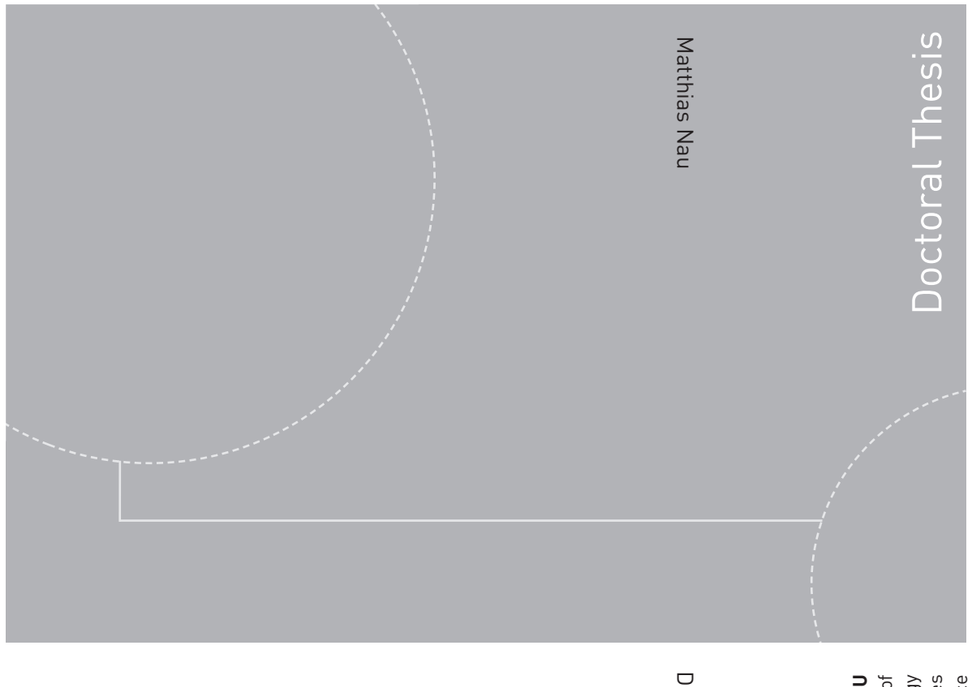
ISBN 978-82-326-4418-6 (printed version)
ISBN 978-82-326-4419-3 (electronic version)
ISSN 1503-8181



Doctoral theses at NTNU, 2020:32

Matthias Nau

Perception and the cognitive map

Deriving a stable world from visual inputs

NTNU
Norwegian University of
Science and Technology
Faculty of Medicine and Health Sciences
Kavli Institute for Systems Neuroscience

NTNU
Norwegian University of
Science and Technology

Doctoral theses at NTNU, 2020:32

NTNU

NTNU
Norwegian University of
Science and Technology

Matthias Nau

Perception and the cognitive map

Deriving a stable world from visual inputs

Thesis for the degree of Philosophiae Doctor

Trondheim, January 2020

Norwegian University of Science and Technology
Faculty of Medicine and Health Sciences
Kavli Institute for Systems Neuroscience

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Medicine and Health Sciences Kavli Institute for Systems Neuroscience

ISBN 978-82-326-4418-6 (printed version)
ISBN 978-82-326-4419-3 (electronic version)
ISSN 1503-8181

Doctoral theses at NTNU, 2020:32

Printed by Skipnes Kommunikasjon as

Supervisor:

Christian F. Doeller

Kavli Institute for Systems Neuroscience, Centre for Neural Computation, NTNU, Trondheim, Norway
Max-Planck-Institute for Human Cognitive and Brain Sciences,
Leipzig, Germany

Co-supervisor:

Edvard I. Moser

Kavli Institute for Systems Neuroscience, Centre for Neural Computation, NTNU, Trondheim, Norway

Thesis assessment committee:

Sylvia Wirth

Institut des Sciences Cognitives Marc Jeannerod, CNRS UMR 5229, University of Lyon, Bron, France

Peter Kok

Wellcome Centre for Human Neuroimaging, University College London,
London, United Kingdom

Jonathan Whitlock

Kavli Institute for Systems Neuroscience, Centre for Neural Computation, NTNU, Trondheim, Norway

Norsk sammendrag

Hjernen skaper stabile og sammenhengende representasjoner av verden fra omskiftelige perseptuelle input. Denne prosessen involverer et omfattende nettverk i hjernen, fra de tidlige synssentrene på lavt nivå til hukommelsessentrene i den mediale tinninglapp på høyere nivå. Langs denne banen gjennomgår informasjon fra visuelle input en rekke endringer før den ender opp som del av det kognitive kartet, som utgjør en ikke-sensorisk representasjon av omgivelsene som understøtter hukommelse og atferd. I mitt doktarbeid undersøkte jeg det nevrale grunnlaget for spatial persepsjon og kognitive kart i menneskehjernen, samt hvordan det kan knyttes til hukommelse og atferd. Prosjektene som presenteres, tilnærmet seg problemstillingene fra flere vinkler ved å fokusere på ulike komputeringstrinn langs den hierarkiske rekkefølgen av kortikale områder.

Vi undersøkte først hvordan synssystemet stabiliserer visuell persepsjon under bevegelse ved å observere menneskelig hjerneaktivitet ved bruk av magnetresonanstomografi (fMRI) under en visuell sporingsoppgave (VSO). Da oppdaget vi at et stort nettverk av områder kodet for visuell bevegelse uavhengig av selvbevegelse, inkludert i de tidlige synssentrene på laveste nivå. Denne mekanismen forankrer visuell persepsjon eksternt i ens omgivelser. Deretter undersøkte vi hvordan høyere kognitive områder representerer ens synlige omgivelser. Fremdeles ved bruk av fMRI og en VSO, observerte vi gittercellelignende aktivitet i entorhinal cortex, et hjerneområde vi vet kartlegger omgivelsene ved navigering. Vi fant en fMRI-signatur som lignet på denne aktiviteten som samsvarte med øyebevegelsene alene. Dette utgjør det første bevis på at entorhinal cortex også koder for et ikke-sensorisk kart over synsfeltet i mennesket. Det siste prosjektet samførte sensoriske og ikke-sensoriske spatiale representasjoner og fokuserte eksplisitt på atferden de understøtter. Jeg utviklet et toppmoderne prediktivt modelleringsrammeverk til studier av naturtro atferd ved bruk av virtuell virkelighet og ultrahøyfelts fMRI i mennesker. Undersøkelser av retningsstuning i regioner for navigasjon og visuell sceneprosessering viste at omgivelsesprosessering på tvers av nettverket er avhengig av korrekt innkoding av omgivelsene.

Samlet sett opplyser dette verket om hvordan menneskehjernen integrerer perseptuelle opplevelser i kognitive kart over omgivelsene. Det demonstrerer sterk gjenjeldelse mellom visuell koding og høyere kognitiv koding, samt belyser videre at det spatiale kartleggingssystemet understøtter domenegenerelle funksjoner som understøtter atferd.

English summary

Our brain derives stable and unified representations of the world from ever-changing perceptual inputs. This process engages a large brain network spanning from early visual to high level memory regions in the medial temporal lobe. Along this pathway, visual information undergoes a series of transformations, which ultimately culminate in the formation of the cognitive map, a non-sensory representation of the environment that guides memory formation and behavior.

In my doctoral work, I examined the neural underpinnings of spatial perception and cognitive mapping in the human brain as well as their relationship to behavior and memory. The projects presented here approach these questions from multiple angles by focusing on different computational stages along the cortical hierarchy.

With my coauthors, I examined how our visual system stabilizes perception when we move. We investigated this by monitoring human brain activity with functional magnetic resonance imaging (fMRI) during a visual tracking task (VTT). We found that a large network of regions encoded visual motion independent of our own movements, including even the earliest visual cortices in the brain. This mechanism anchors our visual perception in space. Next, we investigated how high level regions represent the space we see. Again using fMRI and a VTT, we examined grid-cell like activity in the entorhinal cortex, an area known to map space during navigation. We found an fMRI-signature that resembled this process using eye movements alone, providing the first evidence that the human entorhinal cortex encodes a non-sensory map of visual space as well. The final project then bridged across sensory and non-sensory spatial representations and explicitly focused on the behavior they support. I developed a state-of-the-art predictive modeling framework to study naturalistic behavior using virtual reality and 7T-fMRI in humans. Examining directional tuning in visual scene processing and navigation regions revealed that network-wide environmental processing depends on the successful encoding of the environment.

Collectively, this work illuminates how the human brain integrates perceptual experiences into cognitive maps of the environment. It demonstrates strong reciprocity between visual and high level cognitive coding and further highlights that our spatial mapping system supports domain-general functions in the service of behavior.

Acknowledgements

Christian Doeller. Thank you for your guidance, the advice and the support you have given me, for your drive, your absolute commitment and your optimistic attitude that creates so many opportunities for everybody around you. Thank you for being an inspiring mentor and for exciting and wonderful doctoral years!

Andreas Bartels, **Andreas Schindler** and **Michael Bannert**, you showed me what it means to be a researcher and laid the foundation for everything else to come. Most importantly, you taught me what a privilege it is to be in science and encouraged me to go on this journey. I thank you!

Edvard Moser, thank you for co-supervision and along with **May-Britt Moser** for creating such an amazing institute I feel honored to be part of. I thank all **friends** and fellow researchers, who shared their thoughts, passion and time with me and made Trondheim an incredible place. My deep appreciation goes to the **technical and administrative staff**, without whom none of the science would be possible.

A huge thank you goes to **everybody in the group**, where I always felt like being among best friends. The team spirit and social atmosphere we created together is extraordinary and I never took it for granted. I cherish every experience we shared and learned so much from so many of you. You made this time unforgettable and I look forward to many long-lasting friendships!

Markus Frey. Thank you for being the best office mate I could have ever had. Your enthusiasm, positive vibe and 75° C graphics card created the warm and welcoming atmosphere that made me love every moment with you in the office.

Tobias Navarro Schröder and **Jacob Bellmund**. You were such an integral part of the Donders-Kavli experience that I cannot overstate your importance to me in these past years. Tobias, your balanced and calm aura was the steady pulse that guided me through noisy times. Jacob, you are a brilliant and inspiring mind from whom I learned incredibly much. The discussions with you two on science and life always helped me to focus and to reflect and motivated me to push forward.

Above all, I am deeply grateful to my **family**, especially **my wife Corinna, my parents Ute and Hartmut and my sister Mareike**. You supported me wholeheartedly and unconditionally through many stormy weathers. Corinna, you joined me in this adventure through multiple countries, making difficult decisions and sacrifices along the way. I love and thank you all from the bottom of my heart.

List of publications

Paper 1:

Real-motion signals in human early visual cortex.

Nau, M., Schindler, A., and Bartels, A. (2018).

NeuroImage 175, 379–387.

Paper 2:

Hexadirectional coding of visual space in human entorhinal cortex.

Nau, M., Navarro Schröder, T., Bellmund, J.L.S., and Doeller, C.F. (2018).

Nature Neuroscience 21, 188–190.

Paper 3:

Behavior-dependent directional tuning in the human visual-navigation network.

Nau M., Navarro Schröder T., Frey, M., Doeller C.F. (2019).

BioRxiv

Paper 4 - Extended discussion

How the Brain's Navigation System Shapes Our Visual Experience.

Nau, M. & Julian, J.B., and Doeller, C.F. (2018).

Trends in Cognitive Sciences 22, 810–825.

Other publications & resources:

Deforming the metric of cognitive maps distorts memory.

Bellmund J.L.S., de Cothi, W., Ruiter T.A., **Nau M.**, Barry C., Doeller C.F. (2019).

Nature Human Behavior

Functional imaging of the human medial temporal lobe -
A neuroscientist's guide to fMRI-pulse sequence optimization.

Nau, M. (2019).

Open Science Framework

Memory efficient brain tumor segmentation using an autoencoder-regularized U-Net.

Frey, M., **Nau, M.** (2019).

arXiv

Entorhinal cortex minimises uncertainty for optimal behaviour.

Navarro Schröder T., Towse B.W., **Nau M.**, Burgess N., Barry C., Doeller C.F. (2018).

BioRxiv

Contents

Chapter 1

General introduction

1.1 Aim of this thesis

The ability to integrate sensory information into stable mental representations of the world is a fundamental basis of human cognition. We acquire information with our senses and use it to construct an internal model of the environment and ourselves in it. This process engages a large network of brain regions ranging from early visual cortices to high level mnemonic areas in the medial temporal lobe. Along this pathway, sensory information undergoes a series of transformations that ultimately convert it into world-centered non-sensory representations of space. Together, these high level spatial representations form a 'cognitive map' of the environment and guide memory formation and behavior.

How does the brain derive stable representations of the world from sensory inputs? In my doctoral work, I addressed this question by investigating the neural underpinnings of spatial perception and cognitive mapping in humans. Moreover, I examined the relationship between these processes and the behavior they are thought to support. All thesis chapters address these central questions, but approach them from various angles focusing on different levels of the cortical hierarchy. Because vision is our most prominent means to explore, I use viewing as a model for exploration in general. The first project focuses on computations in early visual cortices that stabilize our perception during movements. The second project highlights how high level mnemonic regions represent non-sensory maps of visual space. Finally, the third project spans across these computational and cortical levels and examines how behavior and memory influence environmental coding in visual and high level mnemonic regions.

This thesis bridges multiple long standing lines of research by examining spatial perception and cognitive mapping from a broad, behavioral and network-level perspective. Before discussing empirical work, the following paragraphs will introduce these topics in three steps. First, I will describe how visual areas represent spatial information during *perception*, how it is stabilized during movements and how environmental features are extracted. I will then introduce the *cognitive map* and how mnemonic areas encode the environment in a non-sensory representational format. Finally, I will *connect the dots* by showing how sensory and non-sensory representations of space combine to form a stable and unified representation of the environment. I will argue that our visual and memory systems are strongly intertwined and that viewing and navigation are guided by a common medial temporal mechanism that allows us to explore the world.

1.2 Perception: the sensory representation of space

1.2.1 Retinotopy

Vision is an active process. We scan our surroundings with our eyes and focus perceptual sampling on features of the world that seem relevant to us. With every eye movement, the image on our retina changes completely, making it necessary for visual information to be transformed into stable representations that anchor our perception in space. This engages a large portion of the brain, starting with the early visual cortex.

The importance of this area for visual perception has long been known. Published in 1881, Hermann Munk lesioned the occipital lobes of dogs and monkeys. He was the first to observe visual deficits following lesions of the particular brain area later known as the early visual cortex (EVC) [1]. By studying his animals over a long time post lesioning, he could also observe partial recovery of their visual abilities, pointing to cortical plasticity within the visual system very early on. Only a few years later, the physician Salomon Henschen described that all his patients with damage in the occipital lobe, or more specifically in the calcarine fissure, had hemianopsia: they were blind in parts of the visual field [2].

