Navigating in a 3D world

Kathryn J. Jeffery¹, Aleksandar Jovalekic ²*, Madeleine Verriotis ³*, and Robin Hayman⁴

¹ (Corresponding author)
Institute of Behavioural Neuroscience
Dept. of Cognitive, Perceptual and Brain Sciences
Division of Psychology & Language Sciences
University College London
26 Bedford Way
London WC1H 0AP
k.jeffery@ucl.ac.uk
www.ucl.ac.uk/jefferylab/

² Institute of Behavioural Neuroscience
Dept. of Cognitive, Perceptual and Brain Sciences
Division of Psychology & Language Sciences
University College London
26 Bedford Way
London WC1H 0AP
and
Axona Ltd
Unit 4U Long Spring, Porters Wood
St. Albans AL3 6EN UK
a.jovalekic@ucl.ac.uk

³ Institute of Behavioural Neuroscience
Dept. of Cognitive, Perceptual and Brain Sciences
Division of Psychology & Language Sciences
University College London
Abstract: The study of spatial cognition has provided considerable insights into how animals (including humans) navigate on the horizontal plane. However, the real world is three-dimensional, having a complex topography including both horizontal and vertical features, which presents additional challenges for representation and navigation. The present article reviews the emerging behavioral and neurobiological literature on spatial cognition in non-horizontal environments. We suggest that three-dimensional spaces are represented in a quasi-planar fashion, with space in the plane of locomotion being computed separately and represented differently from space in the orthogonal axis: a representational structure we have termed “bicoded”. We argue that the mammalian spatial representation in surface–travelling animals comprises a mosaic of these locally planar fragments, rather than a fully-integrated volumetric map. More generally, this may be true even for species that can move freely in all three dimensions such as birds and fish. We outline the evidence supporting this view, together with the adaptive advantages of such a scheme.

Keywords: ethology, grid cells, head direction cells, hippocampus, navigation, neural encoding, place cells, spatial cognition, three-dimensional
This page left intentionally blank.

Please scroll down.
Navigating in a 3D world
Kathryn J Jeffery¹, Aleksandar Jovalekic²*, Madeleine Verriotis³*, Robin Hayman⁴

¹ (Corresponding author)
Institute of Behavioural Neuroscience
Dept. of Cognitive, Perceptual and Brain Sciences
Division of Psychology & Language Sciences
University College London
26 Bedford Way
London WC1H 0AP
k.jeffery@ucl.ac.uk
www.ucl.ac.uk/jefferylab/

²
Institute of Behavioural Neuroscience
Dept. of Cognitive, Perceptual and Brain Sciences
Division of Psychology & Language Sciences
University College London
26 Bedford Way
London WC1H 0AP
and
Axona Ltd
Unit 4U Long Spring, Porters Wood
St. Albans AL3 6EN
UK
a.jovalekic@ucl.ac.uk

³
Institute of Behavioural Neuroscience
Dept. of Cognitive, Perceptual and Brain Sciences
Division of Psychology & Language Sciences
University College London
26 Bedford Way
London WC1H 0AP
m.verriotis@ucl.ac.uk

⁴
Institute of Behavioural Neuroscience
Dept. of Cognitive, Perceptual and Brain Sciences
Division of Psychology & Language Sciences
University College London
26 Bedford Way
London WC1H 0AP
r.hayman@ucl.ac.uk

*These authors contributed equally to the article
Key words ethology, grid cells, head direction cells, hippocampus, navigation, neural encoding, place cells, spatial cognition, three-dimensional

Word counts
Short abstract: 99
Long abstract: 159
Main text: 15266
References: 2302
Entire text: 18014
Short abstract

Spatial cognition research has provided considerable insights into how navigation occurs on the horizontal plane. However, the real world is three-dimensional, presenting additional challenges for representation and navigation. This article reviews emerging behavioral and neurobiological findings on spatial cognition in non-horizontal environments and suggests that in surface-travelling animals, and perhaps even more generally, three-dimensional spaces are represented in a quasi-planar “bicoded” fashion, with space in the plane of locomotion being represented differently from space in the orthogonal axis. We suggest that the mammalian spatial representation comprises a mosaic of these locally planar fragments, rather than a fully-integrated, volumetric map.

Long abstract

The study of spatial cognition has provided considerable insights into how animals (including humans) navigate on the horizontal plane. However, the real world is three-dimensional, having a complex topography including both horizontal and vertical features, which presents additional challenges for representation and navigation. The present article reviews the emerging behavioral and neurobiological literature on spatial cognition in non-horizontal environments. We suggest that three-dimensional spaces are represented in a quasi-planar fashion, with space in the plane of locomotion being computed separately and represented differently from space in the orthogonal axis: a representational structure we have termed “bicoded”. We argue that the mammalian spatial representation in surface –
travelling animals comprises a mosaic of these locally planar fragments, rather than a fully-integrated volumetric map. More generally, this may be true even for species that can move freely in all three dimensions such as birds and fish. We outline the evidence supporting this view, together with the adaptive advantages of such a scheme.
1. Introduction

Behavioural and neurobiological studies of spatial cognition have provided considerable insight into how animals and humans navigate through the world, establishing that a variety of strategies, subserved by different neural systems, can be used in order to plan and execute trajectories through large-scale, navigable space. This work, which has mostly been done in simplified, laboratory-based environments, has defined a number of sub-processes such as landmark recognition, heading determination, odometry (distance-measuring) and context recognition, all of which interact in the construction and use of spatial representations. The neural circuitry underlying these processes has in many cases been identified, and it now appears that the basics of navigation are reasonably well understood.

The real world, however, is neither simple nor two-dimensional, and the addition of vertical space to the other two horizontal dimensions adds a number of new problems for a navigator to solve. For one thing, moving against gravity imposes additional energy costs. Also, moving in a volumetric space is computationally complicated, because of the increased size of the representation needed, and also because of the fact that rotations in three dimensions interact. It remains an open question how the brain has solved these problems, and whether the same principles of spatial encoding operate in all three dimensions or whether the vertical dimension is treated differently.
Studies of how animals and humans navigate in environments with a vertical component are slowly proliferating, and it is timely to review the gathering evidence and offer a theoretical interpretation of the findings to date. In the first part of the article we review the current literature on animal and human orientation and locomotion in 3D spaces, highlighting key factors that contribute to and influence navigation in such spaces. The second part summarises neurobiological studies of the encoding of 3D space, and the third and final part integrates the experimental evidence to put forward a hypothesis concerning 3D spatial encoding – the bicoded model – in which we propose that in encoding 3D spaces, the mammalian brain constructs mosaics of connected surface-referenced maps. The advantages and limitations of such a quasi-planar encoding scheme are explored.

2. Theoretical considerations

Before reviewing the psychological and neurobiological aspects of 3D navigation it is useful to establish a theoretical framework for the discussion, which will be developed in more detail in the last part of this article. Study of navigation needs to be distinguished from the study of spatial perception per se, in which a subject may or may not be moving through the space. Additionally, it is also important to distinguish between space that is encoded egocentrically (i.e., relative to the subject) vs. space that is encoded allocentrically (independently of the subject – that is, relative to the world). The focus of the present article is on the allocentric encoding of the space that is being moved through, which probably relies on different (albeit interconnected) neural systems from the encoding of egocentric space.
An animal that is trying to navigate in 3D space needs to know three things – how it is positioned, how it is oriented, and how it is moving – and all these things require a reference frame: a world-anchored space-defining framework with respect to which position, orientation and movement can be specified. There also needs to be some kind of metric co-ordinate system – that is, a signal that specifies distances and directions with respect to this reference frame. The core question for animal navigation therefore concerns what the reference frame might be, and how the co-ordinate system encodes distances and directions within this frame.

The majority of studies of navigation, at least within neurobiology, have up until now taken place in laboratory settings using restricted environments such as mazes, which are characterized both by being planar (i.e., two-dimensional), and by being horizontal. The real world differs in both of these respects. First, it is often not planar – air and water, for example, allow movement in any direction and so are volumetric. In this article, therefore, we will divide 3D environments into those that are locally planar surfaces, which allow movement only in a direction tangential to the surface, and those that are volumetric spaces (air, water, space and virtual reality) in which movement (or virtual movement) in any direction is unconstrained. Second, even if real environments might be (locally) planar, the planes may not necessarily be horizontal. These factors are important in thinking about how subjects encode and navigate through complex spaces.

What does the addition of the vertical dimension add to the problem of navigation? At first glance, not much – the vertical dimension is a direction one can move in, like any other, and it could be that the principles that have been elucidated for 2D
navigation extend straightforwardly to three. In fact, however, the vertical
dimension makes the problem of navigation considerably more complex, for a
number of reasons. First, the space to be represented is much larger, because a
volume is larger than a plane by a power of 3/2. Second, the directional information
to be encoded is more complex because there are three planes of orientation instead
of just one. Furthermore, rotations in orthogonal planes interact such that sequences
of rotations have a different outcome depending on the order in which they are
executed (i.e., they are non-commutative), which adds to processing complexity.
Third, the vertical dimension is characterized by gravity, and by consequences of
gravity such as hydrostatic or atmospheric pressure, which add both information
and effort to the computations involved. And finally, there are navigationally
relevant cues available for the horizontal plane that are absent in the vertical: for
example, the position of the sun and stars, or the geomagnetic field.

