003. The underpinnings of the BOLD functional magnetic resonance imaging signal. J Neurosci. 23:3963–3971.

Lueschow A, Miller EK, Desimone R. 1994. Inferior temporal mechanisms for invariant object recognition. Cereb Cortex. 4:523–531.

Maddess T, McCourt ME, Blakeslee B, Cunningham RB. 1988. Factors governing the adaptation of cells in area-17 of the cat visual cortex. Biol Cybern. 59:229–236.

Maguire EA, Burgess N, Donnett JG, Frackowiak RSJ, Frith CD, O'Keefe J. 1998. Knowing where and getting there: A human navigation network. Science. 280:921–924.

Maguire EA, Burke T, Phillips J, Staunton H. 1996. Topographical disorientation following unilateral temporal lobe lesions in humans. Neuropsychologia. 34:993–1001.

Maguire EA, Gadian DG, Johnsrude IS, Good CD, Ashburner J, Frackowiak RS, Frith CD. 2000. Navigation-related structural change in the hippocampi of taxi drivers. Proc Natl Acad Sci. 97:4398–4403.

Maguire EA, Nannery R, Spiers HJ. 2006. Navigation around London by a taxi driver with bilateral hippocampal lesions. Brain. 129:2894–2907.

Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage. 19:1233–1239.

Marchette S, Vass L, Ryan J, Epstein R. 2014. Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe. Nat Neurosci. 17:1598–1606.

Marchette SA, Ryan J, Epstein RA. 2017. Schematic representations of local environmental space guide goal-directed navigation. Cognition. 158:68–80.

Markus EJ, Qin Y-L, Leonard B, Skaggs WE, McNaughton BL, Barnes CA. 1995. Interactions between location and task affect the spatial and directional firing of hippocampal neurons. J Neurosci. 15:7079–7094.

Mathis A, Stemmier MB, Herz AVM. 2015. Probable nature of higher-dimensional symmetries underlying mammalian grid-cell activity patterns. Elife. 4:e05979.

Matsumura N, Nishijo H, Tamura R, Eifuku S, Endo S, Ono T. 1999. Spatial- and task-dependent neuronal responses during real and virtual translocation in the monkey hippocampal formation. J Neurosci. 19:2381–2393.

Mazziotta JC, Toga AW, Evans A, Fox P, Lancaster J. 1995. A probabilistic atlas of the human brain: theory and rationale for its development. The International Consortium for Brain Mapping (ICBM). Neuroimage.

McNamara TP, Hardy JK, Hirtle S. 1989. Subjective hierarchies in spatial memory. J Exp Psychol Learn Mem Cogn. 15:211–227.

McNaughton BL, Barnes CA, O'Keefe J. 1983. The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. Exp Brain Res. 52:41–49.

McNaughton BL, Battaglia FP, Jensen O, Moser EI, Moser M-B. 2006. Path integration and the neural basis of the "cognitive map." Nat Rev Neurosci. 7:663–678.

McTighe SM, Mar AC, Romberg C, Bussey TJ, Saksida LM. 2009. A new touchscreen test of pattern separation: effect of hippocampal lesions. Neuroreport. 20:881–885.

Meyer SS, Rossiter H, Brookes MJ, Woolrich MW, Bestmann S, Barnes GR. 2017. Using generative models to make probabilistic statements about hippocampal engagement in MEG. Neuroimage. 149:468–482.

Mhatre H, Gorchetchnikov A, Grossberg S. 2012. Grid cell hexagonal patterns formed by fast self-organized learning within entorhinal cortex. Hippocampus. 22:320–334.

Misaki M, Kim Y, Bandettini P, Kriegeskorte N. 2010. Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. Neuroimage. 53:103–118.

Mittelstaedt H. 1998. Origin and processing of postural information. Neurosci Biobehav Rev. 22:473–478.

Moghaddam M, Kaminsky YL, Zahalka a, Bures J. 1996. Vestibular navigation directed by the slope of terrain. Proc Natl Acad Sci U S A. 93:3439–3443.

Montello D, Pick H. 1993. Integrating knowledge of vertically aligned large-scale spaces. Environ Behav. 25:457–484.