While such lesion studies showed for the first time that the occipital lobe was important for vision, they did not explicitly map the cortical representation of the visual field. This was done in the early 20th century by Tatsuji Inouye [3] and Sir Gordon Holmes [4]. They studied soldiers with ballistic head injuries and systematically related the position of the cortical damage to the position of the blind

spots they perceived. They found that neighbouring locations in the visual field are represented at neighbouring locations in the cortex as well. This cortical organization of visual information is called 'retinotopic', because it follows the same organizational principle as it does on the retina. With this, they had discovered a visual field map that corresponded to the primary visual cortex (V1) in the human brain. They described cortical magnification of the fovea and that different hemispheres represent different parts of the visual field.

The first electrophysiological evidence for the retinotopic organization of the visual system was obtained by recording neurons in the cat and monkey cortex while presenting visual stimuli. This showed that there are multiple visual field maps in the brain [5] and that primary visual neurons respond only to stimuli presented at specific orientations and locations in space [6, 7]. The primary visual receptive field had been discovered, the earliest and most fundamental neuronal representation of visual locations in the neocortex. Since then, the number of reported visual field maps has grown enormously [8]. They form a recursively connected and hierarchically organized brain network in which each region contributes different computations to vision [9–12]. We now know that there are several dozens of brain areas involved in visual processing or associated motor functions [8, 11, 12], many of which are retinotopically organized [13]. In fact, the way different areas represent visual space has become an important criterion for parcellating and studying visual cortex as a whole [14, 15].

In humans, ample neuroimaging studies showed that retinotopic representations can be studied using functional magnetic resonance imaging (fMRI) or magnetoencephalography (MEG). In order for this to work, some functional characteristic of the underlying neurons needs to translate to a population-level activity bias that can be assessed. Importantly, and unlike other cell types we will discuss later, this is the case for retinotopic neurons. Specifically, a subset of the cells does not represent all visual locations homogeneously [16, 17], which induces a stronger fMRI signal when a stimulus is displayed at certain locations that are overrepresented by the population [14, 15]. The resulting population receptive field (pRF) of an fMRI voxel is thought to approximate the cumulative sum of the underlying single cell receptive fields [18, 19]. Today, this circumstance is being utilized routinely in a great number of studies to map visual field representations across the human brain. This includes the early visual cortex [14, 15, 18], the superior occipital lobe [20], the ventral visual stream [21], the intraparietal sulcus [22], the lateral occipital cortex [23], the medial parietal lobe [24, 25], the thalamus [26] and the cerebellum [27]. Modern computational approaches to study pRF's in fMRI have

recently been extended to estimate pRF sizes that approximate the ones of single cells [28]. Inspired by these [18] and related approaches [29–33], I will introduce a similar computational approach in Chapter 4 to study spatial behavior in human fMRI.

In sum, many visual cortices represent retinotopic maps of visual space. Together they form a hierarchical network of areas that extracts and processes increasingly complex information from the visual scene. This retinotopic organization of space is useful, because it might help to bind co-occurring visual features together [34–36]. However, a purely retinotopic system would face a critical challenge: Each movement shifts the retina relative to the environment and changes the retinal representation completely. We move our eyes multiple times per second [37], meaning that the receptive field onto which a given feature is projected also changes perpetually. To maintain a stable perception of the world during self-motion at least some visual representations need to be invariant to these movements. Such representations are called non-retinotopic because they do not move when the retina does. How they are computed will be explained in the following paragraph.

1.2.2 Visual stability

Non-retinotopic representations of space are abundant in the brain [38]. They are useful for the planning of actions and anchor our perception during movement [34]. A typical experiment to study non-retinotopic representations employs a step-wise saccade task, in which two objects are presented successively at the same spatial location separated by a saccade that shifts their retinotopic positions apart. The fact that the two objects interfere on a perceptual level strongly suggests that our vision indeed relies on non-retinotopic coding [34].

There are multiple terms in use to describe the reference frames of non-retinotopic codes. This thesis uses following definition of the terms (Fig.1.1). 'Head-centered' or 'craniotopic' describes all coordinates that are referenced to our head. Head-centered receptive fields are invariant to our own eye movements, but do not move when environmental features do. In contrast, 'world-centered' or 'spatiotopic' coordinates are referenced to the external world explicitly. They could be referenced for example relative to landmarks such as objects or geometric features of the environment. When the reference landmark moves, world-centered receptive fields move along. They do not move when our eyes or head does.

Importantly, all visual information we acquire is retinotopic, making it necessary for the brain to reconstruct non-retinotopic representations from retinotopic ones. This process comprises multiple interconnected mechanisms that are implemented along the cortical hierarchy [38–42]. One important computation is the integration of incoming visual information with internal motoric self-motion signals [43, 44]. These signals are called efference copies, or corollary discharge [41], because they provide a copy of the motor commands that produced the movement. The perceptual role of efference copies is easy to demonstrate. Closing one eye and gently pushing the eyelid of the other with the finger leads to a movement of the perceived visual environment. When the eyes move naturally instead the environment stays stable. Hermann von Helmholtz argued that pushing the eye does not produce an efference copy and hence no compensatory anti-movement effect on perception [43]. This mechanism, or more specifically its neural implementation, received strong experimental support in the past decades [39].

In the context of eye movements, efference copies originate in the superior colliculus, which performs basic but fast image decompositions to guide the eyes across the scene [40, 42, 45]. This signal is then being sent to various oculomotor control regions such as the frontal eye fields (FEF), but also to occipital and parietal lobe regions to modulate perceptual processing directly [39]. Knowing how much the eyes (will) move enables at least four important computations. First, perceptual sensitivity can be reduced to avoid processing blurry information corrupted by movement. Intriguingly, this desensitization of perception at the time of saccades has been observed for both the spatial and the temporal domain [46]. Second, it can be used to compensate self-induced visual motion, a computation implemented in so called 'real-motion neurons'. They were first described in the monkey brain [47,48] and fire only when incoming motion signals were induced by movements in the external world, not when they were self-induced. Third, it can trigger predictive spatial updating, or retinotopic remapping [49], which allows some receptive fields and attention to 'jump ahead' or stretch towards the spatial position they will be occupying after the following saccade. This phenomenon has been observed for example in lateral parietal neurons [50, 51], which encode space in a head-centered frame of reference. Their receptive fields remain fixed at one location in space independent of the eye position [50]. A similar remapping phenomenon has also been described in FEF and V4 neurons, which shift their receptive fields towards the saccade target before the saccade is executed [52]. Finally, efference copies and bottom-up proprioceptive signals [53, 54] together might generate 'gain fields', a modulation of retinotopic neuronal activity by gaze

position. This modulatory gain field effect has been observed in multiple brain regions and modulates the activity of visual neurons even in the dark [55–59]. Proprioceptive signals might also compensate for the consequences of ocular torsion on V1 receptive fields [60].

In humans, comparably few studies examined non-retinotopic coding in the visual system. However, a growing body of literature does suggest that it can be studied with fMRI as well. Visual motion velocity for example has been shown to be encoded in world-centered coordinates in areas V3A and V6 [61], paralleling earlier findings in monkeys [62, 63]. The authors suggested that V3A showed a near complete integration of the incoming motion signal with efference copies [61]. In the cortical hierarchy, V3A and V6 are situated on an intermediate level in between early visual areas and higher-level visual cortex and the parietal lobe [12, 64]. They were proposed to monitor the location of objects in the visual scene to support the motor network in the planning of behavior [64]. Similar non-retinal motion responses have also been shown in the visual area within the cingulate sulcus (CSv) [65–67] and the human motion complex hMT+ [68]. In Chapter 2, I will provide an overview of such world-centered motion responses across the visual system and show that even the earliest stages of visual processing in the brain are involved.

To investigate gain fields, several fMRI studies probed whether visual responses are modulated by gaze position in visual areas [68–71]. A typical fMRI experiment to study non-retinotopic coding utilizes fMRI adaptation, which measures the decrease in activity that follows a repeated exposure to the same stimulus [72]. McKyten & Zohary for example presented objects repeatedly at the same location while having participants fixate at various locations on the screen. They observed adaptation in mid-level visual regions such as the lateral occipital cortex [73], an effect that was later shown to likely rely on active vision and eye movements [74] and needs time to build up [75, 76]. Golomb and colleagues showed that the parahippocampal place area (PPA) shows fMRI adaptation only when the eyes move, not when participants fixate and the stimuli are moved instead [74], demonstrating that the effect is independent of the retinotopic location. By comparing multivariate pattern analyses of fMRI responses between free and gaze-constrained movie watching, others showed that many such high-level regions encode information in a non-retinotopic frame of reference [77]. This is consistent with the idea that the processed information becomes increasingly abstract at higher levels of the cortical hierarchy (for review see [11, 12]). Intriguingly, after a stimulus has been presented at a given location, the parietal cortex of monkeys

encodes a head-centered memory trace of that stimulus, keeping this location in memory for some time [78]. One fMRI study showed that such memory traces are being remapped across eye movements also in the parietal cortex of humans, even if the representation needs to be remapped to the opposite hemisphere [69].

Importantly, spatial updating and self-motion compensation are imperfect and inevitably affected by noise. Retinotopic remapping alone for example would hence lead to the accumulation of updating errors, resulting in a shifting representation of space during movement. One mechanism that could correct such errors post-hoc is 'transsaccadic integration', or 'transsaccadic memory', which allows to detect visual features in both the pre-saccadic and the post-saccadic visual image. The location of salient features can then be matched between the two [79, 80]. Shifting some visual receptive fields back to where they have been before the movement could hence mediate perceptual stability post-hoc. Again, the parietal cortex might be crucial in this process, which might relate to modulations of retinotopic parietal responses by non-retinotopic memory [78]. Substantial empirical evidence suggests that our visual system does indeed use transsaccadic memory to anchor our perception in space (e.g. [81–86]). However, the simple fact that other mechanisms such as retinotopic remapping and self-motion integration exist suggests that perceptual stability does not build on transsaccadic memory alone. It might rather serve as an error correcting mechanism for receptive fields that were imperfectly remapped.

Notably, some studies did explicitly not observe any non-retinotopic coding and suggested that the native coordinate system of visual cortices is exclusively retinotopic [87,88]. Further investigations however showed that this depends on task demand and that non-retinotopic coding is not being maintained continuously [70]. Instead, such representations are likely recruited only when needed by an attentional mechanism to meet the task at hand [38]. Many areas for example switch between retinotopic and non-retinotopic coding depending on whether attention is being focused on the fovea or not [70]. Consistently, some parietal neurons flexibly switch between reference frames to encode visual and gaze related information depending on task demand [89]. This is also consistent with the observation that spatiotopic representations require time to be formed [75, 76]. Together, this suggests that attention recruits non-retinotopic representations when needed, an idea to which I will come back in the Chapter 5.

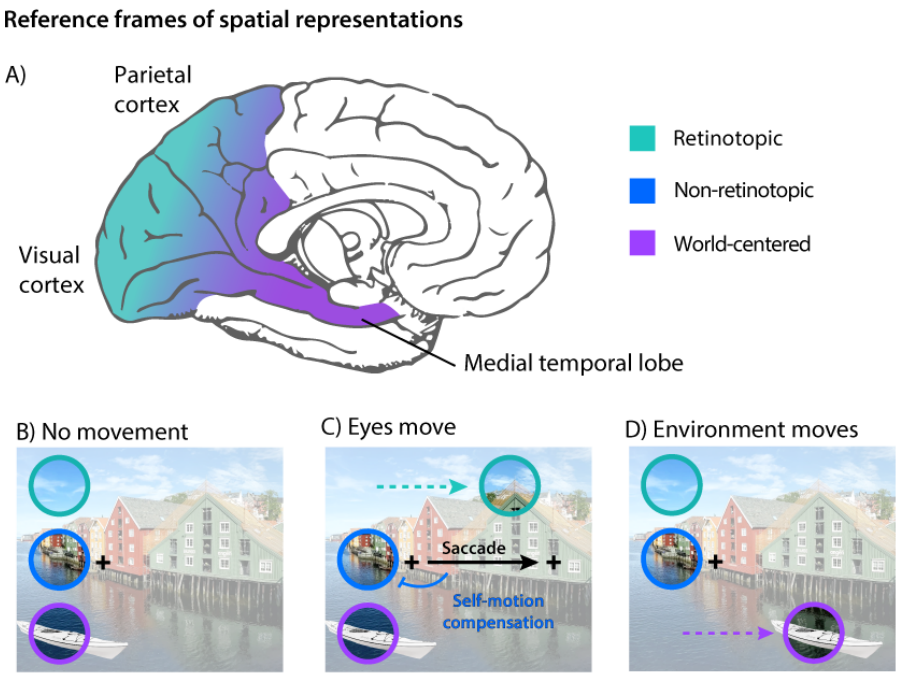


Figure 1.1: Spatial reference frames A) Visual reference frame transformation pathway. Predominant reference frames are color coded: Retinotopic (representation changes during movements of the eyes or head), Non-retinotopic (representation remains stable when the eyes move) and world-centered (receptive fields are referenced to the external world explicitly. World-centered coordinates are invariant to self-motion of the eyes or head and only move if the environment does). B-D) Movement conditions: B) no movement, all receptive fields appear similar. C) eyes move. Retinotopic receptive fields move along. D) environmental feature moves. World-centered receptive fields move along.

Similar visual, motor and proprioceptive integration mechanisms might be in place to compensate not only for movements of the eyes, but also for the ones of the whole body. This likely involves information about the current state of our body and more complex movements in space. Again, the interaction between different sensory and motor cortices likely relies on networks involving the superior colliculus and the parietal cortex [90]. The latter has access to retinotopic [13] and non-retinotopic [38] visual representations, encodes information about body posture [91] and movement [56, 92], and is generally well connected to sensory and motor [12, 64, 93, 94] as well as to higher order mnemonic brain areas [95]. Importantly, the fact that some receptive fields are invariant to our own movements does not imply that they are referenced to the external world. In order to

be considered as world-centered, the receptive fields would need to move whenever the environmental feature does to which they are referenced to (Fig.1.1). The parietal cortex has rather been proposed to provide viewpoint dependent visual information to downstream areas such as the retrosplenial cortex [96].

Together, there are multiple interconnected mechanisms that allow visual receptive fields to represent visual space independent of one's own movements. Retinotopic remapping, the integration of motor efference copies and proprioceptive signals allow to reconstruct non-retinotopic, head-centered receptive fields when needed. Transsaccadic memory might serve as a post-hoc error correcting mechanism. The combination of these mechanisms stabilizes receptive fields during self-motion and hence anchors our perception in space. The following paragraph will describe how high level visual cortices use this stable vision to extract the environmental layout and track the direction of objects as we move [97–99].