The internal spatial representation that the brain constructs for use in large-scale
navigation is often referred to as a “cognitive map” (O'Keefe & Nadel 1978). The
cognitive map is a supramodal representation that arises from the collective activity
of neurons in a network of brain areas that process incoming sensory information
and abstract higher order spatial characteristics such as distance and direction. In
thinking about whether cognitive maps are three-dimensional, it is useful to
consider the various 3D mapping schemes that are possible. Fig. 1 shows three
different hypothetical ways of mapping spaces, which differ in the degree to which
they incorporate a representation of the vertical dimension. In Fig. 1 A, the metric
fabric of the map – the grid – follows the topography but does not itself contain any
information about it: the map is locally flat. Such a surface is called, in mathematics, a manifold, which is a space that is locally Euclidean – that is, for any point in the space, its immediate neighbours will be related to it by the usual rules of geometry, even if more distant points are related in a non-Euclidean way. For example, the surface of a sphere is a manifold, because a very small triangle on its surface will be Euclidean (its interior angles add up to 180 degrees) but a large triangle will be Riemannian (the sum of its interior angles exceeds 180 degrees). Given a small enough navigator and large enough surface undulations, the ground is therefore a manifold. For the purposes of the discussion that follows, we will call a map that follows the topography and is locally Euclidean a surface map. A surface map in its simplest form does not, for a given location, encode information about whether the location is on a hill or in a valley. Although such a map accurately represents the surface distance a navigator would need to travel over undulating terrain, because it does not contain information about height (or “elevation”) above a reference plane it would fail to allow calculation of the extra effort involved in moving over the undulations, and nor would the calculation of shortcuts be accurate.

\footnote{Note that some other part of the animal’s brain may be able to deduce, via postural and kinaesthetic mechanisms, the slope of the terrain. The point is that this information is not integrated into the metric fabric of the map itself.}
Figure 1: Different kinds of 3D encoding. (A) A surface map, with no information about elevation; (B) An extracted flat map in which the horizontal co-ordinates have been generated, by trigonometric inference, from hilly terrain; (C) A bicoded map, which is metric in the horizontal dimension and uses a non-metric scale in the vertical (upper panel = bird’s eye view of the terrain, lower panel = cross-section at the level of the dotted line), or (D) A volumetric map, which is metric in all three dimensions, like this map of dark matter in the universe.

A variant of the surface map is the extracted flat map shown in Fig. 1 B – here, distances and directions are encoded relative to the horizontal plane, rather than to the terrain surface. For a flat environment, a surface map equates to an extracted flat map. For hilly terrain, in order to determine its position in x-y co-ordinates the navigator needs to process the slope in order to determine, trigonometrically, the equivalent horizontal distance. To generate such a map it is therefore necessary to have some processing of information from the vertical dimension (i.e., slope with respect to position and direction) and so this is a more complex process than for a purely surface map. There is evidence, discussed later, that some insects can in fact
extract horizontal distance from hilly terrain, although whether this can be used to generate a map remains moot.

In the map shown in Fig. 1 C, the representation now does include a specific representation of elevation but the encoding method is different from that of the horizontal distances, and hence we have termed this a bicoded map. In a bicoded map, vertical distance is represented by some non-spatial variable, in this case color. If provided with information about x-y co-ordinates together with the non-spatial variable, an agent could work out where they were in 3D space, but would not have metric information (that is, quantitative information about distance) about the vertical component of this position unless they also had a key (some way of mapping the non-spatial variable to distances). Lacking a key, a navigator could still know whether they were on a hill or in a valley, but could not calculate efficient shortcuts because they could not make the trigonometric calculations that included the vertical component. Such a map could be used for long-range navigation over undulating terrain, but determining the most energy-efficient route would be complicated because it would require chaining together segments having different gradients (and thus energetic costs), as opposed to simply computing a shortest-distance route. Although the representational scheme described here is an artificial one used by human geographers, later on we will make the point that an analogous scheme – use of a non-metric variable to encode one of the dimensions in a 3D map – may operate in the neural systems that support spatial representation.

Finally, the volumetric map shown in Fig. 1 D could be used for navigation in a volumetric space. Here, all three dimensions use the same metric encoding scheme,
and the shortest distance between any two points can be calculated straightforwardly using trigonometry (or some neural equivalent), with (on Earth) elevation being simply the distance in the dimension perpendicular to gravity. However, calculating energy costs over an undulating surface would still be complicated, for the same reasons as above – energy expenditure would have to be continuously integrated over the course of the planned journey in order to allow comparisons of alternate routes.

The question concerning the brain’s map of space, then, is: what encoding scheme does it use for navigation over either undulating surfaces or in volumetric spaces? Is the vertical dimension encoded in the same way as the two horizontal ones, as in the volumetric map, is it ignored entirely, as in a surface map, or is it encoded, but in a different way from the horizontal dimensions, as in the bicoded map? These questions motivate the following analysis of experimental data, and will be returned to in the final section, where it will be suggested that a bicoded map is the most likely encoding scheme, at least in mammals, and perhaps even for animals that can swim or fly.

3. **Behavioural studies in three dimensions**

A key question to be answered concerning the representation of 3D space is how information from the vertical dimension is incorporated into the spatial representation. We will therefore now review experimental studies that have explored the use of information from the vertical dimension in animals and humans, beginning with an examination of how cues from the vertical dimension may be
processed, before moving on to how these cues may be used in self-localisation and navigation. The discussion is organised as follows, based on a logical progression through complexity of the representation of the third (usually vertical) dimension:

- Processing of verticality cues
- Navigation on a tilted surface
- Navigation in multilayer environments
- Navigation in volumetric space

3.1 Processing of verticality cues

Processing of cues arising from the vertical dimension is a faculty possessed by members of most phyla, and it takes many forms. One of the simplest is detection of the vertical axis itself, using a variety of mechanisms that mostly rely on gravity, or on cues related to gravity, such as hydrostatic pressure or light gradients in the ocean. This information provides a potential orienting cue which is useful not only in static situations but also for navigation in 3D space. Additionally, within the vertical dimension itself, gravity differentiates down from up and thus polarizes the vertical axis: consequently, animals are always oriented relative to this axis, even though they can become disoriented relative to the horizontal axes. The importance of the gravity signal as an orienting cue is evident in reduced or zero-gravity environments, which appear to be highly disorienting for humans: astronauts report difficulty in orienting themselves in three-dimensional space when floating in an unfamiliar orientation (Oman 2007).
The vertical axis, once determined, can then be used as a reference frame for the encoding of metric information in the domains of both direction and distance. The technical terms for rotations relative to the vertical axis are pitch, roll and yaw, corresponding to rotations in the sagittal, coronal and transverse planes, respectively (Fig. 2A). Note that there is potential for confusion since these rotational terms tend to be used in an egocentric (body-centered) frame of reference, in which the planes of rotation move with the subject. The static terms used to describe the resulting orientation of the subject (or “attitude”) are elevation, bank and azimuth, which describe (1) the angle that the front-back (antero-posterior) axis of the subject makes with Earth-horizontal (elevation), (2) the angle that the transverse axis of the subject makes with Earth-horizontal (bank), and (3) the angle that the anteroposterior axis of the subject makes with some reference direction, usually geographical North (azimuth). Tilt is the angle between the cranial axis and the gravitational vertical.

Figure 2: (A) Terminology describing rotations in the three cardinal planes, and (B) terminology describing static orientation (or “attitude”) in each of the three planes.
The other metric category, in addition to direction, is distance, which in the vertical dimension is called height or (more precisely) elevation when it refers to distance from a reference plane (e.g., ground or water level) in an upwards direction, and depth when it refers to distance in a downwards direction. Determination of elevation/depth is a special case of the general problem of odometry (distance-measuring). There is scant information about the extent to which animals encode information about elevation. A potential source of such information could come from studies of aquatic animals, which are exposed to particularly salient cues from the vertical dimension due to the hydrostatic pressure gradient. Many fish are able to detect this gradient via their swim bladders, which are air-filled sacs that can be inflated or deflated to allow the fish to maintain neutral buoyancy. Holbrook & Burt de Perera (2011) have argued that changes in vertical position would also, for a constant gas volume, change the pressure in the bladder and thus provide a potential signal for relative change in height. Whether fish can actually use this information for vertical self-localization remains to be determined however.

The energy costs of moving vertically in water also add navigationally relevant information. These costs can vary between species (Webber et al. 2000): thus, we might expect that the vertical dimension would be more salient in some aquatic animals than in others. Thermal gradients, which are larger in the vertical dimension, could also provide extra information concerning depth, as well as imposing differential energetic costs in maintaining thermal homeostasis at different depths (Brill et al. 2012; Carey F.G. 1992)
Figure 3: The experiment by Holbrook and Burt de Perera (2009). (A) Line diagram of the choice maze (taken from the original article, with permission), in which fish were released into a tunnel and had to choose one of the exit
arms to obtain a reward. (B) The conflict experiment of that study, in which the fish were trained with the arms at an angle, so that one pointed up and to the left, and the other down and to the right. During testing, the arms were rotated about the axis of the entrance tunnel so that one had the same vertical orientation but a different left-right position, and one the same horizontal (left-right) position but a different vertical level. (C) Analysis of the choices showed that the fish greatly preferred to choose the arm having the same vertical level. Redrawn from the original.

Interestingly, while fish can process information about depth, Holbrook & Burt de Perera (2009) found that they appear to separate this from the processing of horizontal information. While banded tetras (*Astyanax fasciatus*) learnt the vertical and horizontal components of a trajectory equally fast, the fish tended to use the two components independently, suggesting a separation either during learning, storage, or recall, or at the time of use of the information. When the two dimensions were in conflict, fish preferred the vertical dimension (Fig. 3), possibly due to the hydrostatic pressure gradient.