Morel A. 2007. Stereotactic atlas of the human thalamus and basal ganglia. New York:

Informa Healthcare USA. New York: Informa Healthcare USA.

Morey RD. 2008. Confidence intervals from normalized data: A correction to Cousineau (2005). Tutor Quant Methods Psychol. 4:61–64.

Morris RGM, Schenk F, Tweedie F, Jarrard LE. 1990. Ibotenate Lesions of Hippocampus and/or Subiculum: Dissociating Components of Allocentric Spatial Learning. Eur J Neurosci. 2:1016–1028.

Moser EI, Moser M-B, McNaughton BL. 2017. Spatial representation in the hippocampal formation: a history. Nat Neurosci. 20:1448–1464.

Muller RU, Bostock E, Taube JS, Kubie JL. 1994. On the directional firing properties of hippocampal place cells. J Neurosci. 14:7235–7251.

Muller RU, Kubie JL. 1987. The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. J Neurosci. 7:1951–1968.

Nardi D, Bingman VP. 2009a. Pigeon (Columba livia) encoding of a goal location: The relative importance of shape geometry and slope information. J Comp Psychol. 123:204–216.

Nardi D, Bingman VP. 2009b. Slope-based encoding of a goal location is unaffected by hippocampal lesions in homing pigeons (Columba livia). Behav Brain Res. 205:322–326.

Nau M, Navarro Schröder T, Bellmund JLS, Doeller CF. 2018. Hexadirectional coding of visual space in human entorhinal cortex. Nat Neurosci. 21:188–190.

Nili H, Wingfield C, Walther A, Su L, Marslen-Wilson W, Kriegeskorte N. 2014. A Toolbox for Representational Similarity Analysis. PLoS Comput Biol. 10:e1003553.

Nitz DA. 2011. Path shape impacts the extent of CA1 pattern recurrence both within and across environments. J Neurophysiol. 105:1815–1824.

Nolan CR, Vromen JMG, Cheung A, Baumann O. 2017. Evidence against the detectability of a hippocampal place code using functional magnetic resonance imaging. 1–22.

Nonyane BAS, Theobald CM. 2007. Design sequences for sensory studies: achieving balance for carry-over and position effects. Br J Math Stat Psychol. 60:339–349.

Norman KA, Polyn SM, Detre GJ, Haxby J V. 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn Sci. 10:424–430.

O'Keefe J, Burgess N. 1996. Geometric determinants of the place fields of hippocampal neurons. Nature.

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 University

Press.

O'Keefe J, Recce ML. 1993. Phase relationship between hippocampal place units and the EEG theta rhythm. Hippocampus. 3:317–330.

Oman CM. 2007. Spatial orientation and navigation in microgravity. In: Mast F,, Jancke L, editors. Spatial processing in navigation, imagery and perception. Journal of Neuroscience.

Op de Beeck HP. 2010. Against hyperacuity in brain reading : Spatial smoothing does not hurt multivariate fMRI analyses? Neuroimage. 49:1943–1948.

Orban GA. 2013. Which animal model for understanding human navigation in a three-dimensional world? Behav Brain Sci. 36:558–559.

Page HJII, Wilson JJ, Jeffery KJ. 2017. A dual-axis rotation rule for updating the head direction cell reference frame during movement in three dimensions. J Neurophysiol. 119:jn.00501.2017.

Pallis CA. 1955. Impaired identification of faces and places with agnosia for colors: Report of a case due to cerebral embolism. J Neurol Neurosurgery, andPsychiatry. 18:218–224.

Park S, Chun MM. 2009. Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. Neuroimage. 47:1747–1756.

Parvizi J, Kastner S. 2018. Promises and limitations of human intracranial electroencephalography. Nat Neurosci. 21:1–10.

Pengas G, Williams GB, Acosta-Cabronero J, Hong YT, Izquierdo-Garcia D, Fryer TD, Hodges JR, Nestor PJ. 2012. The relationship of topographical memory performance to regional neurodegeneration in Alzheimer's disease. Front Aging Neurosci. 4:1–10.

Poppenk J, Evensmoen HR, Moscovitch M, Nadel L. 2013. Long-axis specialization of the human hippocampus. Trends Cogn Sci. 17:230–240.