1.2.3 From scenes to spaces

A stable spatial perception enables to extract more complex information about the environment from the visual scene. Landmarks and boundaries for example can be recognized and their relative direction, position and distance can be tracked over time (Fig.1.2). Human fMRI and electrophysiological recordings in monkeys revealed that a network of higher order brain areas is involved in this process, together referred to as the high level visual cortex. Among such regions are face selective [100, 101] and object selective cortices [102] as well as regions encoding visual scenes [99]. In the context of cognitive mapping, the latter ones are most relevant here.

Scene perception regions include the retrosplenial cortex (RSC) [96], the occipital place area (OPA) [103] and the parahippocampal place area (PPA) [104]. These areas represent the spatial layout of a scene [104], the 3D surface structure [29], the relative openness [105] and the presence of boundaries [106] as well as navigational affordances [107], that is where an observer would be able to move in the current scene. They overlap with known retinotopic maps [13], even though there is no one to one correspondence between the two concepts [99]. Intruigingly, the selectivity to scenes in these regions is not a consequence of encoding low level features typically associated with scenes. Rather, responses directly reflect whether a given stimulus is perceived as a scene or not [108], also likely independent of the retinotopic position of each feature [77].

Importantly, these regions are not only critical for recognizing which landmarks to rely on, but are also an integral part of the coordinate transformation circuit that interfaces between self-centered and world-centered representations in the brain [96–98]. They likely receive head-centered visual input from parietal neurons, which allows them to be directionally selective, and represent it conjunctively with positional information they receive from the downstream hippocampal formation [109]. Consistent with this idea, the RSC encodes facing direction in rodents [110–112] and also the homologue region in humans has been shown to be directionally tuned [113–115]. Indeed, lesions of the RSC impair the ability to orient oneself relative to landmarks [116]. Furthermore, multiple recent studies showed that the rodent RSC, the striatum, the postrhinal (parahippocampal) cortex and the hippocampal formation contain cells that fire when boundaries or the center of the room are located at a given distance and direction relative to the animal [117–120]. The RSC and parahippocampal regions are well connected to parietal cortex [121, 122] as well as to medial temporal lobe regions, where some of the strongest evidence for world-centered representations of space have been found [123, 124].

Taken together, the scene perception network is a key transformation stage of visuospatial information in the brain [97,99]. Areas such as the retrosplenial cortex and the parahippocampal (postrhinal) cortex transform head-centered, viewpoint dependent representations into a conjunctive directional and positional code [98, 109]. This allows us to orient ourself relative to the environment. Self-centered visual representations are likely mediated by parietal neurons, which combine in scene perception areas with non-sensory spatial representations found in the hippocampal formation [93, 97, 98, 125–127]. Scene processing areas reconstruct world-centered coordinates from head-centered ones, ultimately providing the basis for anchoring our sense of location and direction in space.

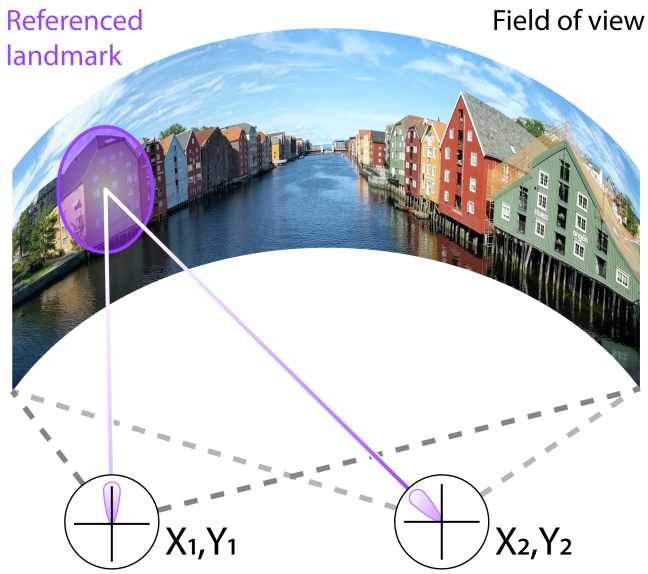


Figure 1.2: The scene perception network extracts environmental information from the scene. This allows to track the relative direction and distance of objects or boundaries (e.g. the yellow house on the left side of the Trondheim old town bridge) from different positions (X,Y) and as we move. This is signaled by cells in retrosplenial and parahippocampal areas, which interface between self-centered and world-centered reference frames likely via a conjunctive representation of place and direction.

It is important to note that all retinotopic, head-centered and world-centered spatial representations mentioned so far are still sensory in nature. Without the ability to store them over time, they would fade away quickly after visual inputs stop. In the following, I will discuss spatial codes in the hippocampal formation that are referenced to the external world just like the ones in the scene processing network, but arise from intrinsic network dynamics and do not reflect a sensory stimulus per se. These high level spatial codes form a non-sensory world-centered representation of space referred to as the cognitive map.

1.3 The cognitive map: non-sensory representations of space

1.3.1 Spatial mapping in the hippocampal formation

Decades of navigation research revealed a 'zoo' of world-centered spatial representations within the rodent hippocampal-entorhinal circuit [123, 124]. In a typical experiment, rats forage for food in an open box while electrophysiological measures of neuronal activity are recorded. This has for example led to the discovery of place cells [128, 129], a spatially tuned cell type in the hippocampus that fires when the animal occupies a certain position in the box. The hippocampus is an evolutionary old brain region present in all vertebrate species [130]. Likewise, place cells have been observed in many species such as mice [131], rats [128], bats [132–134] and humans [135, 136]. Because each place cell prefers a different location, the population of cells covers the entire space equally and is thought to represent a cognitive map of the environment [129, 137]. This map comprises different levels of granularity, which is reflected in the receptive field size of place cells varying across cells [138]. Place cells are controlled by visual cues [139] as well as by path-integration mechanisms that update the representation when the animal moves [140]. When it enters a new environment, place cell receptive fields remap to a new location independent from the previous one [141–145]. On the population level, this creates a unique place cell map for each environment. Importantly, unlike retinotopic neurons, neighboring place cells do not represent neighbouring locations in space. Interestingly, evidence from neuroimaging studies still suggested that place-cell like population responses may nevertheless exist [146–149]. These studies used virtual reality to have participants navigate in a virtual environment using visual cues only. Successfully decoding the spatial location of the participant from hippocampal voxels has been interpreted as evidence for the detectability of the hippocampal place cell code with fMRI. These conclusions however have been challenged and the underlying mechanism remains controversial. When visual and path-related confounds as well as statistical shortcomings are corrected, no place-cell like hippocampal coding could be observed [150]. Interestingly, the receptive field sizes of place cells do follow a cortical topography along the dorsoventral hippocampal axis [138].

In addition to place cells, also grid cells have been shown to represent a map of navigable space [151]. They can typically be found in the entorhinal cortex, an area adjacent to the hippocampus that is responsible for the major part of cortico-hippocampal communication [152]. Strikingly, grid cells fire not at one location,

but at multiple ones that tesselate space with a hexagonal lattice [151]. Similar to place cells, the relative position of this grid pattern as well as the grid field size differ across cells [153]. This provides a unique population activity vector for each location in space [154]. Again, the differences in field size and spacing between fields is topographically organized [153, 155]. Grid cells combine visual and self-motion cues to maintain their firing [154,156] and are thought to provide a stable spatial metric for the hippocampal place cell map [157]. Notably, in contrast to place cells, grid cells likely have functional characteristics that translate to a population response accessible by neuroimaging. As mentioned above, they represent self-location using a hexagonally organized firing pattern tesselating space. The orientation of the grid pattern relative to the environment is shared across cells [153]. On a population level, this has been hypothesized to give rise to sixfold rotationally symmetric activity biases as a function of running direction [158]. First evidence for the presence of this (hexadirectional) population signal has been described in human fMRI studies, which examined entorhinal fMRI-activity while participants navigated in a virtual reality [158]. Here, the entorhinal signal amplitude varied across running directions and followed the hypothesized 60°-modulation. Since then, a large number of studies reported similar hexadirectional modulations of entorhinal activity during virtual navigation [159–167] but also during non-spatial tasks [168, 169]. In Chapter 3, I will discuss how similar principles might allow to study entorhinal codes of visual space as well [170–172]. Importantly, the presence of grid cells in the human brain has been confirmed by electrophysiological studies [172–174].

Place cells and grid cells represent self-location and therefore indicate where we are. However, for spatial orienting it is also necessary to know which way we are facing. Head direction cell firing indicates our head direction relative to the environment and has been observed in many regions in the rodent brain [175] including the anterior thalamus [176], the postsubiculum [177], the parahippcampal region [178], the retrosplenial cortex [110–112] and the medial entorhinal cortex [178–180], but also in monkey [181]. Head direction cells are thought to form a 'neural compass' that mediates our sense of direction and guides path integration and reorientation behavior [175]. Interestingly, there are at least two types of head direction cells in the brain: 'classical' non-sensory head direction cells as well as sensory head direction cells [178]. The firing of non-sensory head direction cells is referenced to the environment but does not reflect a stimulus per se. It is maintained even in the absence of visual cues over some time (but is then subject to drift) [177, 182, 183]. In addition, sensory head direction cells in the

parahippocampal [178] and retrosplenial cortex [112] can switch between active and inactive states. In humans, multiple studies investigated directional representations with fMRI. These typically use directional judgment and mental imagery tasks in combination with virtual reality to make participants mentalize certain directions. Consistent with observations in rodents, evidence for the encoding of direction has been observed for example in the retrosplenial [113, 114, 184, 185] and the parahippocampal cortex [114, 158, 162, 186–190], the hippocampal region [186], the entorhinal cortex [113, 158, 185, 189, 191], the thalamus [113] and the superior parietal cortex [192, 193]. Whether these directional representations reflect the activity of head direction cells however remains unknown. Chapter 4 will elaborate more on this topic and investigate human directional tuning for the first time during active navigation.

Place cells, grid cells and head direction cells are all integral parts of the neural implementation of the cognitive map [129]. They are complemented by other spatially tuned cells such as entorhinal border cells [194, 195], which fire along environmental boundaries, subicular boundary vector cells [196] and entorhinal object vector cells [197], which fire at a given distance and direction to boundaries and objects, as well as entorhinal speed cells, which track the running speed of the animal [198].

It is important to note that many of the spatial coding principles in the hippocampal formation sound similar to the ones observed in visual cortex on a superficial level, yet the receptive fields, their cortical organization and most importantly the presumed nature of their representation are drastically different. Place cells, grid cells, border cells and head direction cells mediate an explicitly non-sensory representation of space. Instead of 'responding' to a specific combination of cues, their firing pattern likely arises from the recurrent activity within the hippocampal-entorhinal circuit. This connectivity gives rise to attractor dynamics [199, 200].

An attractor network is a recurrently wired network of neurons in which each cell excites proximal cells while inhibiting distal cells [199–205]. This leads to network activity that converges onto a stable set of activity states. If the transitions between these states are continuous, the network is referred to as a continuous attractor. Substantial evidence points to such continuous attractor dynamics in head direction cells [200, 206, 207] and grid cells [156, 202]. In grid cells, this idea has received empirical support by multiple recent studies demonstrating that the pairwise cell-to-cell correlation structure across cells remains stable across be-

haviors [208] and even during sleep [209, 210]. This suggests that it is indeed the network connectivity within the entorhinal cortex [152] that gives rise to their firing, rather than external inputs. Similar results have been observed for head direction cells, whose population dynamics follows ring attractor dynamics during walking and during sleep [207]. Strikingly, in fruit flies this head direction ring attractor is anatomically organized in the shape of a ring [211, 212].

In sum, the hippocampal formation contains spatially tuned cell types, which reference self-location or direction relative to the environment. They rely on sensory cues to be anchored to the environment but do not represent a visual stimulus per se. Together, the cells in the hippocampal-entorhinal circuit form a cognitive map of the environment.

1.3.2 A cognitive map of visual space

As described above, most of our knowledge about the cognitive map builds on electrophysiological studies in rats and mice [123, 124]. Rodents however perceive the environment very differently than humans do, relying less on vision and more on their somatosensory and olfactory senses [213–215]. Humans, like other primates rely predominantly on vision and eye movements to explore the environment [213]. This raises the question whether these differences in sensory experience and behavior are reflected also in the way the hippocampal formation represents space. A large number of studies suggested that rodents and primates indeed employ similar spatial coding principles [127, 216–218]. Importantly, while the reported cell types in rodents primarily encode self-location, the homologues in monkey often reflect where the animal is looking [127, 216, 217].

Similar to rodent place cells for example, the hippocampus of monkeys contains cells that are responsive to certain locations in the environment [219]. These neurons however encode the world-centered location of gaze [220]. These 'spatial view cells' maintain their place selectivity even in darkness for some time similar to place cells [221] or when the view is obscured with obstacles [222]. In addition, related hippocampal neurons represent gaze, self-location and head direction in a conjunctive code [223]. Spatial view cells are likely not unique to the primate brain. Recent work suggests for example that the hippocampal formation in rats and bats contains similar cells that represent locations in the environment that are not their own [134, 224].

In addition to place cells, also entorhinal grid cells have a purely visual analogue in monkeys. These neurons encode a map of visual space using a hexagonal firing pattern that strikingly resembles the one observed in rodents during navigation [151, 225]. The representation mediated by (at least some) visual grid cells is world-centered, meaning that their firing pattern moves coherently whenever the stimulus moves [226]. The fact that these patterns emerge in the analysis after pooling data across many stimuli [225] also shows that visual grid cell activity is invariant to the stimulus content (at least of the cells reported to date). Instead, it likely reflects the encoding of a 'content-free', non-sensory metric map of visual space. Also, visual grid cells do not represent the eye position in the eye orbit, but rather the location in space that is currently attended [227]. This is independent of whether the location was attended overtly or covertly [225, 227]. Whether spatial view cells are modulated by spatial attention similar to visual grid cells seems likely but remains an open question. In Chapter 6, we propose that visual grid cells guide our gaze by computing vectors between locations in space similar to the proposed analogous function of grid cells in navigation [127, 154]. Interestingly, computational models recently suggested that this mechanism might allow visual grid cells to support recognition memory [228]. Once the locations of visual features in the scene are associated with unique grid cell population activity, this can be used to compute vectors between these features, which are then forwarded to the oculomotor system.

Similar to how head direction cells encode the direction of the animal's head, entorhinal saccade direction cells encode the direction of future or past eye movements in monkey [229]. The two cell types were proposed to share an evolutionary origin [217]. Akin to rodent border cells, visual border cells in monkeys indicate the proximity of the current gaze position to visual boundaries such as for example the edges of a visual display [225]. Intruigingly, they do not fire along edges within that same stimulus, demonstrating that they do not encode the visual appearance of these edges, but rather the behavioral consequence, or task relevance, associated with it. Similar coding of the task-related 'meaning' of spatial features independent of their visual appearance has recently been observed on a broader scale in other hippocampal cells that encode space into a schema-like representation [230]. Given the strikingly similar firing properties of rodent and monkey spatial representations in the hippocampal-formation, the representations of visual space likely also follow attractor dynamics as described above. However, future studies will need to test this prediction empirically.