Several behavioural studies support the use of elevation information in judging the vertical position of a goal. For instance, rufous hummingbirds (*Selasphorous rufus*) can distinguish flowers on the basis of their relative heights above ground (Henderson et al. 2001; Henderson et al. 2006), and stingless bees (genus *Meliponae*) can communicate the elevation as well as the distance and direction to a food source (Nieh et al. 2003; Nieh & Roubik 1998). Interestingly, honeybees (*Apis mellifera*) do not appear to communicate elevation (Dacke & Srinivasan 2007; Esch & Burns 1995). Rats are also able to use elevation information in goal localisation, and – like the fish in the Holbrook and Burt de Perera study – appear to separate
horizontal and vertical components when they do so. For example, (Grobéty & Schenk (1992a) found that in a lattice maze (Fig. 4A), rats learned to locate the elevation of a food item more quickly than they learned its horizontal coordinates, suggesting perhaps that elevation provides an additional, more salient spatial cue, perhaps because of its added energy cost, or because of arousal/anxiety. By contrast, Jovalekic et al. (2011) found somewhat the opposite. Rats were trained on a vertical climbing wall, studded with protruding pegs to provide footholds (Fig. 4B), to shuttle back and forth between a start and goal location: when a barrier was inserted into the maze, to force a detour, the rats overwhelmingly preferred, in both directions, the route in which the horizontal leg occurred first and the vertical leg second. Similarly, on a three-dimensional lattice similar to the one used by Grobéty and Schenk, the rats preferred the shallow-steep route to the steep-shallow route around a barrier. However, the task differed from that faced by Grobéty and Schenk’s rats in that both the starting location and the goal were well learned by the time the obstacle was introduced. It may be that elevation potentiates initial learning but that in a well-learned situation the animals prefer to solve the horizontal component first, perhaps due to an inclination to delay the more costly part of a journey.

Humans can estimate their own elevation, but with only a modest degree of accuracy. Garling et al. (1990) found that subjects were able to judge, from memory, which one of pairs of familiar landmarks in their hometown has a higher elevation, indicating that information about elevation is both observed and stored in the brain. It is not clear how this information was acquired, although there were
suggestions that the process may be based partly on encoding heuristics, such as encoding the steepest and thus most salient portions of routes first, as opposed to precise metric information. Notwithstanding the salience of the vertical dimension, processing of the vertical dimension in humans appears poor (Montello & Pick 1993; Pick & Rieser 1982; Tlauka et al. 2007).

**Figure 4** Rodent studies of vertical processing. (A) Schematic of the lattice maze of Grobèty and Schenk (1992a), together with a close-up photograph of a rat in the maze (inset). (B) Schematic and inset photograph of the pegboard maze used in the experiments of Jovalekic et al. (2011) and, as described later by Hayman et al. (2011). In the Jovalekic et al. experiment, rats learned to shuttle directly back and forth between start and goal locations as indicated in the diagram on the right. On test trials, a barrier (black bar) was inserted across the usual route, forcing threats to detour. The great majority preferred the routes indicated by the arrows, which had the horizontal leg first and the vertical leg second.
As with the fish and rodent studies above, a dissociation between vertical and horizontal was seen in a human study, in which landmarks arranged in a 3 x 3 grid within a building were grouped by most participants into (vertical) column representations, and by fewer participants into (horizontal) row representations (Buechner et al. 2007). Most participants’ navigation strategies were consistent with their mental representations; as a result, when locating the given target landmark, a greater proportion of participants chose a horizontal-vertical route. The authors suggested that these results might have been reversed in a building that was taller than it was wide, unlike the building in their study. However, given the results from the rats of Jovalekic et al. (2011), in which the climbing wall dimensions were equal, it is also possible that this behaviour reflects an innate preference for horizontal-first routes.

In conclusion, then, studies of both human and non-human animals find that there is clear evidence of the use of elevation information in spatial computations, and also for differential processing of information from vertical vs. horizontal dimensions. The processing differences between the two dimensions may result from differential experience, differential costs associated with the respective dimensions, or differences in how these dimensions are encoded neurally.

Next, we turn to the issue of how the change in height with respect to change in distance – that is, the slope of the terrain – may be used in spatial cognition and navigation. This is a more complex computation than simple elevation, but its value lies in the potential for adding energy and effort to the utility calculations of a journey.
3.2 *Navigation on a sloping surface*

Surfaces extending into the vertical dimension locally form either slopes or vertical surfaces, and globally form undulating or crevassed terrain. Slopes have the potential to provide several kinds of information to a navigating animal (Restat et al. 2004), acting for instance as a landmark (e.g. a hill visible from a distance) or as a local organising feature of the environment providing directional information. Additionally, as described earlier, the added vertical component afforded by slopes is both salient and costly, as upwards locomotion requires greater effort than horizontal locomotion, and an extremely steep slope could be hazardous. As a result, it is reasonable to assume that slopes might influence navigation in several ways.

A number of invertebrates, including desert ants (*Cataglyphis fortis*) and fiddler crabs (*Uca pugilator*) appear able to detect slopes, sometimes with a high degree of accuracy (at least 12 degrees for the desert ant (Wintergerst & Ronacher 2012). It is still unknown how these animals perform their slope-detection, which in the ant does not appear to rely on energy consumption (Schäfer & Wehner 1993), sky polarization cues (Hess et al. 2009) or proprioceptive cues from joint angles (Seidl & Wehner 2008). The degree to which slope is used in ant navigation has been under considerable investigation in recent years, making use of the ability of these animals to “home” (find their way back to the nest) using path integration-based navigation. Path integration (see Wehner 2003; Walls & Layne 2009; Etienne and Jeffery 2004) is the process by which position is constantly updated during an excursion away from a starting point such as a nest. Foraging ants will home after
finding food, and the direction and distance of their journey provides a convenient read-out of the homing vector operative at the start of their homeward journey. Wohlgemuth et al. (2001) investigated homing in desert ants and found that they could compensate for undulating terrain on an outbound journey by travelling the horizontally equivalent distance back to the nest across flat terrain, indicating that they had processed the undulations and could extract the corresponding ground distance. Likewise, fiddler crabs searched for their home burrow at the appropriate ground distance as opposed to the actual travelled distance (Walls & Layne 2009). Whether such encoding of slope is relative to Earth-horizontal (i.e., the plane defined by gravity), or whether it simply relates the two surfaces (the undulating outbound one and the flat return one) remains to be determined. However, despite this ability to extract ground-distance, Grah et al. (2007) found that Cataglyphis do not appear to encode specific aspects of the undulating trajectory in their representations of the journey, since ants trained on outbound sloping surfaces will select homing routes that have slopes even when these are inappropriate (e.g., if the two outbound slopes had cancelled out, leaving the ant at the same elevation as the nest with no need to ascend or descend). Thus, ants do not appear to incorporate elevation into their route memory – their internal maps seem essentially flat.

The ability of humans to estimate the slope of surfaces has been the subject of considerable investigation. Gibson & Cornsweet (1952) noted that geographical slant, which is the angle a sloping surface makes with the horizontal plane, is to be distinguished from optical slant, which is the angle the surface makes with the gaze direction (Fig. 5A). This is important because a surface appears to an observer to
possess a relatively constant slope regardless of viewing distance, even though the angle the slope makes with the eye changes. The observer must somehow be able to compensate for the varying gaze angle in order to maintain perceptual slope constancy, presumably by using information about the angle of the head and eye with respect to the body. However, despite perceptual slope constancy, perceptual slope accuracy is highly impaired – people tend to greatly overestimate the steepness of hills (Kammann 1967), typically judging a 5 degree hill to be 20 degrees, and a 30 degree hill to be almost 60 degrees (Proffitt et al. 1995). This has been linked to the well-established tendency of people to underestimate depth in the direction of gaze (Loomis et al. 1992) – if subjects perceive distances along the line of sight to be foreshortened then they would perceive the horizontal component of the slope to be less than it really is for a given vertical distance, and thus the gradient to be steeper than it is. However, this explanation fails to account for why subjects looking downhill also perceive the hill as too steep (Proffitt et al. 1995). Furthermore, if observers step back from the edge of a downwards slope, there is a failure of slope constancy as well as accuracy: they see it as steeper than if they are positioned right at the edge (Li & Durgin 2009). Durgin and Li have suggested that all these perceptual distortions can collectively be explained by so-called scale expansion (Li & Durgin 2010; Durgin & Li 2011): the tendency to overestimate angles (both gaze declination and optical slant) in central vision by a factor of about 1.5 (Fig. 5B). While this expansion may be beneficial for processing efficiency, it raises questions about how such distortions may affect navigational accuracy.
**Figure 5:** Visual slant estimation. (A) Slope constancy: a subject viewing a horizontal surface with a shallow gaze ($\beta_1$) vs. a steep gaze ($\beta_2$) perceives the slope of the ground to be the same (i.e., zero) in both cases, despite the optical slant ($\beta$) being different in the two cases. In order to maintain constant slope perception, the nervous system must therefore have some mechanism of compensating for gaze angle. (B) The scale expansion hypothesis: An observer misperceives both the declination of gaze and changes in optical slant with a gain of 1.5, so that all angles are exaggerated, but constancy of (exaggerated) slant is locally maintained. Increasing distance compression along the line of sight additionally contributes to slant exaggeration (Li & Durgin, 2010).
Other, non-visual distortions of slope perception may occur too. Proffitt and colleagues have suggested that there is an effect on slope perception of action-related factors such as whether people are wearing a heavy backpack or are otherwise fatigued, or elderly (Bhalla & Proffitt 1999), whether the subject is at the top or bottom of the hill (Stefanucci et al. 2005), their level of anxiety (Stefanucci et al. 2008) or mood (Riener et al. 2011) or even level of social support (Schnall et al. 2012). However, the interpretation of many of these observations has been challenged (Durgin et al., in press).

Creem & Proffitt (1998; 2001) have argued that visuomotor perceptions of slope are dissociable from explicit awareness. By this account, slope perception comprises two components: an accurate and stable visuomotor representation and an inaccurate and malleable conscious representation, which may be why subjects who appear to have misperceived slopes according to their verbal reports nevertheless tailor their actions to them appropriately. Durgin et al (Durgin et al. 2011) have challenged this view, suggesting that experimental probes of visuomotor perception are themselves inaccurate. The question of whether there are two parallel systems for slope perception thus remains open.