Proffitt DR, Bhalla M, Gossweiler R, Midgett J. 1995. Perceiving geographical slant. Psychon Bull Rev. 2:409–428.

Pruessner JC, Köhler S, Crane J, Pruessner M, Lord C, Byrne A, Kabani N, Collins DL, Evans AC. 2002. Volumetry of temporopolar, perirhinal, entorhinal and parahippocampal cortex from high-resolution MR images: considering the variability of the collateral sulcus. Cereb Cortex. 12:1342–1353.

Quirk GJ, Muller RU, Kubie JL. 1990. The firing of hippocampal place cells in the dark depends on the rat's recent experience. J Neurosci. 10:2008–2017.

Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I. 2005. Invariant visual representation by single neurons in the human brain. Nature. 435:1102–1107.

Qutob Z. 1963. The swimbladder of fishes as a pressure receptor. Arch Néerlandaises Zool. 15:1–67.

Reagh ZM, Yassa M a. 2014. Object and spatial mnemonic interference differentially engage lateral and medial entorhinal cortex in humans. Proc Natl Acad Sci. 111:E4264–E4273.

Renninger LW, Malik J. 2004. When is scene identification just texture recognition? Vision Res. 44:2301–2311.

Riecke BE, Sigurdarson S, Milne AP. 2012. Moving through virtual reality without moving? Cogn Process. 13:293–297.

Rotenberg A, Mayford M, Hawkins RD, Kandel ER, Muller RU. 1996. Mice expressing activated CaMKII lack low frequency LTP and do not form stable place cells in the CA1 region of the hippocampus. Cell. 87:1351–1361.

Savelli F, Yoganarasimha D, Knierim JJ. 2008. Influence of boundary removal on the spatial representations of the medial entorhinal cortex. Hippocampus. 18:1270–1282.

Schapiro AC, Turk-Browne NB, Norman KA, Botvinick MM. 2016. Statistical learning of temporal community structure in the hippocampus. Hippocampus. 26:3–8.

Schuck NW, Cai M, Wilson RC, Niv Y. 2016. Human orbitofrontal cortex represents a cognitive map of state space. Neuron. 91:1402–1412.

Shaffer DM, Flint M. 2011. Escalating slant: increasing physiological potential does not reduce slant overestimates. Psychol Sci. 22:209–211.

Shine JP, Valdés-Herrera JP, Hegarty M, Wolbers T. 2016. The human retrosplenial cortex and thalamus code head direction in a global reference frame. J Neurosci. 36:6371–6381.

Smith M Lou, Milner B. 1981. The role of the right hippocampus in the recall of spatial location. Neuropsychologia. 19:781–793.

Solstad T, Boccara CN, Kropff E, Moser M-B, Moser EI. 2008. Representation of geometric borders in the entorhinal cortex. Science. 322:1865–1868.

Solstad T, Moser EI, Einevoll GT. 2006. From grid cells to place cells: A mathematical model. Hippocampus. 16:1026–1031.

Spector RH. 1990. Chap 116 Visual fields. In: Walker H,, Hall W,, Hurst J, editors. Clinical methods: The history, physical, and laboratory examinations. 3rd edition.

Spiers HJ, Burgess N, Hartley T, Vargha-Khadem F, O'Keefe J. 2001. Bilateral hippocampal pathology impairs topographical and episodic memory but not visual pattern matching. Hippocampus. 11:715–725.

Spiers HJ, Hayman RMA, Jovalekic A, Marozzi E, Jeffery KJ. 2015. Place field repetition and purely local remapping in a multicompartment environment. Cereb

Cortex. 25:10–25.

Spiers HJ, Maguire EA. 2006. Thoughts, behaviour, and brain dynamics during navigation in the real world. Neuroimage. 31:1826–1840.

Stackman R, Taube J. 1998. Firing properties of rat lateral mammillary single units: head direction, head pitch, and angular head velocity. J Neurosci. 18:9020–9037.

Stackman RW, Golob EJ, Bassett JP, Taube JS. 2003. Passive transport disrupts directional path integration by rat head direction cells. J Neurophysiol. 90:2862–2874.