In Chapter 6, I will discuss the function and computational tasks these codes might solve in more detail. In short, the spatial codes in rodents and monkeys might not only be related, but could both be expressions of one common mechanism that maps the environment and guides exploratory behavior. This hippocampal-entorhinal mechanism might have been co-opted and adapted according to the sensory experience and the behavior, which the different animals express.

1.4 How sensory and non-sensory representations combine

Hippocampal-entorhinal codes mediate a cognitive map that is generated by internal network dynamics, not by sensory inputs. How can such non-sensory representations of space be anchored to the external world? To understand this process it could help to consider a practical example (Fig.1.3). As mentioned before, the firing of head direction cells for example is governed by a continuous attractor [207]. The network always maintains a certain level of population activity even though the cells that are active change over time. Because nearby cells are excited and distal cells are inhibited, the population activity forms an 'activity bump' in the network. Because head direction is a circular variable, all possible population activity states this network can occupy fall onto a circle in state space. When the animal moves, vestibular inputs bias the activity bump to switch to a new state [199,231]. In the case of head direction cells, this means that the network now encodes a new head direction. By moving the activity bump via vestibular inputs the network can do path integration, meaning that even in the absence of visual cues the current head direction of the animal can be tracked. This explains why many of the spatial cell types can maintain a certain activity level and with it a representation over some time.

However, path integration has at least one limitation: it accumulates error with every update [156]. Just like retinotopic updating, the vestibular self-motion integration process is not perfect and subject to noise. The representation hence needs to be anchored frequently to the external world in order to remain useful and to avoid drift. This is likely done by visual inputs that have been processed as described in Sections 1 and 2. Computational models proposed that the anchoring is achieved by an 'outer ring' of visual cells that are directly connected to the head direction cells in the ring attractor network [199, 200, 203–205, 232]. These visual cells indicate the direction of salient landmarks or boundaries. Each visual cell provides input to very few head direction cells in the network, firing only when the observer faces towards the landmark they are referenced to. If the attractor

activity was misaligned, this visual input biases it back towards the 'correct' state. This process is called 'cue control' or 'landmark control', which simply describes the process of anchoring the ring attractor to these landmarks.

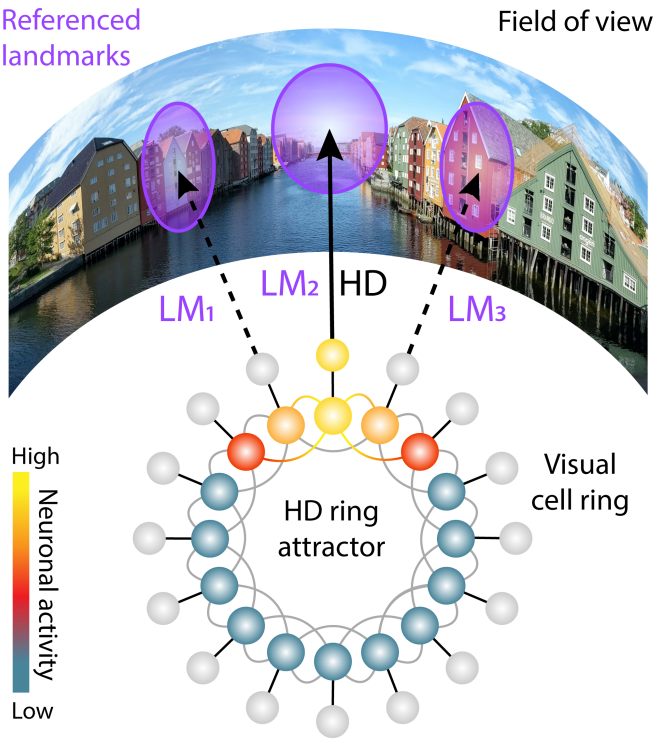


Figure 1.3: Anchoring attractor dynamics to the visual world via cue control. Head direction cells are thought to follow continuous attractor dynamics. The currently represented head direction (HD, black arrow) is reflected in the activity bump (hot colors) of the HD ring attractor ('inner ring'). A second 'outer ring' of visual cells represents the direction of environmental landmarks (LM), which could be objects or other features that define the environmental layout. When the observer faces towards a landmark (distal or proximal), the corresponding visual cell would fire and bias the activity bump in the ring attractor to represent the correct HD. In this example, the visual neuron representing the bridge straight ahead (LM2) is exhibiting cue control.

In this model, the direction of the landmarks is being tracked by upstream regions such as the retrosplenial cortex [96, 98, 99] as described in Section 1.2. Because attractor dynamics sustain the activity over time, and because path integration mechanisms can update the representations to a large degree, this anchoring mechanism does not need to be maintained continuously. Rather, it could be recruited as needed by additional attentional mechanisms for example during

reorientation [233]. This would be in line with the observation that also non-retinotopic representations are non-continuous and modulated by attention in visual cortices [38, 70]. It is also in line with parietal neurons switching reference frames depending on task demand [89] and with the fact that non-retinotopic representations need time to build up, because the above-described sequential computations first need to be conducted.

In the case of grid cells, the possible network states would not occupy a ring in state space, but instead a torus, reflecting the 2D spatial representation they mediate [151, 154, 202]. Note however that there is evidence that such spatial codes might also represent 3D space [187, 234–236] or even higher dimensional non-spatial variables [168, 169, 237–239]. Hippocampal activity likely follows similar attractor dynamics as suggested by computational models [240, 241] and empirical findings in rodents [242] and humans [243]. Notably, while hippocampal place cells typically remap when an animal enters a new environment [141–145], some of them, not all, also remap when only few of the environmental features are altered [143]. This phenomenon is called 'partial remapping' and shows that the population activity of place cells does not always converge on one stable state. This shows that there is not only one attractor network controlling all spatially tuned cell types. Instead, there might be multiple ones, likely including continuous and discrete attractors to allow smooth transitions between states while still encoding our experiences using high-dimensional and orthogonal representations [205, 244–246].

In sum, the spatial representations in the hippocampal formation are fundamentally different from visually evoked responses or from self-motion compensated non-retinotopic ones in visual cortices. Instead, they mediate representations of space that are referenced to the external world, but are non-sensory in nature. They build on the intrinsic hippocampal-entorhinal network connectivity, which gives rise to attractor dynamics. Visual input supports the anchoring of the representations to the external world. Collectively, these representations are thought to form the basis of the cognitive map [137, 139].

1.5 Thesis outline

The above-described mechanisms allow the brain to perceive a stable visual world and to form cognitive maps of the environment. This comprises several complex multi-sensorimotor integration processes, which ultimately interface between sensory and non-sensory representations of space. While most of our knowledge about this topic is grounded in electrophysiological recordings in rodents and monkeys, the neural systems that support a stable perception and cognitive mapping remain poorly understood in humans.

This thesis addresses the central question how the human brain derives cognitive maps from visual inputs and how they relate to spatial perception, behavior and memory. The following projects approach these topics from multiple angles and focus on several cortical and computational stages of this process.

Empirical paper 1 investigates how the human visual system integrates self-motion information to maintain a stable perception during eye movements [247]. This likey builds on the integration of visual input with motoric self-motion signals, a computation that is critical to anchor our perception in space. In addition to showing that many regions of the human visual motion network are involved in this process, this project provided the first evidence that the human early visual cortex signals visual motion velocity in a world-centered frame of reference. This suggests that extra-retinal signals modulate even the earliest stages of visual processing in the human cortex in support of a stable perception of space.

Empirical paper 2 examines how high-level regions implicated in environmental mapping represent what we see [170]. Specifically, we ask whether the human entorhinal cortex represents a grid-like map of visual space akin to its monkey homologue [225]. We tested this by examining human fMRI-proxies of grid-cell-like representations during a viewing task, which suggested that this was indeed the case. Our results demonstrate that the entorhinal cortex maps visual space with a grid-like code and provided the first evidence of a non-sensory visual field map in the human brain. Our results support the proposal of the cognitive map as a general organizational principle of neural information coding and further emphasizes its domain-general function.

Empirical paper 3 spans across these levels of processing, probing world-centered directional tuning in both low-level visual and high-level scene processing and navigation regions during active spatial behavior [248]. Here, we ask how behavior and spatial memory influence environmental coding in the human visual-navigation network. Using virtual navigation during 7T-fMRI and a newly developed voxel-wise behavioral encoding model, we demonstrated directional tuning in this network for the first time during active spatial behavior. More importantly however, our results revealed that this directional code depends on the behavior of our participants as well as on how well the environment has been encoded. This demonstrates a direct link between neural population tuning and human behavior and shows that network-wide perceptual processing directly interacts with the high level cognitive mapping process.

After presenting this empirical work, I will discuss its implications in the 'General discussion' Chapter 5 as well as in the 'Extended discussion' Chapter 6. The latter has been published as **Review paper 4** [127] and reviews the growing body of literature implicating the hippocampal formation in representing non-sensory map of visual space. The underlying cell types provide the optimal solution to multiple computational challenges that are shared between navigation and vision. They may provide world-centered coordinates for the encoding of space, the planning and guidance of behavior as well as for the formation and recall of episodic and contextual memories. I will propose that these high-level visuospatial codes shape perceptual processing on a large scale and that navigation and viewing are two behavioral expressions of a common spatial mapping and exploration system in the medial temporal lobe.

Methods Box 1

Functional magnetic resonance imaging (fMRI)

In order to study human brain activity non-invasively I use fMRI. As discussed throughout this thesis, fMRI is well suited to examine spatial representations in the brain and has been successfully used to study the retinotopic [13] and non-retinotopic [38] organization of visual cortices as well as spatial representations in the medial temporal lobe [165]. It measures changes in the oxygenation and the volume of blood, which typically follow the dynamics of neural population activity for following reason [249].

Neuronal signaling requires energy, which is produced via glycolytic oxygenation in the tissue of need. The local oxygen storages are too small to support the activity of larger populations of neurons and need to be refilled after periods of enhanced activity. This leads to increased blood flow supplying the tissue with fresh oxygen. Both, the change in blood-oxygenation level and in blood volume affect the local magnetic field strength of a region, in turn inducing currents in receiver coils placed around the participants head. These changes in current can be used to reconstruct a 3D image of the brain, digitally sliced into thousands of voxels. By monitoring each voxel over time, one obtains signal time courses that serve as a basis for all imaging analyses presented in this thesis. The experimental paradigms themselves were presented on a screen placed outside the scanner bore, which participants could see via a mirror display.

Eye tracking

For the first two experiments it was critical to know where participants were looking and when they moved their eyes. We therefore recorded their viewing behavior online during fMRI scanning. The eye trackers were infrared-based video cameras placed outside the scanner bore, emitting low-energy light and recording the reflections of the cornea and retina. Because the spatial arrangement of these reflections is unique for each location on the screen, tracking it allowed to monitor gaze locations continuously.

Methods Box 2

Visual tracking task

In many instances, it is important not only to know where participants look, but to control their gaze behavior and therefore their experience tightly. This is accomplished for example by using visual tracking tasks, typically comprising a fixation dot at which participants are asked to fixate for the duration of the experiment. By moving the dot, one also controls what participants see. This has at least two great advantages over unconstrained viewing. First, it allows to balance the sampling of different behaviors (e.g. eye movement directions) and can therefore greatly improve the efficiency of experimental designs. Second, it reduces the number of confounding factors that need to be accounted for when interpreting the results. In the present thesis, the first two papers build on such visual tracking tasks, which were superimposed with another stimulus or task such as optic flow in project 1 and an object-location memory task in project 2.

Virtual reality

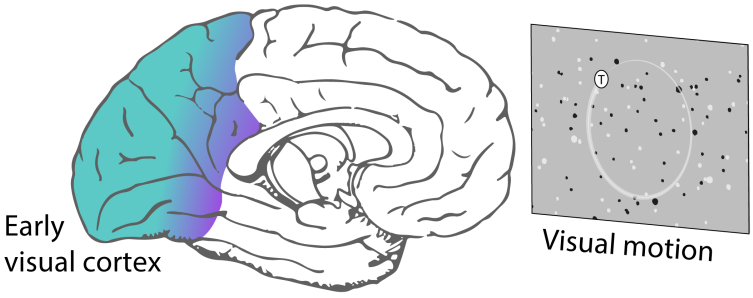
Paradoxically, the strength of the visual tracking task, i.e. the tight control over the participants behavior and experience, is also its weakness. It allows to study certain cognitive functions with high certainty, it does so however in an unrealistic, artificial setting. Our brain evolved in a highly complex and engaging environment, making the use of more naturalistic paradigms imperative when trying to understand its function. Virtual reality is a powerful tool to do so by providing a flexible and realistic environment, which is still easy to control and to monitor. Most importantly, virtual reality enabled us to record fMRI data during active navigation behavior inside the MRI machine. In project 3 and as a control condition in project 2, we used a simple virtual environment in the form of a circular arena, which participants saw from first person perspective. They could navigate freely within this arena using key presses while performing an object location memory task.

Object-location memory task

Object-location memory tasks were used in projects 2 and 3. Participants memorized and reported the location of objects either on the screen or within a virtual environment. In the course of different trials, we cued these objects, prompting the participants to either move a cursor to the location they remembered the object at (project 2) or to navigate to that location themselves and press a button (project 3 and control condition of project 2). This task not only ensured that the participants remained attentive, but also allowed to assess their spatial memory performance later on.

Chapter 2

Real-motion signals in human early visual cortex



This chapter has been published as:
Real-motion signals in human early visual cortex.
Nau, M., Schindler, A., Bartels, A. (2018).
NeuroImage 175, 379–387. https://doi.org/10.1016/j.neuroimage.2018.04.012

The data used in this work were collected at the Max Planck Institute for Biological Cybernetics in Tübingen and were partially used in my Master Thesis at the University of Tübingen, Germany. The analysis, obtaining the present results, the writing of the manuscript and the publication were all part of the doctoral work at NTNU.

2.1 Abstract

Eye movements induce visual motion that can complicate the stable perception of the world. The visual system compensates for such self-induced visual motion by integrating visual input with efference copies of eye movement commands. This mechanism is central as it does not only support perceptual stability but also mediates reliable perception of world-centered objective motion. In humans, it remains elusive whether visual motion responses in early retinotopic cortex are driven by objective motion or by retinal motion associated with it. To address this question, we used fMRI to examine functional responses of sixteen visual areas to combinations of planar objective motion and pursuit eye movements. Observers were exposed to objective motion that was faster, matched or slower relative to pursuit, allowing us to compare conditions that differed in objective motion velocity while retinal motion and eye movement signals were matched. Our results show that not only higher level motion regions such as V3A and V6, but also early visual areas signaled the velocity of objective motion, hence the product of integrating retinal with non-retinal signals. These results shed new light on mechanisms that mediate perceptual stability and real-motion perception, and show that extra-retinal signals related to pursuit eye movements influence processing in human early visual cortex.