How does slope perception affect navigation? One mechanism appears to be by potentiating spatial learning. For instance, tilting some of the arms in radial arm mazes improves performance in working memory tasks in rats (Brown & Lesniak-Karpiak 1993; Grobéty & Schenk 1992b). Similarly, experiments in rodents involving conical hills emerging from a flat arena that rats navigated in darkness found that the presence of the slopes facilitated learning and accuracy during
navigation (Moghaddam et al. 1996; Fig. 6B). In these experiments, steeper slopes enhanced navigation to a greater extent than shallower slopes. Pigeons walking in a tilted arena were able to use the 20 degree slope in order to locate a goal corner; furthermore they preferred to use slope information when slope and geometry cues were placed in conflict (Nardi et al. 2010; Nardi & Bingman 2009). In these examples, it is likely that slope provided compass-like directional information as well as acting as a local landmark with which to distinguish between locations in an environment. In humans, similar results were reported in a virtual reality task involving navigational tasks within a virtual town with eight locations, in which participants navigated using a bicycle simulator (Restat et al. 2004; Steck et al. 2003; Fig. 6D). Participants experiencing the sloped version of the virtual town (4 degree slant) made fewer pointing and navigation errors than those experiencing the flat version of the virtual town, indicating that humans can also use slope to orient with greater accuracy, although interestingly, this appears to be more true for males than for females (Nardi et al. 2011). When slope and geometric cues were placed in conflict, geometric information was evaluated as more relevant (Kelly 2011). This is the opposite of the finding reported in pigeons discussed above. There are of course a number of reasons why pigeons and humans may differ in their navigational strategies, but one possibility is that different strategies are applied by animals that habitually travel through volumetric space rather than travel over a surface, in which geometric cues are arguably more salient.
The presence of sloping terrain adds complexity to the navigational calculations required in order to minimise energy expenditure. A study investigating human participants’ ability to find the shortest routes linking multiple destinations (the so-called travelling salesman problem, or “TSP”) in natural settings that were either flat or undulating found that performance (as measured by total distance travelled) was “better” in the flat condition (Phillips et al. 2009). This was perhaps because
participants avoided straight-line trajectories in the undulating terrain, and in so doing increased travel distance, to minimise hill-climbing. However, it is not known to what extent they minimized energy expenditure. Interestingly, Grobéty & Schenk (1992a) found that rats on the vertical plane made a far greater number of horizontal movements, partly because they made vertical translations by moving in a stair pattern rather than directly up or down, thus minimising the added effort of direct vertical translations. This reflects the findings of the TSP in the undulating environments mentioned above (Phillips et al., 2010) and is in line with the reported horizontal preferences in rats (Jovalekic et al., 2011).

In summary, then, the presence of sloping terrain not only adds effort and distance to the animal’s trajectory, but also adds orienting information. The extent to which slope is explicitly incorporated into the metric fabric of the cognitive map, as opposed to merely embellishing the map, is as-yet unknown.

3.3 Navigation in multilayer environments

Multilayer environments such as trees, burrow systems or buildings are those in which Earth-horizontal x-y co-ordinates recur as the animal explores, due to overlapping surfaces that are stacked along the vertical dimension. They are thus, conceptually, intermediate between planar and volumetric environments. Their theoretical importance lies in the potential for confusion, if – as we will argue later – the brain prefers to use a plane rather than a volume as its basic metric reference.

Most behavioural studies of spatial processing in multilayer environments have involved humans, and conversely, most human studies of 3D spatial processing
have involved multilayer environments (with the exception of studies in microgravity, discussed below). The core question concerning multilayer environments is the degree to which the layers are represented independently vs. treated as parts within an integrated whole. One way to explore this issue is to test whether subjects are able to path integrate across the layers. Path integration, introduced earlier, consists of self-motion tracking using visual, motor and sensory-flow cues, in order to continuously update representations of current location and/or a homing vector (for a review, see Etienne & Jeffery, 2004). An early study of path integration in three dimensions explored whether mice could track their position across independent layers of an environment (Bardunias & Jander 2000). The mice were trained to shuttle between their nest-box on one horizontal plane and a goal location on a parallel plane directly above the first, by climbing on a wire mesh cylinder joining the two planes. When the entire apparatus was rotated by 90 degrees in darkness while the animals were climbing the vertical wire mesh cylinder, all mice compensated for this rotation by changing direction on the upper plane, thus reaching the correct goal location. This suggests that they had perceived and used the angle of passive rotation and mapped it onto the overhead space. A second control experiment confirmed that they were not using distal cues to accomplish this, and thus suggested that they were able to rely on internal path integration mechanisms in order to remain oriented in three dimensions (Bardunias & Jander 2000). However, it is not necessarily the case that the mice calculated a three-dimensional vector to the goal, because in this environment the task of orienting in three dimensions could be simplified by maintaining a sense of direction in the horizontal dimension while navigating vertically. This is generally
true of multilayer environments, and so to show true integration, it is necessary to show that subjects incorporate vertical angular information into their encoding.

Montello & Pick (1993) demonstrated that human subjects who learned two separate and overlapping routes in a multilevel building could subsequently integrate the relationship of these two routes, as evidenced by their ability to point reasonably accurately from a location on one route directly to a location on the other. However, it was also established that pointing between these vertically aligned spaces was less accurate and slower than performances within a floor. Humans locating a goal in a multi-storey building preferred to solve the vertical component before the horizontal one, a strategy that led to shorter navigation routes and times (Holscher et al. 2006). In essence, rats and humans appear in such tasks to be reducing a three-dimensional task to a one-dimensional (vertical) followed by a two-dimensional (horizontal) task. However, their ability to compute a direct shortest route was obviously constrained by the availability of connecting ports (stairways) between the floors and so this was not a test of true integration.

(Wilson et al. 2004) investigated the ability of people to integrate across vertical levels in a virtual reality experiment in which subjects were required to learn the location of pairs of objects on three levels of a virtual multilevel building. Participants then indicated, by pointing, the direction of objects from various vantage points within the virtual environment. Distance judgements between floors were distorted, with relative downward errors in upward judgements and relative upward errors in downward judgements. This effect is interesting as the sense of vertical space seems to be biased towards the horizontal dimension. However, in
this study there was also a (slightly weaker) tendency to make rightwards errors to objects to the left and leftwards errors to objects to the right: the results may therefore reflect a general tendency for a bias towards making errors in the direction towards the center point of the spatial range, regardless of dimension (although this was more pronounced for the vertical dimension). Another interesting finding from this study is that there appears to be a vertical asymmetry in spatial memories, with a bias in favour of memories for locations that are on a lower rather than higher plane. (Tlauka et al. 2007) have confirmed the findings of Wilson et al. (2004) and suggest that it might be a “contraction bias” due to uncertainty, reflecting a compressed vertical memory. The authors speculated that such biases might be experience-dependent, as humans pay more attention to horizontal space directly ahead than to regions above or below it. This view is interesting and might explain the experimental results seen with rats in Jovalekic et al. (2011). Finally, young children appear to have more problems in such three-dimensional pointing tasks (Pick & Rieser 1982), once again suggesting that experience may modulate performance in vertical tasks.

In the second part of this article, it will be argued that the mammalian spatial representation may be fundamentally planar, with position in the vertical dimension (orthogonal to the plane of locomotion) encoded separately, and in a different (non-metric) way, from position in the horizontal (locomotor) plane. If true, this has implications for the encoding of multilevel structures, in which the same horizontal coordinates recur at different vertical levels. On the one hand, this may cause confusion in the spatial mapping system – if the representations of horizontal and
vertical dimensions are not integrated, then the navigational calculations that use horizontal coordinates may confuse the levels. As suggested above, this may be why, for example, people are confused by multi-level buildings. On the other hand, if there is, if not full integration, at least an interaction between the vertical and horizontal representations, then it is possible that separate horizontal codes are able to be formed for each vertical level, with a corresponding disambiguation of the levels. Further studies at both behavioural and neurobiological levels will be needed to determine how multiple overlapping levels are represented and used in navigation.

3.4 Navigation in a volumetric space

Volumetric spaces such as air and water allow free movement in all dimensions. Relatively little navigational research has been conducted in such environments so far. One reason is that it is difficult to track animal behaviour in volumetric spaces. The advent of long-range tracking methods that use global positioning system (GPS) techniques or ultrasonic tagging is beginning to enable research in to foraging and homing behaviours over larger distances (Tsoar et al. 2011). However, many of these studies still only address navigation in the Earth-horizontal plane, ignoring possible variations in flight or swim height. Laboratory studies have started to explore navigation in 3D space, though the evidence is still limited.

One of the earliest studies of three-dimensional navigation in rats, the Grobety and Schenk (1992a) experiment mentioned earlier, involved a cubic lattice, which allowed the animals to move in any direction at any point within the cube (Grobéty
& Schenk 1992a; see Fig. 4A). As noted above, it was found that rats learnt the correct vertical location before the horizontal location, suggesting that these elements may be processed separately. Grobéty & Schenk suggested that the rats initially paid greater attention to the vertical dimension both because it is salient and because greater effort is required to make vertical as opposed to horizontal translations, therefore, minimising vertical error reduced their overall effort. By contrast, Jovalekic et al. (2011) found that rats shuttling back and forth between diagonally opposite low and high points within a lattice maze or on a vertical subset of it, the pegboard (Fig. 4B) did not tend to take the direct route (at least on the upward journey) but preferred to execute the horizontal component of the journey first. However, the separation of vertical from horizontal may have occurred because of the constraints on locomotion inherent in the maze structure: the rats may have found it easier to climb vertically than diagonally. Another experiment, however, also found vertical/horizontal differences in behavior. Jovalekic et al. (2011) examined foraging behaviour in the Grobety and Schenk lattice maze and found that rats tended to retrieve as much of the food as possible on one horizontal layer before moving to the next. They observed similar behaviour on a simplified version of this apparatus, the pegboard, where again, foraging rats depleted one layer completely before moving to the next. The implication is that rather than using a truly volumetric representation of the layout, the animals tend to segment the environment into horizontal bands.