Stangl M, Shine J, Wolbers T. 2017. The GridCAT: A toolbox for automated analysis of human grid cell codes in fMRI. Front Neuroinform. 11:47.

Stark M. 1996. Impairment of an egocentric map of locations: Implications for perception and action. Cogn Neuropsychol. 13:481–524.

Steck SD, Mochnatzki HF, Mallot HA. 2003. The role of geographical slant in virtual environment navigation. In: Freksa C, Wilfried B, Habel C, Wender KF, editors. Spatial Cognition III. Berlin, Heidelberg: Springer Berlin Heidelberg. p. 62–76.

Stella F, Treves A. 2015. The self-organization of grid cells in 3D. Elife. 4:e05913.

Stensola H, Stensola T, Solstad T, FrØland K, Moser MB, Moser EI. 2012. The entorhinal grid map is discretized. Nature. 492:72–78.

Stensola T, Stensola H, Moser M-B, Moser EI. 2015. Shearing-induced asymmetry in entorhinal grid cells. Nature. 518:207–212.

Strange BA, Witter MP, Lein ES, Moser EI. 2014. Functional organization of the hippocampal longitudinal axis. Nat Rev Neurosci. 15:655–669.

Sulpizio V, Committeri G, Galati G. 2014. Distributed cognitive maps reflecting real distances between places and views in the human brain. Front Hum Neurosci. 8:716.

Takahashi N, Kawamura M, Shiota J, Kasahata N, Hirayama K. 1997. Pure topographic disorientation due to right retrosplenial lesion. Neurology. 49:464–469.

Taube J, Muller R, Ranck J. 1990. Head-direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. J Neurosci. 10:436–447.

Taube JS. 2007. The head direction signal: Origins and sensory-motor integration. Annu Rev Neurosci. 30:181–207.

Taube JS, Stackman RW, Calton JL, Oman CM. 2004. Rat head direction cell responses in zero-gravity parabolic flight. J Neurophysiol. 92:2887–2997.

Taube JS, Valerio S, Yoder RM. 2013. Is navigation in virtual reality with fMRI really navigation? J Cogn Neurosci. 25:1008–1019.

Taube JS, Wang SS, Kim SY, Frohardt RJ. 2013. Updating of the spatial reference frame of head direction cells in response to locomotion in the vertical plane. J Neurophysiol. 109:873–888.

Tavares RM, Mendelsohn A, Grossman Y, Williams CH, Shapiro M, Trope Y, Schiller D. 2015. A map for social navigation in the human brain. Neuron. 87:231–243.

Taylor GK, Holbrook RI, Burt de Perera T. 2010. Fractional rate of change of swim-bladder volume is reliably related to absolute depth during vertical displacements in teleost fish. J R Soc Interface. 7:1379–1382.

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.

Thibault G, Pasqualotto A, Vidal M, Droulez J, Berthoz A. 2013. How does horizontal and vertical navigation influence spatial memory of multifloored environments? Atten Percept Psychophys. 75:10–15.

Tolman EC. 1948. Cognitive maps in rats and men. Psychol Rev. 55:189–208.

Tse D, Langston RF, Kakeyama M, Bethus I, Spooner PA, Wood ER, Witter MP, Morris RGM. 2007. Schemas and memory consolidation. Science. 316:76–82.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage. 15:273–289.

Vann SD, Aggleton JP, Maguire EA. 2009. What does the retrosplenial cortex do? Nat Rev Neurosci. 10:792–802.

Vass LK, Copara MS, Seyal M, Shahlaie K, Farias ST, Shen PY, Ekstrom AD. 2016. Oscillations go the distance: Low-frequency human hippocampal oscillations code spatial distance in the absence of sensory cues during teleportation. Neuron. 89:1180–1186.

Vass LK, Epstein RA. 2013. Abstract representations of location and facing direction in the human brain. J Neurosci. 33:6133–6142.

Vidal M, Amorim M-AA, Berthoz A. 2004. Navigating in a virtual three-dimensional maze: how do egocentric and allocentric reference frames interact? Cogn brain Res. 19:244–258.