2.2 Introduction

When objects change their position in the visual field, the brain needs to infer whether the object moved ('Objective motion') or whether the eye moved and therefore shifted the visual field relative to the object ('Pursuit Motion'). Even though both options result in retinal motion, only the former is perceived as world-centered 'Real-motion'. Previous behavioral and electrophysiological studies suggest that the discrimination between these sources for visual motion is mediated by integrating efference copies of eye movement commands with visual input, reflected in responses of real-motion neurons [47] von Helmholtz, 1867; von Holst and Mittelstaedt, 1950), allowing to separate self-induced from world-centered objective visual motion [47,48,62,250]. Similar non-retinotopic coding of visual motion and visual locations has been found in various areas in the monkey brain such as MST [251], V3A [62], V6 [252], VIP [50] and early visual areas V1 [47,60] and V2 [250].

In the human brain, comparably few studies examined objective motion responses, and a systematic overview across visual regions is still missing. A recent functional magnetic resonance imaging (fMRI) study showed that human areas V3A and V6 respond almost exclusively to planar objective motion, suggesting a near complete integration of efference copies with visual input [61]. Both areas, along with hMST have also been shown to encode objective motion when visual stimuli simulate head motion [66, 67], and the same regions encode visual stimuli in a spatiotopic (head- or world-centered) reference frame at fixed eye-positions [38, 68, 70]. Among the motion responsive regions in the cingulate sulcus, areas Pc [253] and CSv [254], the latter has been shown to encode objective motion to a limited extent [65]. Finally, there is evidence that human VIP [69] as well as several regions in the intraparietal sulcus (IPS) [255] have access to both retinal and eye movement information and might hence show objective motion responses. While early visual cortex does respond to objective motion in monkey [47,250], its involvement in the human brain has not been examined before.

To examine early visual cortex in this context and to provide the yet missing overview across the visual system, we used fMRI to investigate the responses of sixteen visual brain areas to objective motion during eye movements. We used retinotopic mapping (early visual areas V1-hV4, parietal regions IPS-0 to IPS-4) and dedicated motion localizers (hMT/V5, hMST+, V3A, V6, CSv, Pc, VIP) to independently identify these regions in every participant prior to examining their objective motion responses.

We used planar dot motion and a moving fixation disc to induce different objective motion velocities while pursuit speed as well as retinal motion were matched. Given that early as well as high level visual motion responses have previously been shown to be speed tuned [66, 256–258] we expected different objective motion velocities to yield differential responses. Eye tracking was employed throughout the experiment and participants conducted an independent task at fixation to keep attention balanced across conditions. We specifically tested which regions were involved in perceptually relevant estimation of objective motion velocity while physical input parameters were exactly matched. This estimation can only be achieved by the multi-modal integration of efference copies with visual signals, and it is essential to maintain a stable perception of the visual environment during self-motion.

2.3 Methods

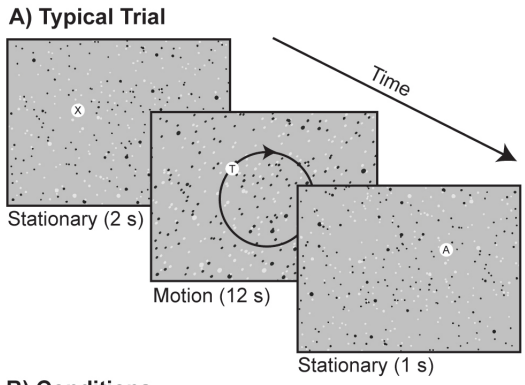
Participants

A total of 18 participants (7 male, 11 female) participated in this study (23–34 years of age, normal or corrected to normal vision, no neurological pathologies). Prior to scanning, participants were instructed about the experiment in spoken and written form, performed several test trials and gave written consent. The study was approved by the joint ethics committee of the university clinics Tuebingen and the Max Planck Institute Tuebingen.

Stimulus

The stimulus was written in MATLAB 2013b (http://www.mathworks.de) using Psychtoolbox (http://psychtoolbox.org) and was projected onto a translucent screen in a back-projection setup via a gamma corrected NEC PE401h projector. The display covered 22×16.4 visual degrees. The stimulus (Fig.2.1) consisted of a fixation disc (light grey, 0.9° in diameter) as well as a random dot pattern (black and white dots, 100% contrast) on a grey background (mean luminance: 10 participants with 487cd/m^2, 8 participants with 244cd/m^2. The luminance split between participants was due to technical reasons and unlikely to affect any question of interest). The planar random dot pattern consisted of an average of 1540 dots varying randomly in size between 0.08 and 0.2°, with all dots moving coherently and simultaneously while keeping their relative distances fixed. The random dots hence moved together as one background image, inducing global planar visual motion. Both fixation disc and the random dots in the background moved on a circular trajectory

with a radius of 4.1° (quarter of the screen height). The random dot pattern could however move at different velocities relative to the fixation disc, yet always on the same trajectory (see experimental conditions below). The motion radius was chosen such that the area of controlled visual stimulation was maximal: the nearest border to the screen edge was at all times further away than 4 visual degrees, leading to controlled visual stimulation within at least 8×8 visual degrees [61]. The rotation direction and starting point of the fixation disc was randomized and counter balanced across trials within each participant.



B) Conditions

Objective Motion [°/s] \ Pursuit [°/s]	2	3	
	1	2	***'Slower'*** than pursuit
	2	3	***'Matched'*** with pursuit
	3	4	***'Faster'*** than pursuit

Figure 2.1: Illustration of stimulus timeline and experimental conditions. A) Stimulus timeline across a typical trial. Each trial started with a stationary background and stationary fixation disc, shown for 2s. This was followed by the condition-specific motion (of background and fixation) for 12s. Note that the circular trajectory of fixation disc and background plane were identical, but that background motion could be faster, matched, or slower relative to pursuit. Trials ended with a stationary period lasting 1s. B) At two different pursuit speeds (2° /s and 3° /s) we presented objective motion that was 1° /s slower ('Slower' conditions), matched ('Matched' conditions) or 1° /s faster ('Faster' conditions) relative to pursuit. Data was pooled across pursuit speeds to obtain the 3 (pursuit-speed invariant) conditions Faster, Matched and Slower.

Fixation task

Participants were instructed to always fixate the fixation disc on which we displayed a character repetition-detection task to balance attention across condi-

tions. Random black characters (A-Z) were presented in random order with a frequency of 1Hz. Every three to eight presentations one of these characters was repeated, which participants had to report via button press. As a control measure we recorded hit rate, false alarms rate and response time across all conditions. A button press was counted as a hit if occurred within 1s after a character repetition; otherwise it was counted as false alarm. Response time depicts the time between a character repetition and the corresponding hit. For each of these measures we performed repeated-measures ANOVA to test for condition dependent effects.

Trial timeline

A typical trial started with a stationary stimulus for 2s, followed by the movement of the stimulus for 12s, ending with another stationary period of 1s (Fig.2.1). The motionless pre- and post-stimulus periods provided enough time to facilitate the saccade from the end-point of the fixation disk of the preceding trial to start-point of the current trial, as well as stable fixation before the actual motion stimulus was shown.

Experimental conditions

Pursuit eye velocity was experimentally determined through movement of the fixation disc and objective motion by movement of the background dots. Retinal motion is the difference between the two. We presented a total of nine conditions that resulted from combinations of two pursuit speeds and three objective motion speeds. Data was pooled across the two pursuit speeds (2°/s and 3°/s) to provide invariance with respect to pursuit. Hence, only three (pooled) conditions were of interest for this study: objective motion was either 1°/s slower ('Slower' conditions), exactly matched ('Matched' conditions) or 1°/s faster ('Faster' conditions) relative to pursuit. Faster and Slower conditions differed only in objective motion, not in pursuit or retinal motion velocity. Differences in responses between these two conditions hence reflect responses to objective motion alone, and constitute the key contrast in this study. An additional three conditions presented 0°/s, 1°/s and 3°/s objective/retinal motion during stationary fixation (Supplementary figure 1). These three non-pursuit conditions are not of interest for this particular study and will be neglected in the following. Each condition was presented 18 times in a pseudo-randomized, history-matched order.

Data acquisition

Functional T2*-weighted gradient-echo multiplexed echoplanar images (EPI) were acquired on a Siemens MAGNETOM Prisma 3 T magnetic resonance tomograph using a 64 channel phased-array head coil. The pulse sequence included GeneRalized Autocalibrating Partially Parallel Acquisition (GRAPPA) with factor 2 and multiband radiofrequency excitation with factor 4 (for more information see [259]). Images were acquired using the following parameters: repetition time (TR)=870ms, echo time (TE)=30ms, voxels size 2 ×2 × 2 mm^3, flip angle 56°, field of view (FOV) 192×112mm. In 4 runs a total of 2902 volumes were scanned, each containing 56 slices with 96×96 pixels. An anatomical scan was recorded for each participant using a T1-weighted ADNI-sequence with 1mm^3 voxel size.

fMRI preprocessing and statistical analysis

Data were preprocessed using SPM12b (www.fil.ion.ucl.ac.uk/spm/). All images were corrected for head motion and field distortions, spatially normalized to MNI space and smoothed with a full width at half maximum (FWHM) kernel of 4mm for regions of interest (ROI) analysis and 9mm for group analysis. The data of each participant was then analyzed using a General Linear Model (GLM). The model involved one regressor per condition as well as one for button presses. These regressors were convolved with the hemodynamic response function. The six realignment parameters (translations of X, Y, and Z coordinates, pitch, roll, and yaw) as well as the global signal of each volume orthogonalized to the rest of the design matrix were modeled as nuisance regressors to further reduce movement artifacts and global signal fluctuations induced by scanner or physiological noise [260]. Low frequency drifts were removed using a high-pass cutoff filter of 128s.

Regions of interest definition of motion responsive areas

To identify motion responsive areas hMT/V5, hMST+, V3A, V6, VIP, CSv and Pc (Fig.2.2), a separate localizer paradigm was scanned. The stimulus resembled the one used in the main experiments involving a fixation disc with attention task and random dot patterns on grey background. It differed in the actual motion pattern that was shown as follows. In total the paradigm involved seven conditions: 1) fixation on static dots, 2) pursuit on static dots, 3) pursuit and objective (planar) motion matched in their trajectories, 4) full field 3D (expansion-contraction) optic flow, 5) random motion (derived from and trajectory-matched with 4), 6) left and 7) right hemi-field optic flow.

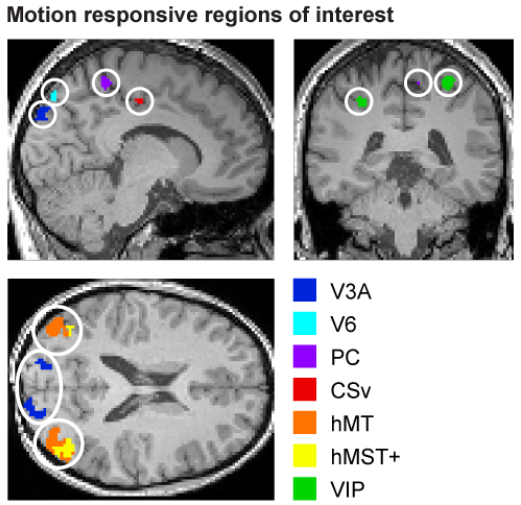


Figure 2.2: Illustration of locations of the regions of interest defined by our motion localizer on the example of a representative participant. Motion responsive regions were color coded and shown overlaid on sections of the participant's structural T1-scan at MNI-coordinates X=10, Y=-39, Z=20.

Data acquisition (one scanning run of approximately 10min), preprocessing and statistical analysis were identical to the main experiment and ROI definition was done using MarsBar region of interest toolbox for SPM [261]. To this end, t-maps of respective motion contrasts (explained in the following) were overlaid on individual structural scans and thresholded at ROI-specific p-values as amplitudes of motion responses differed substantially between areas. Thresholds were chosen so that ROI sizes were approximately matched across participants and that all areas were identifiable in all participants. ROIs were then defined by selecting significant clusters according to the following criteria. The hMT+/hV5+ complex was identified using standard methods [262] and subdivided into hMT/V5 and hMST+. Area hMST+ was defined as the ipsilateral response to optic flow versus static dots within the hMT+/hV5+ complex, leaving the contralateral response without hMST+ as hMT/V5 [262, 263]. We use the term hMST+ as several additional motion responsive satellite regions of monkey MT/V5 have receptive fields extending into the ipsilateral hemifield like those of MST that are most likely included in our hMST+ [264, 265].

We used a previously established motion localizer to robustly localize V3A [61]. Area V3A was defined as the strong response on the superior occipital lobe to pursuit with co-moving objective motion versus pursuit on a static background, corresponding to contrasting objective with retinal motion, which replicably leads to selective activation of voxels overlapping with retinotopically defined V3A [61]. While area V6 also exhibits consistent responses in this contrast [61, 252], we localized V6 [24, 61] along with areas VIP [56], CSv [65] and Pc [253] on the basis of their preference for optic flow compared to random motion. Note that the region defined here as V6 is located slightly anterior compared to previous reports [266] and most likely also includes voxels of the neighboring V6Av as this region also marginally prefers coherent compared to random dot motion [266, 267]. Detailed information about ROIs defined in this study is given in Table 1.

Retinotopic mapping

Retinotopic regions of interest were delineated in 10 participants using standard retinotopic mapping techniques [14,22]. The stimulus consisted of a wedge shaped checkerboard (90° in size, 100% contrast, 6Hz contrast inversion flicker, check sizes increased logarithmically with eccentricity) rotating around the central fixation dot (55.7s period, matched clockwise and counterclockwise rotation) in front of gray background. An attention task (button press when red dot appeared after on average every 4sat random location) was superimposed onto the checkerboard to facilitate parietal responses [22]. Image acquisition and preprocessing was identical to the main experiment except that images were not smoothed. Polar angle maps were calculated as described in previous studies [14], and ROI definitions were performed on inflated brain surfaces of individual participants. In short, we estimated the phase of an oscillatory BOLD-activity induced by the rotating stimulus and mapped it across all voxels of the visual brain. In retinotopic regions, neighboring voxels respond to neighboring visual field locations and hence estimated phases of neighboring voxels are similar. Since the orientation (among other features) of the represented visual field differs across regions, the resulting polar angle maps allow to trace borders and to delineate visual areas [17]. We used this approach to identify areas V1, V2, V3 and hV4 as well as IPS0 (V7), IPS1, IPS2, IPS3 and IPS4. Left and right hemispheres as well as dorsal and ventral parts of these areas were combined to obtain one ROI per area and participant. We did not subdivide the V3AB-complex into V3A and V3B as this would have required eccentricity maps which were not obtained in this study. Here, we identified V3A using a dedicated functional motion localizer (see above). Surface reconstruction [268], analysis of polar angle maps [14] as well as retinotopic ROI

Table 2.1: ROI coordinates in SPM normalized space (approximating MNI space) (meanSEM) for left (L) and right (R) hemispheres, number of voxels for each ROI(meanSEM), and number of hemispheres in which each ROI was found. Note that not all areas were localizable in all participants and hemispheres.