A similar propensity to treat 3D environments as mainly horizontal has been reported in humans by Vidal et al (Vidal et al. 2004) who studied participants in a
virtual 3D maze and found that they performed better if they remained upright and had therefore aligned their egocentric and allocentric frames of reference in one of the dimensions (Vidal et al, 2004; similar results were also reported in Aoki et al. 2005). Similarly, astronauts in a weightless environment tended to construct a vertical reference using visual rather than vestibular (i.e., gravitational) cues, and remain oriented relative to this (Lackner & Graybiel 1983; Tafforin & Campan 1994; Young et al. 1984).

Finally, one intriguing study has shown that humans may be able to process not only three- but also four-dimensional (4D) information (Aflalo & Graziano 2008). In this virtual reality study, in addition to the usual three spatial dimensions, a fourth dimension was specified by the ‘hot’ and ‘cold’ directions such that turns in the maze could be left, right, up, down, hot or cold, and the usual trigonometric rules specified how movements in the fourth dimension related to movements in the other three. Participants were required to path integrate by completing a multi-segment trajectory through the maze and then pointing back to the start. They eventually (after extensive practice) reached a level of performance exceeding what they could have achieved using three-dimensional reasoning alone. This study is interesting because it is highly unlikely that any animals, including humans, have evolved the capacity to form an integrated 4D cognitive map, and so the subjects’ performance suggests that it is possible to navigate reasonably well using a metrically lower-dimensional representation than the space itself. The question, then, is whether one can in fact navigate quite well in three dimensions using only a 2D cognitive map. We will return to this point in the next section.
4. **Neurobiological studies in three dimensions**

So far, we have reviewed behavioural research exploring three-dimensional navigation and seen clear evidence that 3D information is used by animals and humans, although the exact nature of this use remains unclear. A parallel line of work has involved recording from neurons that are involved in spatial representation. This approach has the advantage that it is possible to look at the sensory encoding directly, and make fine-grained inferences about how spatial information is integrated. While most of this work has hitherto focussed on flat, two-dimensional environments, studies of three-dimensional encoding are beginning to increase in number. Here, we shall first briefly describe the basis of the brain’s spatial representation of two-dimensional space, and then consider, based on preliminary findings, how information from the third dimensions is (or might be) integrated.

The core component of the mammalian place representation comprises the hippocampal place cells, first identified by O'Keefe and colleagues (O'Keefe & Dostrovsky 1971), which fire in focal regions of the environment. In a typical small enclosure in a laboratory, each place cell has one or two, or sometimes several, unique locations in which it will tend to fire, producing patches of activity known as place fields (Fig. 7 A). Place cells have thus long been considered to represent some kind of marker for location. Several decades of work in the years since place cells were discovered have shown that the cells respond not to raw sensory inputs from the environment such as focal odours, or visual snapshots, but rather, to higher-order sensory stimuli that have been extracted from the lower-order sensory
data – examples include landmarks, boundaries and contextual cues such as the color or odor of the environment (see Jeffery 2007; Moser et al. 2008 for review).

For any agent (including a place cell) to determine its location, it needs to be provided with information about directions and distances in the environment. Directional information reaches place cells via the head direction system (Taube et al. 1990a; Taube 2007), which is a set of structures in areas surrounding hippocampus whose neurons are selectively sensitive to the direction in which the animal’s head is pointing (Fig. 7B). Head direction (HD) cells do not encode direction in absolute geocentric coordinates – rather, they seem to use local reference frames, determined by visual (or possibly also tactile) cues in the immediate surround. This is shown by the now well-established finding that rotation of a single polarizing landmark in an otherwise unpolarized environment (such as a circular arena) will cause the HD cells to rotate their firing directions by almost the same amount (Goodridge et al. 1998; Taube et al. 1990b). Interestingly, however, there is almost always a degree of under-rotation in response to a landmark – this is thought to be due to the influence of the prevailing “internal” direction sense, sustained by processing of self-motion cues such as vestibular, proprioceptive and motor signals to motion. The influence of these cues can be revealed by removing a single polarizing landmark altogether – the HD cells will maintain their previous firing directions for several minutes, although they will eventually drift (Goodridge et al. 1998).

The internal (sometimes called “idiothetic”) self-motion cues provide a means of stitching together the directional orientations of adjacent regions of an environment
so that they are concordant. This was first shown by Taube and Burton (1995) and replicated by Dudchenko & Zinyuk (2005). However, although there is this tendency for the cells to adopt similar firing directions in adjacent environments, especially if the animal self-locomoted between them, this is not absolute and cells may have discordant directions, flipping from one direction to the other as the animal transitions from one environment to the next (Dudchenko & Zinyuk 2005). This tendency for HD cells to treat complex environments as multiple local fragments is one that we will return to later.

Figure 7: Firing patterns of the three principal spatial cell types. (A) Firing of a place cell, as seen from an overhead camera as a rat forages in a 1 m square environment. The action potentials (“spikes”) of the cell are shown as black spots and the cumulative path of the rat over the course of the trial is shown in grey. Note that the spikes are clustered towards the North-East region of the box, forming a “place field” of about 40 cm across. (B) Firing of a head direction cell, recorded as a rat explored an environment. Here, the heading direction of the cell is shown on the x-axis and the firing rate on the y-axis. Note that the firing intensifies dramatically when the animal’s head faces in a particular direction. The cell is mostly silent otherwise. (C) Firing of a grid cell (from Hafting et al., 2005), depicted as for the place cell in (A). Observe that the grid cell produces multiple firing fields in a regularly spaced array. Adapted from Jeffery and Burgess (2006).
The source of distance information to place cells is thought to lie, at least partly, in the grid cells, which are neurons in the neighbouring entorhinal cortex that produce multiple place fields that are evenly spaced in a grid-like array, spread across the surface of the environment (Hafting et al. 2005); for review see (Moser & Moser 2008). Fig. 7 C. The pattern formed by the fields is of the type known as hexagonal close-packed, which is the most compact way of tiling a plane with circles.

Grid cell grids always maintain a constant orientation for a given environment (Hafting et al. 2005), suggesting a directional influence that probably comes from the head-direction system, as suggested by the finding that the same manipulations that cause HD firing directions to rotate also cause grids to rotate. Indeed, many grid cells are themselves also HD cells, producing their spikes only when the rat is facing in a particular direction (Sargolini et al. 2006). More interesting, however, is the influence of odometric (distance-related) cues on grid cells. Because the distance between each firing field and its immediate neighbours is constant for a given cell, the grid cell signal can theoretically act as a distance measure for place cells, and it is assumed that this is what they are for (Jeffery & Burgess 2006) though this has yet to be proven. The source of distance information to the grid cells themselves remains unknown. It is likely that some of the signals are self-motion-related, arising from the motor and sensory systems involved in commanding, executing and then detecting movement through space. They may also be static landmark-related distance signals – this is shown by the finding that subtly rescaling an environment can cause a partial rescaling of the grid array in the rescaled dimension (Barry et al. 2007), indicating environmental influences on the
grid metric. It seems that when an animal enters a new environment, an arbitrary grid pattern is laid down, oriented by the (also arbitrary) HD signal, and this pattern is then “attached” to this environment by learning processes so that when the animal re-enters the (now-familiar) environment, the same HD cell orientation, and same grid orientation and location, can be reinstated.

Given these basic building blocks of the mammalian spatial representation, we turn now to the question of how these systems may cope with movement in the vertical domain, using the same categories as previously: verticality, slope, elevation, multilayer environments and volumetric spaces.

4.1 Neural processing of verticality cues

The simplest form of 3D processing is detection of the vertical axis. Verticality cues are those that signal the direction of gravity, or (equivalently) the horizontal plane, which is defined by gravity. The brain can process both static and dynamic verticality cues: static information comprises detection of the gravitational axis, or detection of slope of the terrain underfoot, while dynamic information relates to linear and angular movements through space and includes factors such as effort. Both static and dynamic processes depend largely on the vestibular apparatus in the inner ear (Angelaki & Cullen 2008), and also by a second set of gravity detectors ("graviceptors") located in the trunk (Mittelstaedt 1998). Detection of the vertical depends on integration of vestibular cues together with those from the visual world and also from proprioceptive cues to head and body alignment (Angelaki et al. 1999; Merfeld et al. 1993; Merfeld & Zupan 2002).
The vestibular apparatus is critical for the processing not only of gravity but also of movement-related inertial 3D spatial cues. The relationship of the vestibular apparatus is shown in Fig. 8 A, while the close-up in Fig. 8 B shows the core components: these are the otolith organs, of which there are two – the utricle and the saccule – and the semicircular canals. The otolith organs detect linear acceleration (including gravity, which is a form of acceleration) by means of specialised sensory epithelium: in the utricle the epithelial layer is oriented approximately horizontally and primarily detects Earth-horizontal acceleration, and in the saccule it is oriented approximately vertically and primarily detects vertical acceleration/gravity. The two otolith organs work together to detect head tilt. However, because gravity is itself a vertical acceleration, there is an ambiguity concerning how much of the signal is generated by movement and how much by gravity – this “tilt-translation ambiguity” appears to be resolved by means of the semicircular canals (Angelaki & Cullen 2008). These are three orthogonal fluid-filled canals oriented in the three cardinal planes, one horizontal and two vertical, which collectively can detect angular head movement in any rotational plane.