Vogt BA, Miller MW. 1983. Cortical connections between rat cingulate cortex and visual, motor, and postsubicular cortices. J Comp Neurol. 216:192–210.

von Gierke HE, Parker DE. 1994. Differences in otolith and abdominal viscera graviceptor for motion sickness body position dynamics : Implications and perceived body position. Aviat Sp Environ Med. 65:747–752.

Wallace DJ, Greenberg DS, Sawinski J, Rulla S, Notaro G, Kerr JND. 2013. Rats

maintain an overhead binocular field at the expense of constant fusion. Nature. 498:65–69.

Walther A, Nili H, Ejaz N, Alink A, Kriegeskorte N, Diedrichsen J. 2016. Reliability of dissimilarity measures for multi-voxel pattern analysis. Neuroimage. 137:188–200.

Wang RF, Simons DJ. 1999. Active and passive scene recognition across views. Cognition. 70:191–210.

Watrous AJ, Fried I, Ekstrom AD. 2011. Behavioral correlates of human hippocampal delta and theta oscillations during navigation. J Neurophysiol. 105:1747–1755.

Weiskopf N, Hutton C, Josephs O, Deichmann R. 2006. Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: a whole-brain analysis at 3 T and 1.5 T. Neuroimage. 33:493–504.

Wiener J, Schnee A, Mallot H. 2004. Navigation strategies in regionalized environments. Max Plank Inst Biol Cybern. Technical report No. TR – 121.

Wills TJ, Cacucci F. 2014. The development of the hippocampal neural representation of space. Curr Opin Neurobiol. 24:111–119.

Wilson JJ, Harding E, Fortier M, James B, Donnett M, Kerslake A, O'Leary A, Zhang N, Jeffery K, O'Leary A, Zhang N, Jeffery K. 2015. Spatial learning by mice in three dimensions. Behav Brain Res. 289:125–132.

Woollett K, Maguire EA. 2011. Acquiring "the knowledge" of London's layout drives structural brain changes. Curr Biol. 21:2109–2114.

Xiang JZ, Brown MW. 1998. Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. Neuropharmacology. 37:657–676.

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. 2011. Large-scale automated synthesis of human functional neuroimaging data. Nat Methods. 8:665–670.

Yartsev MM, Ulanovsky N. 2013. Representation of three-dimensional space in the hippocampus of flying bats. Science. 340:367–372.

Yartsev MM, Witter MP, Ulanovsky N. 2011. Grid cells without theta oscillations in the entorhinal cortex of bats. Nature. 479:103–107.

Yoon K, Lewallen S, Kinkhabwala AA, Tank DW, Fiete IR. 2016. Grid cell responses in 1D environments assessed as slices through a 2D lattice. Neuron. 89:1086–1099.

Yushkevich PA, Amaral RSC, Augustinack JC, Bender AR, Bernstein JD, Boccardi M, Bocchetta M, Burggren AC, Carr VA, Chakravarty MM, Chételat G, Daugherty AM, Davachi L, Ding SL, Ekstrom A, Geerlings MI, Hassan A, Huang Y, Iglesias JE, La Joie R, Kerchner GA, LaRocque KF, Libby LA, Malykhin N, Mueller SG, Olsen RK, Palombo DJ, Parekh MB, Pluta JB, Preston AR, Pruessner JC, Ranganath C, Raz

N, Schlichting ML, Schoemaker D, Singh S, Stark CEL, Suthana N, Tompary A, Turowski MM, Van Leemput K, Wagner AD, Wang L, Winterburn JL, Wisse LEM, Yassa MA, Zeineh MM. 2015. Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in in vivo MRI: Towards a harmonized segmentation protocol. Neuroimage. 111:526–541.

Zeidman P, Maguire EA. 2016. Anterior hippocampus: the anatomy of perception, imagination and episodic memory. Nat Rev Neurosci. 17:173–182.

Zwergal A, Schöberl F, Xiong G, Pradhan C, Covic A, Werner P, Trapp C, Bartenstein P, la Fougère C, Jahn K, Dieterich M, Brandt T. 2016. Anisotropy of human horizontal and vertical navigation in real space: Behavioral and PET correlates. Cereb Cortex. 26:4392–4404.