Brain region	Hemisphere	X [mm]	Y [mm]	Z [mm]	#Voxels	#Hemispheres
Functional Motion Localizer (n = 18 participants):						
V3A	L	-16 ± 4	-91 ± 4	25 ± 5	107 ± 7	36
	R	19 ± 4	-89 ± 4	27 ± 5		
V6	L	-14 ± 5	-78 ± 4	40 ± 7	100 ± 11	35
	R	17 ± 6	-77 ± 4	42 ± 5		
CSv	L	-10 ± 3	-19 ± 4	42 ± 3	15 ± 2	31
	R	10 ± 3	-19 ± 4	45 ± 3		
Pc	L	-10 ± 4	-43 ± 4	51 ± 5	21 ± 2	28
	R	10 ± 4	-41 ± 5	54 ± 6		
hMST+	L	-45 ± 4	-71 ± 5	8 ± 5	82 ± 9	35
	R	48 ± 5	-67 ± 3	8 ± 6		
hMT/hV5	L	-45 ± 3	-76 ± 2	8 ± 5	264 ± 16	35
	R	47 ± 2	-70 ± 3	7 ± 6		
VIP	L	-26 ± 5	-52 ± 5	53 ± 5	25 ± 2	32
	R	27 ± 6	-52 ± 8	53 ± 4		
Retinotopic Mapping (n = 10 participants):						
V1	L	-7 ± 1	-91 ± 2	-1 ± 4	1110 ± 26	20
	R	10 ± 1	-88 ± 2	2 ± 2		
V2	L	-8 ± 2	-88 ± 3	1 ± 4	911 ± 41	20
	R	12 ± 2	-86 ± 3	1 ± 4		
V3	L	-18 ± 3	-86 ± 3	-2 ± 3	700 ± 41	20
	R	21 ± 3	-83 ± 3	0 ± 3		
hV4	L	-27 ± 3	-78 ± 4	-13 ± 1	445 ± 31	20
	R	30 ± 3	-77 ± 5	-11 ± 2		
IPS0	L	-23 ± 4	-79 ± 2	32 ± 5	598 ± 61	20
	R	26 ± 4	-77 ± 2	36 ± 6		
IPS1	L	-21 ± 5	-70 ± 3	40 ± 7	418 ± 45	19
	R	23 ± 3	-68 ± 4	44 ± 7		
IPS2	L	-20 ± 5	-68 ± 4	49 ± 6	259 ± 27	19
	R	22 ± 5	-64 ± 3	54 ± 4		
IPS3	L	-22 ± 6	-65 ± 6	57 ± 5	274 ± 39	17
	R	23 ± 5	-57 ± 3	59 ± 5		
IPS4	L	-19 ± 7	-58 ± 6	59 ± 4	397 ± 44	17
	R	24 ± 7	-52 ± 6	60 ± 6		

definition was done in FreeSurfer (http://surfer.nmr.mgh.harvard.edu/). See Table 1 for ROI coordinates and volumes.

Spatial localizer control

Since participants' visual fields extended beyond the size of the stimulus display, we localized voxels in the early visual cortex that responded only to positions we could control the motion stimulation for. We presented checkerboard stimuli similar to the one used in the retinotopic mapping experiment (100% contrast, 6Hz contrast inversion flicker, logarithmically increasing check size), while participants were fixating a fixation cross at the screen center. We presented three condition: 1) Stimulation only at central 0.9° (size of the fixation disc in main paradigm), 2) full-field stimulation without central 4° and 3) stimulation only at central 4° (quarter of screen size) without area of fixation disc (Fig.2.3). Data acquisition, preprocessing and statistical analysis were identical to the main experiment. We then contrasted responses to condition 3 versus 2 and 1 to localize voxels responding only to stimulation at the central 4×4 degree of the visual field surrounding the fixation disc. For a control analysis, ROIs were then intersected with the resulting t-map, thresholded at $p<0.001$ (uncorrected) on single participant level. Please note that voxels with population receptive field sizes exceeding 4° would likely be lost in this contrast and would hence be excluded in this control analysis. Typically, higher level motion- and parietal regions have population receptive field sizes far bigger than this, which is why we limit this control analysis to early visual cortex whose population receptive fields are typically smaller than 4° visual angle [18].

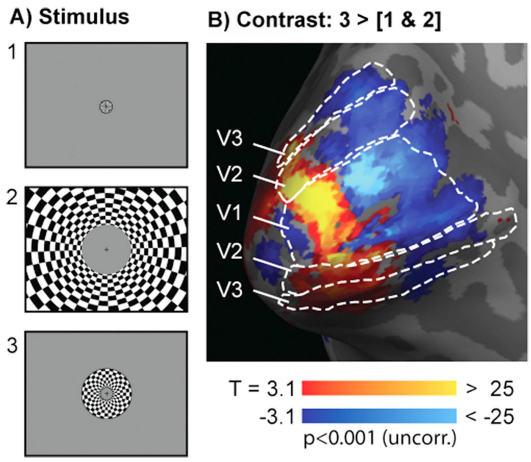


Figure 2.3: Spatial localizer. A) Stimulus: flickering checkerboard, three conditions: 1) localize area of fixation disc, 2) localize periphery, 3) localize center of the visual field surrounding fixation disc (central 4.1 visual degree, quarter of screen height). Figures were cropped for visualization and do not show true stimulus proportions. B) Localizer contrast: condition 3 contrasted against conditions 1 and 2. Corresponding T-map thresholded at T (667)=3.1, p=0.001 of an exemplary participant overlaid on his respective inflated brain surface for illustration purposes. This contrast was applied to early visual ROIs of each participant to localize voxels preferring stimulation within the central 4.1 visual degrees to stimulation at the fixation disc and in the periphery beyond 4.1°. Shown is the left hemispheric occipital lobe from a posteromedial view. Dotted lines show ROIs for early visual cortex as obtained by the retinotopic localizer.

Regions of interest analysis

For regions of interest analysis, beta estimates for each condition were extracted and averaged for each ROI and hemisphere of each individual participant. Beta estimates were averaged across the two pursuit speeds for each condition, as our interest for this study was on activity related to relative speeds between objective motion and pursuit. This resulted in three averages per ROI: 'Slower', 'Matched' and 'Faster' objective motion compared to pursuit. One-tailed paired t-tests were performed to test whether Faster conditions generally elicited a stronger response than Slower conditions. In this contrast, retinal- as well as pursuit motion were exactly matched while only objective motion velocity differed. It hence constitutes the cleanest and most powerful test to determine whether an area showed velocity-dependent objective motion responses in the absence of retinal- or pursuit motion biases. The Faster vs. Slower contrast constitutes the key comparison on which the main conclusion of this study are based. In order to examine these responses in more detail, we additionally used one-tailed paired t-tests to test

whether Faster conditions elicited a stronger, and Slower conditions a weaker response compared to the Matched condition.

Eye tracking

We used a video-based infrared eye tracker with long-range optics (Eye-Trac6; Applied Science Laboratories, Bedford, MA, USA) to track the gaze at 60Hz online during the experiment in 14 of the 18 participants (in 4 participants the procedure failed due to technical reasons). The stationary periods at the beginning and end of each trial as well as blinks and saccades were removed from analysis. Saccades and blinks were identified using a threshold of five times the median-based standard deviation of velocity (saccades) and pupil aspect ratio (blinks). Median x and y position was set to zero for both gaze and fixation target. Data were linearly detrended to remove slow signal drifts and smoothed with a running-average kernel of 200ms to reduce high-frequency noise. As a measure of fixation accuracy the root-mean-squared error (RMSE) between fixation disc and eye position was calculated for each condition averaged across trials. To examine differences in fixation accuracy across conditions, a repeated-measures ANOVA was then performed across all conditions involving pursuit.

2.4 Results

Functional responses of visual areas

We aimed to determine which regions differentiated between distinct velocities of objective motion while pursuit-velocity as well as retinal velocity were matched. The only way the brain could distinguish objective motion velocities was by comparing the relative direction of retinal motion with that of pursuit. Retinal motion with the same direction as pursuit would indicate faster objective motion, while retinal motion in the opposite direction would indicate slower objective motion.

A whole-brain group analysis of the contrast of Faster compared to Slower conditions revealed activity in the occipital lobe, with its peak centered on area V3A, and also involving early visual regions (Fig.2.4). Importantly, Faster and Slower conditions only differed in objective motion, not in pursuit or retinal motion velocity. Responses in this contrast hence reflect activity related to objective motion alone.

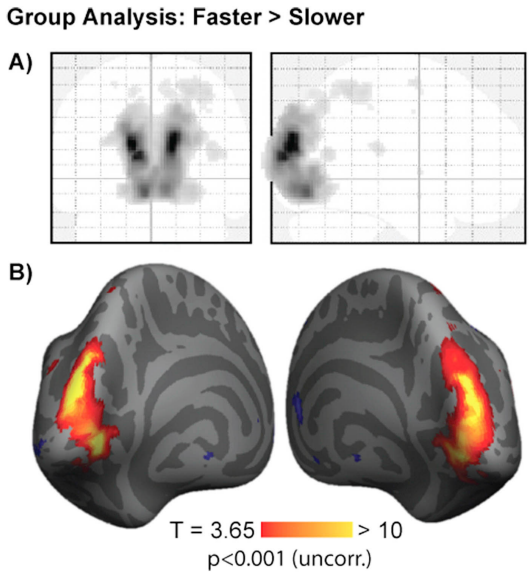


Figure 2.4: Whole-brain group analysis for Faster versus Slower conditions (n=18 participants). A) Group responses are shown for illustration purposes with liberal threshold (p<0.001, uncorrected) to ensure that no major response was overlooked in the ROI analysis. B) Same analysis and statistical thresholding as in (A) but shown overlaid on an inflated average surface brain viewed from a posteromedial view. Global peak activity was located at X=14, Y=-88, Z=24 and coincides with coordinates for area V3A (see Table 1).

First we focus on functional responses of the motion responsive regions hMT/V5, hMST+, V3A, V6, CSv, Pc and VIP. In line with previous evidence [61], areas V3A [$t(17) = 8.51$, $p = 7.8*10^{-8}$] and V6 [$t(17) = 5.21$, $p = 3.5*10^{-5}$] preferred objective motion faster than pursuit to objective motion slower than pursuit (Fig.2.5A). In addition, we found the same response profile in area VIP [$t(15) = 4.45$, $p = 2.3*10^{-4}$] and to a lesser extent in area Pc [$t(17) = 3.48$, $p = 0.0014$]. hMT/V5 and hMST+ were not significantly modulated while CSv reached significance in uncorrected tests. To investigate these findings in more detail, we then compared Faster and Slower to Matched conditions. Note that these contrasts include responses to both, objective motion and to retinal motion. Both areas V3A and V6 preferred Matched to Slower motion conditions [$t(17) = -7.76$, $p = 2.74*10^{-7}$], [$t(17)=-3.47$, $p=0.0015$] with V3A showing the strongest effects among all regions tested.

Among the motion responsive regions, areas V3A, V6, VIP and Pc hence signaled the direction of retinal motion relative to the pursuit direction thus encoded objective motion velocity during eye movements.

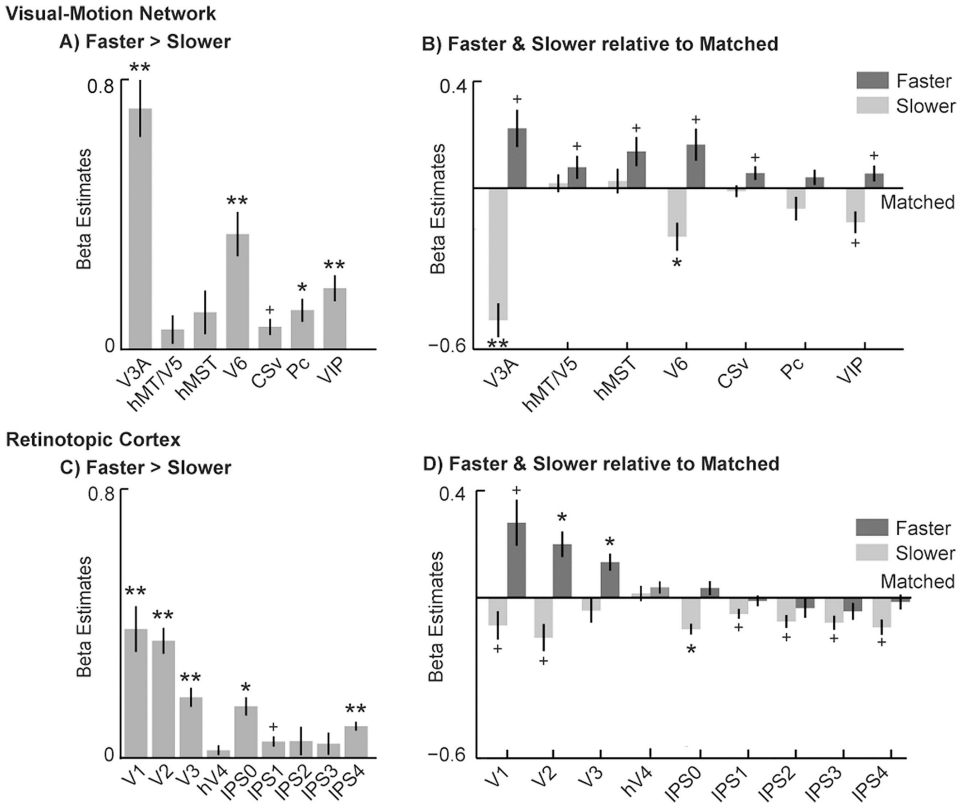


Figure 2.5: ROI analyses of motion regions (A & B) as well as of retinotopic early visual and parietal regions (C & D). A,B) ROI responses of the visual motion regions (n=18 participants, see Table 1) to (A) difference between Faster and Slower and (B) responses to Faster and Slower relative to Matched. C,D) ROI responses of retinotopically defined areas V1-hV4 and IPS0-IPS4 (n=10 participants, see Table 1) to (C) Faster versus Slower and (D) responses to Faster and Slower compared to Matched. **p < 0.01 (Bonferroni corrected), *p < 0.05 (Bonferroni corrected), +p < 0.05 (uncorrected). Error bars depict the SEM. ROIs for visual motion network and retinotopic cortex were defined based on two independent localizer scans. Bonferroni correction for the visual motion network (A,B) was performed for n = 21 tests (upper panel), for retinotopic cortex (C,D) n = 27 tests (lower panel).