Neural encoding of direction in the vertical plane (tilt, pitch, roll etc) is not yet well understood. Head direction cells have been examined in rats locomoting on a vertical plane and at various degrees of head pitch, to see whether the cells encode vertical as well as horizontal direction. Stackman & Taube (1998) found pitch-sensitive cells in the lateral mammillary nuclei (LMN) – however, these were not true head direction cells, because their activity was not at all related to horizontal heading direction. Nor did they correspond to the vertical-plane equivalent of HD
cells: most of the observed cells had pitch preferences clustered around the almost-vertical, suggesting that these cells were doing something different from simply providing a vertical counterpart of the horizontal direction signal. By contrast, the HD cells in LMN were insensitive to pitch, and their firing was uniformly distributed in all directions around the horizontal plane. Thus, vertical and horizontal directions appear to be separately and differently represented in this structure.

Figure 8: (Taken from [http://weboflife.nasa.gov/learningResources/vestibularbrief.htm](http://weboflife.nasa.gov/learningResources/vestibularbrief.htm))
The vestibular apparatus in the human brain. (A) Diagram showing the relationship of the vestibular apparatus to the ear and skull. (B) Close-up of the vestibular organ showing the detectors for linear acceleration (the otolith organs – comprising the utricle and saccule) and the detectors for angular acceleration (the semi-circular canals, one in each plane).

In a more formal test of vertical directional encoding, (Stackman et al. 2000) recorded HD cells as rats climbed moveable wire mesh ladders placed vertically on the sides of a cylinder. When the ladder was placed at an angular position corresponding to the cell’s preferred firing direction, the cell continued to fire as the animal climbed the ladder, but did not fire as the animal climbed down again.
Conversely, when the ladder was placed on the opposite side of the cylinder, the reverse occurred: now the cell remained silent as the animal climbed up, but started firing as the animal climbed down. This suggests that perhaps the cells were treating the plane of the wall in the same way that they usually treat the plane of the floor. Subsequently, (Calton & Taube 2005) showed that HD cells fired on the walls of an enclosure in a manner concordant with the firing on the floor, and also informally observed that the firing rate seemed to decrease in the usual manner when the rat’s head deviated from the vertical preferred firing direction. This observation was confirmed in a follow-up experiment in which rats navigated in a spiral trajectory on a surface that was either horizontal or else vertically aligned in each of the four cardinal orientations (Taube 2005). In the latter conditions, the cells showed preferred firing directions on the vertical surface in the way that they usually do on a horizontal surface. When the vertical spiral was rotated, the cells switched to a local reference frame and maintained their constant preferred firing directions with respect to the surface.

What are we to make of these findings? In the Calton & Taube (2005) experiment, it appears that the HD cells seemed to be acting as if the walls were an extension of the floor, meaning, in other words, that it is as if the pitch transformation, when the rat transitioned from horizontal to vertical, had never happened. A possible conclusion is that HD cells are insensitive to pitch, which accords with the LMN findings of Stackman and Taube (1998). This has implications for how 3D space might be encoded and the limitations thereof, which we return to in the final section. In the spiral experiment described in Taube (2005), it further appears that
the reference frame provided by the locomotor surface became entirely
disconnected from the room reference frame – when the rat was clinging to the
vertical surface, the cells became insensitive to the rotations in the azimuthal plane
that would normally modulate their firing on a horizontal plane. The idea of local
reference frames is one that we will come back to later.

The Calton and Taube (2005) experiment additionally replicated an observation
made in microgravity, which is that head direction cells lose their directional
specificity or are greatly degraded during inverted locomotion on the ceiling (Taube
et al. 2004). In another study, head direction cells were monitored while animals
were held by the experimenter and rotated by 360 degrees (relative to the horizontal
plane) in the upright and in the inverted positions, as well as by 180 degrees in the
vertical plane (Shinder & Taube 2010). Although the animal was restrained, head
direction cells displayed clear directionality in the upright and vertical positions,
and no directionality in the inverted position. Additionally, directionality in the
vertical position was apparent until the animal was almost completely inverted.
These findings confirm those of the previous studies (Calton & Taube 2005;Taube
et al. 2004), that loss of HD cell directionality is a feature of the inverted position.
This breakdown in signalling may have arisen from an inability of the HD cells to
reconcile visual evidence for the 180 degree heading-reversal with the absence of a
180 degree rotation in the plane to which they are sensitive (the yaw plane). This
finding is consistent with a recent behavioural study by Valerio et al. (2010). Here,
rats were trained to locomote while clinging upside-down to the underside of a
circular arena, in order to find an escape hole. Animals were able to learn fixed
routes to the hole but could not learn to navigate flexibly in a mapping-like way, suggesting that they had failed to form a spatial representation of the layout of the arena. Taking these findings together, then, it appears that the rodent cognitive mapping system is unable to function at all possible orientations of the animal.

What about place and grid cells in vertical space? Processing of the vertical in these neurons has been explored in a recent study by Hayman et al. (2011). They used two environments: a vertical climbing wall (the pegboard), and a helix (see Fig. 9; Hayman et al, 2011). In both environments, place and grid cells produced vertically elongated fields, with grid fields being more elongated than place fields, thus appearing stripe-like (Fig. 9 E). Thus, although elevation was encoded, its representation was not as fine-grained as the horizontal representation. Furthermore, the absence of periodic grid firing in the vertical dimension suggests that grid cell odometry was not operating in the same way as it does on the horizontal plane. One possible explanation is that grid cells do encode the vertical axis, but at a coarser scale, such that the periodicity was not evident in these spatially restricted environments. If so, then the scale or accuracy of the cognitive map, at least in the rat, may be different in horizontal vs. vertical dimensions, possibly reflecting the differential encoding requirements for animals that are essentially surface-dwelling. An alternative possibility, explored below, is that while horizontal space is encoded metrically, vertical space is perhaps encoded by a different, possibly even non-grid-cell-dependent, mechanism.
Figure 9: Adapted from Hayman et al (2011). (A) Photograph of the helical maze. Rats climbed up and down either 5 or 6 coils of the maze, collecting food reward at the top and bottom, while either place cells or grid cells were recorded. (B) The firing pattern of a place cell (top, spikes in blue) and a grid cell (bottom, spikes in red), as seen from the overhead camera. The path of the rat is shown in grey. The place cell has one main field with a few spikes in a second region, and the grid cell has three fields. (C) The same data as in (B) but shown as viewed from the size, with the coils unwound into a linear strip. The single place field in (B) can be seen here to repeat on all of the coils, as if the cell is not discriminating elevation, but only horizontal coordinates. The grid cell, similarly, repeats its firing fields on all the coils. (D) Photograph of the pegboard, studded with wooden pegs that allowed the
rat to forage over a large vertical plane. (E) The firing pattern of a place cell (top, spikes in blue) and a grid cell (bottom, spikes in red) as seen from a horizontally aimed camera facing the pegboard. The place cell produced a single firing field, but this was elongated in the vertical dimension. The grid cell produced vertically aligned stripes, quite different from the usual grid-like pattern seen in Fig. 8C.

Note that because the rats remained horizontally oriented during climbing, it is also possible that it is not the horizontal plane so much as the current plane of locomotion that is represented metrically, while the dimension normal to this plane (the dorso-ventral dimension with respect to the animal) is represented by some other means. Indeed, given the HD cell data discussed above, this alternative seems not only possible but likely. Thus, if the rats had been vertically oriented in these experiments, perhaps fields would have had a more typical, hexagonal pattern.

It is clear from the foregoing that much remains to be determined about vertical processing in the navigation system – if there is a vertical head-direction-cell compass system analogous to the one that has been characterized for the horizontal plane then it has yet to be found, and if elevation is metrically encoded, the site of this encoding is also unknown. The cells that might have been expected to perform these functions – HD, place and grid cells – do not seem to treat elevation in the same way as they treat horizontal space, which argues against the likelihood that the mammalian brain, at least, constructs a truly integrated volumetric map.
We turn now to the question of what is known about the use of 3D cues in navigationally relevant computations – processing of slopes, processing of multilayer environments and processing of movement in volumetric space.

4.2 *Neural encoding of a sloping surface*

Neural encoding of non-horizontal surfaces is only in the early stages, but preliminary studies have been undertaken in both the presence and the absence of gravity. In normal gravity conditions, a slope is characterized by its steepness with respect to earth-horizontal, which provides important constraints on path integration. If the distance between a start location and a goal includes a hill, additional surface ground has to be covered to achieve the same straight-line distance. Furthermore, routes containing slopes require more energy to traverse than routes without slopes, requiring a navigator to trade off distance against effort in undulating terrain. Thus, one might imagine that slope should be incorporated into the cognitive map, and neural studies allow for an investigation of this issue that is not feasible with behavioural studies alone.

The effect of terrain slope has been explored for place cells but not yet for grid cells. The earliest place cell study, by Knierim & McNaughton (2001), investigated whether place fields would be modulated by the tilt of a rectangular track. *A priori*, one possible outcome of this experiment was that place fields would expand and contract as the tilted environment moved through intrinsically three-dimensional, ovoid place fields. However, fields that were preserved following the manipulation did not change their size, and many place cells altered their firing altogether on the
track, evidently treating the whole tilted track as different from the flat track. One interpretation of these findings is that the track remained the predominant frame of reference (the “horizontal” from the perspective of the place cells), and that the tilt was signalled by the switching on and off of fields, rather than by parametric changes in their morphology.

This study also showed that slope could be used as an orienting cue: cells were more likely to be oriented by the track (as opposed to the plethora of distal cues) in rotations of the tilted track than in rotations of the flat track. This observation was replicated in a subsequent study by Jeffery et al. (2006) who found that place cells used the 30° slope of a square open arena as an orienting cue after rotations of the arena, and preferred the slope to local olfactory cues. Presumably, since the cells all reoriented together, the effect was mediated via an effect on the head direction system. Place cells have also been recorded from bats crawling on a near vertical (70° tilt) open-field arena (Ulanovsky & Moss 2007). Firing fields were similar to rodent place fields on a horizontal surface, a pattern that is consistent with the fields using the surface of the environment as their reference plane (because if Earth-horizontal were the reference, then fields should have been elongated along the direction of the slope, as in the Hayman et al. (2011) experiment (Fig. 9E).