We next examined responses in retinotopically mapped areas V1-hV4 and IPS0-IPS4. Early visual areas V1 [$t(9) = 5.64$, $p = 1.6*10^{-4}$], V2 [$t(9) = 9.04$, $p = 4.1*10^{-6}$] and V3 [$t(9) = 6.45$, $p = 5.9*10^{-5}$], but not hV4, preferred objective motion faster than pursuit to objective motion slower than pursuit (Fig.2.5C). We also found significant differences between Faster and Slower conditions in IPS0 [$t(9) = 4.24$, $p = 0.0011$] and IPS4 [$t(8) = 5.29$, $p = 3.7*10^{-4}$]. While areas V2 [$t(9) = 4.21$, $p = 0.0011$] and V3 [$t(9) = 4.20$, $p = 0.0012$] additionally preferred Faster to Matched, IPS0 preferred Matched to Slower [$t(9) = -4.39$, $p = 8.75*10^{-4}$] (Fig.2.5D).

Our results indicate that in addition to above mentioned areas of the functional motion network, also early visual areas V1, V2 and V3 differentiated between distinct objective motion velocities during pursuit.

Spatial localizer control

To ensure that responses in early visual cortex were unaffected by motion energy originating outside of the stimulation display, we limited the analysis to voxels of the early visual cortex that responded only to the 4° surrounding the fovea, but not to the fixation disc or the periphery. Areas V1 [$t(9) = 4.06$, $p = 0.0014$], and V2 [$t(9) = 4.92$, $p = 4.1*10^{-4}$] still clearly preferred faster to slower objective motion during pursuit (Fig.2.6).

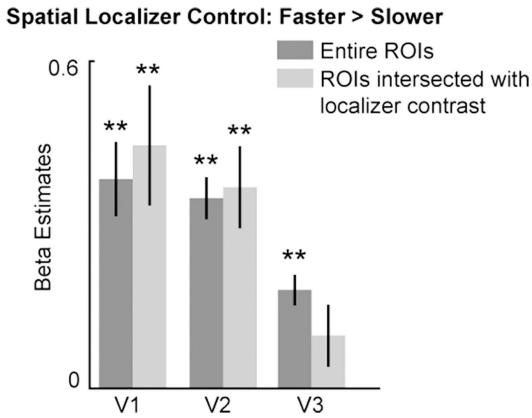


Figure 2.6: ROI analysis of early visual regions for entire ROIs and limited to voxels responding to stimulation surrounding the fovea but not to periphery or the fixation disc. For the latter, ROIs were intersected with t-maps for the independent spatial localizer contrast threshold at $p < 0.001$. **$p < 0.01$ (Bonferroni corrected for $n = 6$ tests). Amount of voxels in entire ROIs: V1: 1110 ± 26, V2: 911 ± 41, V3: 700 ± 41 and ROIs intersected with localizer contrast: V1: 525 ± 42, V2: 400 ± 29, V3: 372 ± 30. Error bars depict the SEM.

Eye tracking and behavioral performance

We performed repeated-measures ANOVAs across all pursuit conditions for fixation error and for behavioral measures related the fixation task. Behavioral measures did not reveal any influence of condition on hit rate [$F(5, 85)=1.298$, $p=0.272$], false alarms rate [$F(5, 85)=1.312$, $p=0.267$] or response time [$F(5, 85)=0.83$, $p=0.532$]. Fixation accuracy monitored with eye tracking did not reveal differences between conditions ($F(5, 65)=1.75$, $p=0.135$). Root mean square errors for fixation accuracy were as follows: RMSE for 2 deg/s pursuit: Slower: 1.84 ± 0.16, Matched: 1.78 ± 0.15, Faster: 1.83 ± 0.12, and for 3 deg/s pursuit: Slower: 1.85 ± 0.17, Matched: 1.77

± 0.15, Faster: 2.0 ± 0.22], in line with previous reports from our group using a similar task [61].

2.5 Discussion

We examined the ability of sixteen visual areas to differentiate between distinct velocities of objective motion. Our paradigm kept retinal motion as well as eye movement velocity constant while only varying objective motion velocity. Visual regions could hence distinguish objective velocities only by integrating the direction of retinal motion with that of eye movements. All other factors – visual and non-visual – were matched. In line with previous evidence, we found the strongest differential responses in higher-level motion areas V3A and V6 [61] and additionally also in VIP and to a weaker extend in IPS0, IPS4 and area Pc. Importantly, our results provide clear evidence that in addition also early visual areas V1-V2 and possibly V3 signal objective motion velocity in the absence of retinal and pursuit motion biases. It is unclear whether this is related to bottom-up or feedback related processing. Neither fixation error nor any other measure related to the fixation task indicated a condition-dependent influence on our results.

Visual motion regions

Among all motion responsive regions, V3A had the strongest response preference to faster compared to slower objective motion, similar to results obtained before with fewer participants and fewer ROIs (see expt. 4 in [61]). These results were consistent across two pursuit eye movement velocities and are compatible with a high fraction of real-motion [62] and gaze-dependent neurons in area V3A in monkey [269].

Similarly, the present findings confirm prior results for human V6 [61] and are consistent with high percentages of real-motion cells in area V6 in monkey [252]. Human V6 has been shown to respond to coherent optic flow, to have receptive fields extending well into the periphery [24, 253, 270, 271], and to prefer near-field over far-field stimulation [272]. Together with its connections with grasping related areas, V6 is well suited to provide other areas with information about movements of objects in reachable space in context of vision-for-action [64, 252, 266, 267].

Human VIP was localized here according to a prior study relying on multi-modal cues [273], a known property of monkey VIP [92, 274]. Its strong response to objective motion demonstrated here is compatible with monkey reports showing VIP neurons to represent visual space and heading information invariantly in

head-centered coordinates [50, 275]. Similarly, our findings are compatible with a great portion of VIP neurons being speed and direction selective [276], and indeed changing their preferred direction depending on whether visual motion was self-induced or not [277].

Moreover, monkey VIP was shown to causally influence heading perception during eye movements [278]. Even though monkey VIP receives inputs from V5/MT and especially from MST [275,279], its functional responses to planar objective motion clearly differed from the ones of hMT/V5 and hMST+ in the present experiment. Areas hMT/V5 and hMST+ did not signal retinal motion direction relative to pursuit. Our prior study saw only weak objective motion response in hMST+ [61], even though it contains robust responses at the neural level [48]. Interestingly, for fixed eye-positions, there is good evidence for spatiotopic coding in human V5/MT [68], even though it is strongly attention-dependent [38, 70], suggesting that motion-pursuit-integration differs in terms of the neural machinery involved. The stronger objective motion response in VIP could be explained by pooling, as it is thought to do for flow [273,275,280], but also through its input from V6 [281]. This contrasts VIP also to all other intraparietal regions examined here. In fact, even though IPS0 and IPS4 reached significance, retinal motion generally seemed to attenuate responses in the IPS during pursuit. Further, we found objective motion responses in area Pc as well as to a lesser extent (not surviving Bonferroni correction) yet consistent with previous reports in the cingulate sulcus visual region CSv [65].

Early visual cortex

The current results demonstrate for the first time objective motion responses in human early visual cortex and reveal early visual speed tuning for objective motion: faster objective motion led to higher responses than slower objective motion while retinal motion as well as eye velocity were matched. Prior research has established both velocity and direction tuned responses in early visual cortex [7, 282, 283], with similar findings in human neuroimaging [284–287]. Electrophysiological reports have demonstrated the presence of extra-retinal signals in monkey early visual cortex in the form of real-motion neurons [47,250] as well as neurons that compensate consequences of ocular counter-rolling in V1 [60]. Further, V1 neurons in primates have been shown to differentiate between real and simulated microsaccades [288].

The present result that responses to Faster and Matched differed significantly while responses to Matched and Slower did not can be easily explained by responses to both, retinal and objective motion: retinal motion positively con-

tributed not only to Faster but also to Slower conditions, hence counteracting the positive contribution of objective velocity that increased from Slower, to Matched, to Faster.

Areas V1 and V2 strongly preferred faster to slower objective motion even when we limited the analysis to the central 4×4 degree of the visual field excluding the fixation disc. Since we controlled visual stimulation in at least the central 8×8 visual degrees, effects of pursuit-induced retinal motion originating outside of the stimulus display cannot explain these early visual responses. Statistical tests for area V3 however did not reach significance in this control analysis, indicating weaker objective motion sensitivity for the central visual field in human V3 compared to the periphery. This is consistent with real-motion neurons being found in V1-V2 [47, 250] and V3A [62] but to our knowledge not in V3. Finally, a recent monkey fMRI study showed a preference to self-induced versus stimulus-induced motion in central early visual cortex whereas the opposite was true in higher-level regions [289]. While related to our question, the results cannot directly be compared as the monkey fMRI results were driven by differences in eye movements (whereas they were balanced here), and secondly we used pursuit whereas the monkey results were related to saccadic processing [289]. It would hence be interesting to re-examine the present paradigm in context of saccadic eye movements. While area V4 responds to visual motion in a direction-specific manner in monkey [290], we did not observe any significant modulation of human V4 in any of the contrasts tested.

Feedback to early visual cortex?

There is evidence that conscious vision is possible without a functioning V1 [291, 292] and that parieto-occipital regions beyond V1 are necessary for perceptually separating self-induced from objective motion [293]. In fact, early visual motion responsive neurons are relatively unspecialized, often confusing visual motion with contrast flicker [294, 295].

In order to respond to objective motion, neurons need to integrate eye movement commands with visual input, a process that is considered to be implemented at a higher stage of the visual processing hierarchy. It therefore seems possible that early visual responses to objective motion observed in this study reflect top-down feedback rather than bottom-up processes. An abundant amount of studies demonstrate such feedback from areas like hMT+/hV5+, V6 and V3A to V1 [281, 284, 296–300], some even indicating its contribution to visual awareness of motion [300, 301]. Besides, fMRI is very susceptible to local processing and feedback,

since BOLD responses mainly reflect synaptic activity rather than actual output spiking [302, 303]. In the context of predictive coding theory [299], feedback has been demonstrated to account for signal in early visual cortex in processing of various features such as color [304], size [305] and shape [306]. In context of motion, feedback was reasoned to explain early visual preferences to incoherent to coherent motion as these two motion types differ in predictability [307, 308]. However, since retinal motion and therefore predictability was matched in faster and slower conditions, the present results cannot be sufficiently explained in the same manner. Indeed, there is clear evidence showing that a particular class of neurons in V1 code visual input in head-centered coordinates to compensate for a particular type of eye movements, so-called ocular counter-roll [60]. Even though the mechanisms are likely unrelated to those of pursuit, the evidence shows, along with reports of real-motion cells [47], that V1 receives extra-retinal signals related to eye movements.

2.6 Conclusion

We found that early visual cortex signals the direction of retinal motion relative to the eye movement direction to a similar extent as high-level motion areas V6 and VIP, yet less so than V3A. In addition, our results show objective motion responses in areas Pc, a cingulate area which has so far only rarely been included in visual motion studies as well as in intraparietal areas IPS0 and IPS4. Apart from providing the first evidence of objective motion responses in human early visual cortex, the current study gives an overview on objective motion encoding across the visual system. Our results point to a role of early visual cortex in compensating for self-induced visual motion and thus to a potential role in the stable perception of our visual environment.

2.7 Supplementary material

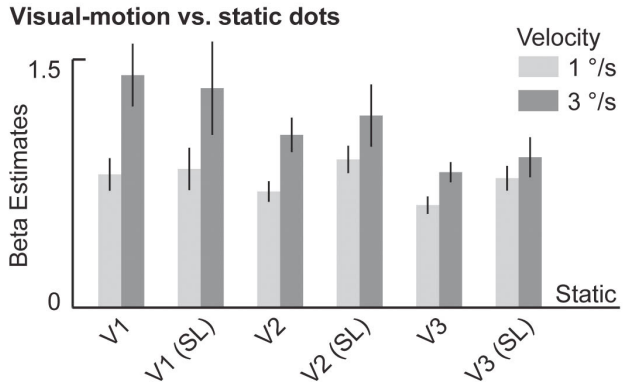
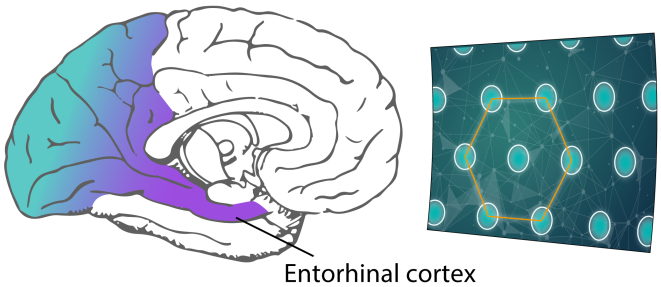


Figure 2.7: Visual-motion vs. static dots during central fixation. We plot responses to 1° /s and 3° /s visual-motion relative to static dots for unmasked early visual cortex ROIs (V1 -V3) andmasked with a spatial localizer (SL) for central visual field only (V1 (SL) -V3 (SL)). In this contrast, objective motion and retinal motion are equivalent. Barsdepict the mean, error bars depict the SEMacross n = 10 participants.

Chapter 3

Hexadirectional coding of visual space in human entorhinal cortex



This chapter has been published as:
Hexadirectional coding of visual space in human entorhinal cortex.
Nau, M., Navarro Schröder, T., Bellmund, J.L.S., Doeller, C.F. (2018).
Nature Neuroscience 21, 188-190.
https://doi.org/10.1038/s41593-017-0050-8

The data used in this work were collected at the Donders Institute for Brain, Cognition & Behavior at the Radboud University in Nijmegen, the Netherlands. All analyses, the writing of the manuscript and the publication were part of the doctoral work at NTNU only.

3.1 Abstract

Entorhinal grid cells map the local environment, but their involvement beyond spatial navigation remains elusive. We examined human functional MRI responses during a highly controlled visual tracking task and show that entorhinal cortex exhibited a sixfold rotationally symmetric signal encoding gaze direction. Our results provide evidence for a grid-like entorhinal code for visual space and suggest a more general role of the entorhinal grid system in coding information along continuous dimensions.

3.2 Introduction

Grid cells in entorhinal cortex (EC) exhibit place-responsive firing patterns that represent self-location [151, 153, 173, 309]. These firing patterns are sixfold rotationally symmetric (hexadirectional) and are implicated in navigation and the formation of maps of the local environment. Investigations in nonhuman primates show that EC neurons further encode eye-movement direction [229] and visual space [225]. Here grid cells fire as a function of gaze location rather than self-location, suggesting visuospatial coding in EC that is strikingly different from retinotopic representations typically found in visual cortex [13]. In humans, functional MRI (fMRI) has demonstrated that EC activity depends on virtual running direction in a hexadirectional manner, in line with a putative grid cell population response [158]. However, it is currently unknown whether human EC codes for visual space and eye movements beyond navigation.