Grid cells have not yet been recorded on a slope – the pattern they produce will be informative, because their odometric properties will allow us to answer the question of whether the metric properties of place fields in horizontal environments arise
from distances computed in Earth-horizontal, or with respect to environment surface. The results discussed above lead us to predict the latter.

Only one place cell study has been undertaken in the absence of gravity, in an intriguing experiment on the Neurolab Space Shuttle mission of 1998 by Knierim and colleagues (Knierim et al. 2000; Knierim et al. 2003). Rats were implanted with microelectrodes for place cell recording, and familiarized with running on a rectangular track on Earth, before being carried in the shuttle into space. Under microgravity conditions in space, the animals were then allowed to run on a three dimensional track, the ‘Escher staircase,’ which bent through 3D space such that three consecutive right-angle turns in the yaw plane of the animal, interspersed with three pitch rotations, led the animal back to its starting point (Fig. 10 A). The question was whether the cells could lay down stable place fields on the track, given that the visual cues might have conflicted with the path integration signal, if the path integrator could not process the 3D turns. On the rats’ first exposure to the environment on flight day 4 the cells showed somewhat inconsistent patterns, with those from one rat showing degraded spatial firing, those from a second rat showing a tendency to repeat their (also degraded) firing fields on each segment of the track and those from a third looking like normal place fields. By the second recording session on the 9th flight day, place fields from the first two rats had gained normal-looking properties, with stable place fields (Fig. 10 B) suggesting either that the path integrator could adapt to the unusual conditions and integrate the yaw and pitch rotations, or else that the visual cues allowed the construction of multiple independent planar representations of each segment of the track.
Figure 10: (From Knierim et al., 2000 with permission) (A) The “Escher staircase” track, on which rats ran repeated laps while place fields were recorded. (B) The firing field of a place cell, showing a consistent position on the track, indicating either the use of visual cues or of 3D path integration (or both) to position the field.

It will be important to answer the question of which of these explanations is true. If the cells could integrate turns in three dimensions, this would imply a highly sophisticated ability of the head direction system to co-ordinate turns in the three axes of rotation, which – given that they respond mainly to yaw – suggests an integrative process outside the HD cells themselves.

4.3 Multilayer environments

Very little work has examined neural encoding of space in multi-layer environments, which is unfortunate given the theoretical importance of this issue, discussed earlier. Stackman et al. (2000) compared the firing of HD cells between the floor of a cylinder, and an annular mezzanine “annulus” located 74 cm above the floor, and found a 30% increase in firing rate on the annulus. However, the
wire-mesh studies of Stackman et al. (2000) and Calton & Taube (2005) did not report increases in firing rate as animals climbed the walls, so the annulus effect may perhaps have been due to arousal or anxiety on the annulus. Directional firing preferences were shifted slightly on the upper level with respect to the lower, which may have resulted from transient slight disorientation during the climb.

There is also little work on place and grid cells in multilayer environments. Fenton et al. (2008) recorded place cells in a large environment, the walls of which supported stairs leading to an upper ledge housing reward drinking ports. They found that frequently, fields that were present on the floor were also present on the corresponding part of the stairs/ledge. This could mean that either the fields were three dimensional and extended from the floor to the staircase, or else that the cells were only responding to horizontal co-ordinates and ignoring the vertical. Results from the Hayman et al. (2011) experiment on the helical track support the latter interpretation, in that the data showed that firing tended to repeat on each successive coil (Fig. 9B and C). There was modulation of firing rate across the coils, which is not inconsistent with the notion that the fields may have been intrinsically 3D (albeit very elongated in the vertical dimension). However, it could also simply be rate modulation of the same field on each coil. The rate-modulation of place fields has been seen on other situations (Leutgeb et al. 2004) and suggests an influence of some other cues (in this case, elevation) on field production. Grid cells in this experiment showed a similar pattern, with field location repeating from one coil to the next, but also (as with the place cells) a slight degree of rate
modulation – something, interestingly, that has not been reported in other experiments on grid cells.

The implication is that the system was treating the environment as a stack of horizontal planes, rather than as an integrated 3D space. However, the helical track experiment is not a true test of multi-layered environment representation, because there was no discrete boundary between layers as there is in, say, a multi-storey building. Also, the stereotypical nature of the route the animals ran might have precluded the formation of independent maps for each layer. A truly multi-storey apparatus will be required to resolve this issue.

4.4 Neural encoding of a volumetric space

The final and most important question concerning 3D spatial encoding is whether there is metric encoding of volumetric space. As with multilayer environments, there have been no studies done to date on neural responses to free movement through true volumetric spaces, and so the answer to this question remains open. However, it is interesting to speculate about what kinds of evidence we should look for.

A volumetric map requires directions and distances to be calculated for both the horizontal and vertical dimensions. We have already seen that there is no evidence as yet for 3D head direction cells, at least in rodents, so either there are 3D HD cells in some other region of the brain, or else the cognitive map has a way of combining the planar signal from the classic HD areas together with pitch or elevation information, perhaps from the lateral mammillary nuclei.
What about distance in 3D? Since grid fields seem to provide the odometric representation for horizontal distance by virtue of their regularly spaced grid fields, we could predict that the periodicity observed in the horizontal dimension might be part of a three-dimensional, lattice-like arrangement of spherical fields. There are two ways of close-packing spheres in a volume, termed hexagonal and face-centered cubic, respectively (Fig. 11 A and B). Since we already know that the grid array on the horizontal plane is hexagonal close-packed, if either of these volumetric arrangements were to pertain then we would predict it should be the hexagonal form.

**Figure 11**: The two ways of efficiently filling a volumetric space with spheres, are (A) hexagonal close-packing, in which the spheres on each layer are planar hexagonal close-packed, and offset between one layer and the next, and (B) Face-centered cubic, in which the spheres on each layer are in a square arrangement, with each sphere in a given layer nestling in the dip formed by four of those in the layers on either side. (C) The result of transecting a 3D hexagonal close-packed array is a regular set of circles. (D) The resulting of transecting a set of horizontally hexagonal close-packed columns is stripes, resembling the grid cell pattern of Fig. 9E.
Although no recordings have yet been made of grids in a volumetric space, the experiment of Hayman et al. (2011) of grids recorded on a vertical plane might have been expected to reveal evidence of periodicity in the vertical dimension, akin to what one would find by transecting a close-packed volume (Fig 6 C). However, that experiment found that the grid fields formed stripes on the vertical surface, implying instead that the true underlying pattern of grids in a volumetric space might be in the form of hexagonal close-packed columns (Fig 6 D). Whether this would also hold true for animals moving in an unconstrained way through the space remains to be seen. It is also possible that there is periodicity in the vertical dimension, but at a much greater scale than could be seen in a small laboratory setting.

5. Is the cognitive map bicoded?

At the beginning of this article we reviewed the encoding schemes that might exist to map three-dimensional space, highlighting three different mapping schemes that vary in their use of elevation information (Fig. 1). Based on the foregoing review of behavioural and neurobiological studies, we return now to the question of map types. Below, it will be argued that the data collected to date favour a metrically planar map of the form we have termed “bicoded” – that is, using a different encoding scheme for horizontal than for vertical space. Indeed, we will go a step further and suggest that the schemes are referenced not to horizontal and vertical, but – more generally – to the plane of locomotion (horizontal in the animal’s canonical orientation) vs. the axis orthogonal to this.
First, however, let us consider the alternative hypotheses for the possible structure of the cognitive map. These are (i) that the map is a surface map, lacking information about vertical travel, or (ii) that it is a fully integrated volumetric map. It is unlikely that the cognitive map could be a surface map, because the evidence reviewed suggests that almost all animals investigated show some processing of vertical space. A volumetric map, by contrast, is metric (i.e., distances and directions are explicitly encoded) in all three dimensions. No studies have yet been undertaken in animals that can move freely in all three dimensions, such as those that fly or swim, and it may be that through evolution and/or through developmental experience, a fully integrated volumetric map may have formed in such animals. In particular, animals that have to solve three-dimensional spatial problems for which there is no planar approximation, such as those that have to path integrate across a volumetric space like the ocean, may be expected to possess a volumetric cognitive map if such a thing exists. Studies in birds, fish or swimming and flying mammals such as bats or cetaceans have the potential to reveal whether there are spatial metric computations for travel in all three dimensions, and also whether these are integrated. If the cognitive map in animals that have a truly three-dimensional ecology is volumetric, we might expect to discover a new class of head direction cells in these animals that encode the vertical plane in the same way that the rodent HD cells that have been discovered to date encode the horizontal plane. Alternatively, these animals might even possess HD cells that point uniquely to positions in 3D space (that is, that have tuning curves
restricted in all three dimensions). We might also expect to see spherical grid cell grids in a 3D hexagonal close-packed array like those shown in Figure 11 A and B.

5.1 Arguments for a bicoded cognitive map

The above notwithstanding, we suggest here that even ecologically three-dimensional animals may turn out to form planar, or quasi-planar, cognitive maps. Such animals often live in environments that are relatively more extended horizontally than vertically, making a planar map useful (particularly since almost all animal maintain a canonical body orientation with respect to gravity). Also, as discussed below, encoding three dimensions poses considerable technical challenges for a compass system. Indeed, there is little evidence for such integrated encoding in the species that have been studied so far. Head direction cells evidently operate in a plane rather than a volume, and grid cells are not periodic in the vertical dimension (or perhaps not in the dimension perpendicular to the animal’s body plane).

In a bicoded map, there is a spatially metric representation of a reference plane – usually the horizontal plane – and a non-metric, or at least differently-metricized, representation for the axis perpendicular to this plane. In the geographical example given earlier, the non-spatial indicator of distance was color, which varies as a function of elevation but from which elevation cannot be extracted without a key. Loosely speaking, it is as if the non-metric variable conveys a coarse signal along the lines of “high” or “very high”, but does not contain explicit information about distance which could enter into a trigonometric calculation. Obviously the cognitive
map does not use color as its non-metric variable, but it could plausibly use an analogous signal which we call here, for want of a better word, “contextual” – that is, information that pertains to a space but need not itself contain distance or direction information. A real-world example of a contextual elevation cue might be hydrostatic pressure, which Burt de Perera and colleagues put forward as a possible mechanism for fish to localize themselves in the vertical axis (Burt de Perera et al. 2005). For terrestrial animals it could be visual or kinaesthetic cues to ascent/descent, or aspects of the visual panorama, and for flying animals it could be depth cues from stereopsis or motion parallax, or perhaps olfactory gradients in the atmosphere. Our hypothesis then is that the cognitive map is such a bicoded map – it is metric in a given plane (the plane of locomotion), and contextually modulated orthogonal to this. It is possible that contextual information concerning elevation could be processed by brain areas that are located outside the currently established navigation circuitry, although it is likely that these signals would interact at some point..

What evidence supports this bicoding view? Existing behavioural data from a number of studies, reviewed earlier, suggest a potential separation between horizontal and vertical dimensions. However, this does not prove that the encoding itself is different, merely that execution of navigational plans is different. Theoretically, this could be because the energetic costs of moving vertically are taken into account when planning, even if the planning is based on a fundamentally volumetric, 3D map. The neural data, however, tell a more revealing story. The observation that head direction cells appear to use the plane of locomotion as a
local reference (Stackman et al. 2000) implies that the directional signal for the vertical plane (or rather, the plane normal to locomotion) must lie elsewhere, in an as-yet-undiscovered region. Furthermore, the finding that grid and place cells have different metric properties in the vertical dimension, at least for horizontally oriented body posture, implies that odometry does not take place equally in all three dimensions – the map is anisotropic (Hayman et al. 2011). Note that it remains to be determined experimentally whether there is something special about the horizontal plane in the cognitive map, or whether it is simply that gravity primarily orients the locomotor plane, and it is this plane that is metrically encoded. Recordings of grid cells on steeply sloping surfaces may answer this question.

A bicoded scheme seems suboptimal, but in fact a surface-travelling animal could do quite well with such a map, because it could use the metrically encoded locomotor plane to calculate the directions needed for travel, and the topography of the landscape would constrain the animal to be at the correct elevation for that position. Although a bicoded map could not be used for path integration in volumetric space, an animal could divide the navigational task into two parts – it could first calculate the horizontal trajectory, and then ascend to the correct elevation (as indicated by the contextual gradient) or vice versa (or, indeed, both simultaneously but in separate computational circuits). Indeed, the findings of Grobéty & Schenk (1992a) seem to suggest that this is indeed what happens, at least in rats. However, in the case where “horizontal” is not true, Earth-horizontal, but lies on a slope, then it would be advantageous for the map to encode, in some form, information about that slope, since this would affect the energy costs of
navigational routes involving that surface. Thus, there may be some mechanism for encoding the angle of a given planar fragment with respect to earth-horizontal.

A bicoded map has a planar metric reference, but as noted earlier, the real world is far from planar. How could a bicoded map be used on undulating topology? One possibility is that an undulating surface could be encoded as a manifold – not a single unified map, but a mosaic of contiguous map fragments (Fig. 12). In such a scheme, each fragment would comprise a local reference frame defined by a reference plane, which would be horizontal on a horizontal surface but need not be, and an axis orthogonal to this (which would correspond to the gravitationally-defined vertical for an animal locomoting in its canonical orientation on a horizontal surface). Figure 12 shows an illustrative example of such a mosaicized representation. A rat, in a fairly short excursion, moves from the ground to the trunk of a tree to a branch, each of these planes being orthogonal to the one before. To maintain orientation it must switch from a planar map referenced to the ground, to a new one referenced to the tree trunk, to a third referenced to the branch.

If the planar fragments of the cognitive map need not be oriented horizontally, then how does the animal process the vertical axis, which remains a special, and important, reference due to the effects of gravity? To form an integrated map of large-scale space using mosaic fragments, there also should be some means of associating the local reference frames, both to each other (to enable navigational between fragments) and to Earth-horizontal (to enable appropriate energy and other gravity-related calculations to be made). A plausible means of achieving this could be the vestibular system, which tracks changes in orientation via the semi-circular
canals and could therefore provide some means of linking the map fragments together at, so to speak, their edges. With such a scheme it would be possible to use heuristic methods to compute, for example, approximate shortcuts between map fragments. For more sophisticated navigation, however, it would be preferable to weld fragments together into a larger, unitary frame that would allow for precise metric computations across the space. It may be that such a reference frame adaptation could occur with experience, and the head direction cell studies of Stackman et al (1998), in which reference frames could be local to the surface or more globally referenced to the surrounding room, support the possibility of such experience-dependent plasticity.

The possibility that the cognitive map is formed of a mosaic of locally planar fragments is consistent with the sudden reorientations experienced by humans and animals as they transition from one region of the environment to the next (Wang & Spelke 2002). It is a commonplace experience of urban navigators when taking an unfamiliar route between two locally familiar places that a sudden, almost dizzying re-alignment of reference frames occurs across the interface between the two regions. This subjective experience may reflect the sudden reorientation of head direction cells that occurs when animals cross between local connected regions having different head direction cell orientations (Dudchenko & Zinyuk 2005; Taube & Burton 1995; Zugaro et al. 2003). The reorientation episodes experienced by astronauts in conditions of microgravity may have a similar underlying cause (Oman 2007).
Figure 12: Hypothetical structure of the cognitive map in a dimensionally complex environment. The map is assumed to be metric in the plane of locomotion, which is the ground for the space between the rock and the tree, the vertical trunk as the animal climbs, and the horizontal branch. The color gradient represents the non-metric encoding, which conveys coarse information about distance in the plane orthogonal to the locomotor plane, which is encoded differently from the precise metric of the locomotor plane.

What about animals that navigate in volumetric spaces, such as fish, birds, bats, and (in humans) divers, aviators and astronauts? Do they use a surface map too? In these environments, a plane-referenced map would seem much less useful.

However, evidence from microgravity environment studies shows that even in a weightless environment, astronauts tend to find a reference plane to serve as a local “floor” (Oman 2007). This is usually the surface beneath the feet, but if they drift too far away from this and too far towards the “ceiling”, or switch visual attention, they frequently report visual reorientation illusions (VRIs), where surfaces abruptly and unexpectedly exchange identity (Oman et al. 1986), testifying to the salience of the reference plane.
5.2 Why would animals have a metrically planar map?

Why would animal cognitive maps be (quasi-)planar? On the face of it, this seems maladaptive, because a planar map, even a bicoded one, has inherently less information than a fully integrated volumetric one. The first possibility is an ontogenetic one: that development of the map during development is constrained by the experience of the infants, and the animals that have been investigated to date have, mainly for technical reasons, been those species that are primarily surface-dwelling. An exception is fish, in which we also saw a processing separation between horizontal and vertical space (Holbrook & Burt de Perera 2009), but this could have been a behavioural adaptation superimposed on an underlying volumetric map. It may be that if rats and mice were raised from birth in a volumetric space – for example, in microgravity, or at least in a lattice system in which they could move in all directions – then we would see true three-dimensional HD cells, and path integration behaviour that seamlessly encompassed all three dimensions.

Second, it may be that phylogenetic development of the map was so constrained, and therefore, surface-dwelling animals like rats, mice and humans would never develop a 3D map even if raised in conditions that would allow it. In other words, surface-dwellers have either lost or else never evolved the neurological capacity to fully represent 3D space. If we were to study HD cells in animals that can move freely in all dimensions then we may find 3D HD cells in these species. Emerging studies in bats may hopefully soon answer this question. It should be noted, however, that rats and mice naturally inhabit dimensionally complex environments,
and might be expected to show integrated 3D encoding if this ever did evolve in vertebrates.

The third and final possibility, and one for which we argue here, is that a fully 3D map has never evolved in any species, because of the difficulties inherent in stably encoding an allocentric 3D space using egocentric sensory receptors. A volumetric map requires three co-ordinates for position and three for direction. Monitoring position and travel in such a space is complicated because the vestibular inputs to the head direction system originate in the semicircular canals, which are themselves planar (one aligned in each plane). In order to extract 3D heading, the system would need to modulate the rotational signal from each canal with each of the other two, and do so dynamically and instantaneously. It is possible that the cost-benefit ratio of the required extra processing power is not sufficiently great, even for animals that have a three-dimensional ecological niche. As we saw in the 4D experiment of Aflalo and Graziano (2011), it is possible to navigate quite well in a space using a lower-dimensional representation together with some heuristics, and what applies in four dimensions may equally well apply in three.

6. Conclusion

To conclude, then, we have reviewed emerging experimental evidence concerning the encoding of three-dimensional space in animals and humans. We find clear evidence of ability to process information about travel in the vertical dimension, particularly in the domain of slope intensity and slope direction. There is also some evidence of the processing of elevation information. Evidence for true volumetric
coding is, however, weak, and both behavioural and neural studies suggest the
alternative hypothesis that the neural representation of space, in a wide variety of
species, is metrically flat (referenced to a plane – usually the plane of locomotion)
and is modulated in the vertical dimension in a non-metric way. We have termed
such a map “bicoded” to reflect its essentially anisotropic nature, and suggest that a
bicoded representation is a universal feature of vertebrate (and possibly
invertebrate) cognitive maps.

Acknowledgements

The work was supported by grants to KJ from the Wellcome Trust
(GR083540AIA), the European Commission Framework 7 (‘Spacebrain’), the
Medical Research Council (G1100669) and the Biotechnology and Biological
Sciences Research Council (BB/J009792/1) and a grant to Axona Ltd from the
European Commission Framework 7 (‘Spacebrain’).
References