We used fMRI to examine human EC responses during a visual tracking and object location memory task. Our experiment (Fig.3.1 and Supplementary Figs.3.3 and 3.4) ensured attentional focus on visuospatial information while balancing directional sampling of smooth-pursuit and saccadic eye movements. We exposed participants to a virtual arena from bird's eye view and presented a fixation target that moved within the environment in a highly controlled and directionally balanced fashion. We sampled eye movement directions with 10° angular resolution and monitored fixation accuracy with a magnetic resonance-compatible eye-tracking system. To facilitate spatial attention, participants memorized object locations within the environment (arena shown in two orientations: 0° and 180°) that were cued only when the fixation target crossed over them.

Imaging data were analyzed using a bootstrapping-based, leave-one-out threefold cross-validated symmetry test (see Methods). We fitted a general linear model to obtain beta estimates for each voxel and eye-movement direction for three data partitions (run blocks within which directional sampling was balanced) for every participant. For every EC voxel (Supplementary Fig. 3.4), we then estimated the phase of the putative six-peaked oscillatory modulation (putative grid orientation) of fMRI activity across directions, based on a normalized average of two of the data partitions. Using the third, independent, data partition, we then tested whether directions aligned to the putative grid orientation (0 modulo 60°) exhibited stronger EC activity than directions misaligned (30 modulo 60°) to it. This process was iterated until every data partition served as the test set once and as the training set twice. Using a bootstrapped distribution obtained by

shuffling the direction labels of the test set 1,000 times, resulting contrast coefficients were transformed to z-scores, averaged across iterations and EC voxels and taken to the group level (Supplementary Fig. 3.6). Since the distribution of z-scores was non-normal (one-sample Kolmogorov-Smirnov goodness-of-fit test for EC: K=0.367, P=5.2x10^{-4}), z-score we performed nonparametric one-sided Wilcoxon signed-rank tests to examine whether z-scores for aligned directions were systematically higher than for misaligned directions.

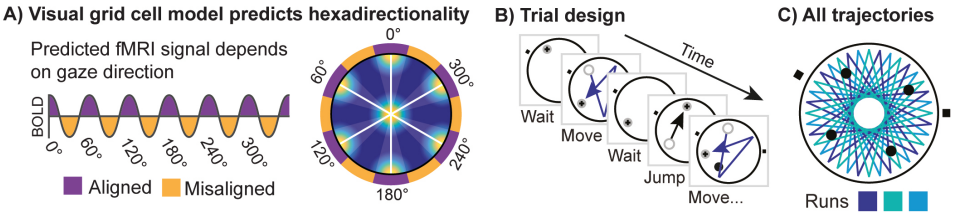


Figure 3.1: a, Hypothesis. The number of firing fields crossed depends on eye movement direction (more fields are crossed for directions aligned to grid axes (white lines) compared to directions misaligned to it). Because grid orientation3 and spatial phase4 cluster across cells and because conjunctive direction-tuning has been reported to align to the grid axes8, this relationship translates to hexadirectional biases of putative grid cell population activity, in turn predicting a stronger fMRI signal for aligned versus misaligned directions. BOLD, blood oxygen level-dependent contrast. b, Visual tracking task. Experiment consisted of 3 blocks of 3 runs incorporating 9 trials each. Each run tested a different set of 12 directions. Typical trial: each trial started with the fixation target (black cross) being stationary (1s), then moving (6-10s; blue arrows) to sample 3-5 eye-movement directions sequentially in 2-s blocks, followed by a waiting period (1s) and finally jumping (black arrow) to a random position and continuing moving from there. Objects were shown when the fixation target moved over predefined locations (black circle). c, All trajectories tested in this experiment. Each run block tested all trajectories (36 directions) equally with 10° angular resolution. External landmarks: black squares. Objects were cued at four predefined locations, depicted as black circles.

3.3 Results

Eye-movement directions aligned to the putative grid orientation consistently induced stronger EC activity than misaligned directions (z=2.38, P=8.7x10^{-3}; Fig.3.2a). To investigate these results in more detail, we aligned each voxel's beta estimates, cross-validated before bootstrapping (also see Supplementary Fig. 3.7), to the respective estimated putative grid orientation and found that, across participants, the six aligned directions generally elicited stronger EC responses than the six misaligned directions (Fig.3.2c). The hexadirectional effect in EC was therefore not driven by a single direction but rested upon a systematic sixfold rotational

symmetry in response to eye movement direction. We repeated the main analysis for several control regions as well as biologically implausible control symmetries (Fig.3.2d). Among all symmetries tested, and after Bonferroni correction, only sixfold symmetry yielded a significant modulation (control symmetries: fourfold: z=1.04, P=0.150; fivefold: z=0.43, P=0.333; sevenfold: z=1.12, P=0.130; eightfold: z=0.69, P=0.245). Among all control regions, and again after Bonferroni correction, only EC showed a significant hexadirectional modulation (control regions: early visual cortex: z=0.84, P=0.200; motor cortex: z=-0.35, P=0.635; frontal lobe: z=1.08, P=0.140; parietal cortex: z=1.19, P=0.117). A subpart of the frontal lobe was weakly hexadirectionally modulated in line with previous studies [] but did not survive Bonferroni correction (Supplementary Fig. 3.8). To examine whether EC generally responded to eye movements in our task irrespective of direction, we then compared responses to eye movements to a visual motion control in which the fixation target remained at the screen center while the arena moved instead, matching retinal motion in the main experiment (see Methods). Among our regions of interest (Fig.3.2b), only early visual cortex ($z=4.55$, $P=5.3\times10^{-6}$) and parietal cortex ($z=-2.82$, $P=4.8\times10^{-3}$), but not EC (z=1.50, P=0.133), exhibited significant responses in this contrast (Fig.3.2d). EC hence did not respond to eye movements or to visual motion per se, but responded to gaze in a highly direction-specific fashion. We did not find evidence for hexadirectional modulation in the visual motion control condition (z=-0.39, P=0.651). Next, we estimated one putative grid orientation per voxel with maximized signal-to-noise ratio (direction-specific beta estimates averaged across partitions) and quantified how much these orientations varied across voxels, as well as around the participant's across-voxel mean orientation (Fig.3.2e). We found that putative grid orientations were not random across voxels but clustered around the participant's mean orientation (V-test for nonuniformity, $V=31.80$, $P=3.46\times10^{-6}$), with higher coherence within compared to across hemispheres ($t(28)=-3.67$, P=0.001, confidence interval [-0.03, -0.01]; Supplementary Fig. 3.9). More specifically, voxels exhibiting a stronger hexadirectional signal varied less around the across-voxel mean than those showing a weaker effect ($t(28)=-6.67 P=1.5\times10^{-7}$, confidence interval $[-\infty, -0.19]$, also true for unsmoothed data; Supplementary Fig. 3.9). On the one hand, this indicates that not all EC voxels were engaged in our task (there is noise), but on the other hand it also indicates that others indeed represent activity with a predominant orientation.

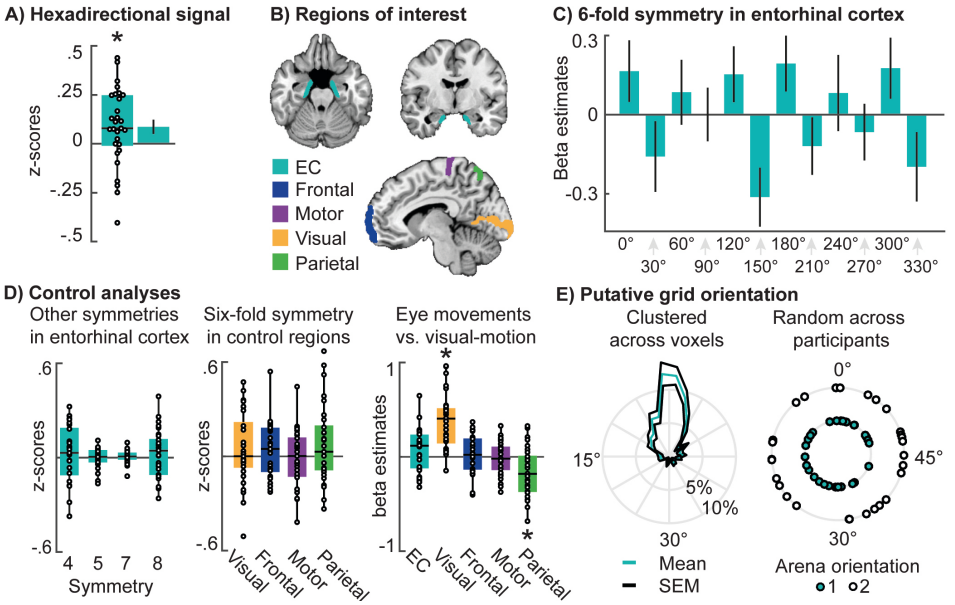


Figure 3.2: a, EC exhibited higher fMRI activity for aligned vs. misaligned directions (Bonferroni-corrected one-sided Wilcoxon signed-rank test, z=2.38, *P=8.7×10^{-3}, n=29). Left: aligned vs. misaligned contrast for single-participant data overlaid on whisker-and-box plots (center, median; box, 25th to 75th percentiles; whiskers, 1.5×interquartile range) and mean and s.e.m. across participants. Right: same data depicted as a bar plot, showing mean and s.e.m. across participants. b, Regions of interest (ROIs) displayed on the T1 template (Montreal Neurological Institute (MNI) coordinates: X=-8, Y=-11, Z=-20). c, To visualize the effect in a in more detail, we plot cross-validated beta estimates for aligned and misaligned directions, sorted according to putative grid orientation (0°). Each bar represents an averaged 20° bin of eye-movement directions averaged across the three iterations and across participants (see Methods). Error bars depict s.e.m. across 29 participants. d, Control analyses. Left: aligned vs. misaligned contrast for control symmetries (four-, five-, seven- and eightfold periodicity). Middle: sixfold symmetry effect for control regions. Right: eye movements vs. visual motion contrast. Activity in early visual cortex (z=4.55, *P=5.3×10^{-6}) and parietal cortex (z=-2.82, *P=4.8×10^{-3}) was significantly different in these conditions (Bonferroni-corrected two-sided Wilcoxon signed-rank tests, n=29). As in a, we plot single-participant data overlaid on whisker-and-box plots. e, Left: circular distances between putative grid orientation of each voxel and the circular mean across voxels; radial axes represent percent of voxels. Right: putative grid orientations across participants for both arena orientations.

Consistent with previous reports of first-person navigation in nonpolarized environments8, putative grid orientations were not clustered across participants (Rayleigh's tests for nonuniformity, arena orientation 1: z=0.083, P=0.922; orientation 2: z=0.141, P=0.871; see Methods). Participants were able to report all object locations equally well both from the bird's eye view (F(3,81)=0.64, P=0.593) and in

subsequent first-person navigation ($F(3,81)=1.9$, $P=0.136$; Supplementary Fig. 3.4). Eye tracking analysis did not reveal any confounding symmetry in fixation accuracy relevant to our imaging results ($F(4,92)=1.7$, $P=0.156$), and average eye velocity matched the one of the fixation targets equally well for all directions ($F(35, 805)=1.1$, $P=0.313$; see Methods and Supplementary Fig. 3.10).

3.4 Discussion

Our results provide (to our knowledge) the first evidence of gaze-dependent hexadirectional signals in human EC and are consistent with reports of eye-movement direction encoding [229] and visuospatial grid cells [225] in (non-human) primates. In contrast to the aforementioned studies, our experiment involved both smooth-pursuit and (low-amplitude) saccadic eye movements; however, it allowed us to examine eye movements with predominant directions (critical for our analysis), was highly controlled and had balanced sampling across directions, and reduced perceptual distortion effects known to occur in (high-amplitude) saccades [310]. Gaze location has been shown to drive spatial view cells in monkey hippocampus [221], and primates generally rely strongly on vision as they explore space [213]. While such hippocampal viewing responses have not yet been reported in rodents, the hippocampal formation has been shown to integrate visual information [311]. In contrast to retinotopic [13] and spatiotopic [12, 312] representations typically found in visual cortex, however, our results implicate human EC in encoding visual space with a grid-like code, a representation drastically different from any found in visual cortex, as well as from visually driven responses [104] and memory-guided visual processing [313] in other mediotemporal lobe regions. One alternative explanation for our results could be that participants imagine themselves navigating as they track the fixation disc. While hexadirectional signals have been observed during mental imagery [162], our results are unlikely to be explained in the same manner because, in contrast to the aforementioned study, our experiment engaged participants in an active attention task and because EC was not hexadirectionally modulated in the visual motion control. While saccade-direction cells [229] exhibiting hexadirectional population activity could provide a potential nongrid explanation of our results, such asymmetries in saccade-direction tuning have not yet been observed. However, grid cell population responses can provide a parsimonious yet substantiated explanation for the signal we observed (Fig.3.1a). It remains to be shown how eye-movement-invariant visuospatial information reaches EC, as many areas contribute to visual constancy

during eye movements [312]. However, given its strong involvement in visuospatial processing in both monkeys and humans [12, 312], and despite not being hexadirectionally modulated in our task, parietal cortex appears to be a prime candidate for future studies in this context.

Conclusions

Recent reports implicate EC in coding not spatial information per se, but generally continuous feature dimensions also including sound frequency [237] or abstract knowledge [168]. The present study supports this notion and further suggests a domain-general coding regime in EC that is also employed to encode visual space and potentially contributes to memory-guided viewing behavior orchestrated by the hippocampal formation [314]. Our results are in line with reports of hexadirectional signals during virtual navigation [158] and mental simulation [162] and show that eye movement and visual information is processed in human EC. They further suggest a grid-like code for representing visual space and point toward a broader role of EC in cognition beyond spatial navigation [168, 237].

3.5 Methods

Participants

A total of 36 participants (22 female, age range: 18-35, normal or corrected to normal vision) were recruited for this study via the Radboud University Nijmegen online recruitment system. Participants were not grouped and hence no sample randomization or blinding was performed. Two participants were excluded due to strong head movements, frequently larger than the voxel size of 2mm. Another 5 participants were excluded because eye tracking revealed low tracking-task performance (deviating more than one s.d. from the average fixation error across participants). In total, 29 participants entered the analysis (in line with a priori G*Power sample size estimation (https://stats.idre.ucla.edu/other/gpower) at 80% power, alpha=0.05 and assumed medium effect size of d=0.5, n=27 participants; see the Life Sciences Reporting Summary for more information). The study was approved by the local ethics committee (CMO Arnhem-Nijmegen, The Netherlands), all relevant ethical regulations were followed, and participants gave written consent before the experiment.

Stimulus and experimental procedure

Prior to scanning, participants navigated a circular virtual arena (generated with Unreal Engine 2, https://www.unrealengine.com) in first-person perspective for 2-3min (Supplementary Fig. 3.4c). During scanning we then exposed them to the same virtual arena from bird's eye view (screen size: 24°×18°, arena diameter: 17° visual angle, Supplementary Fig. 3.4a) and presented a fixation target that moved within the environment in a highly controlled yet (to the participant) unknown fashion (fixation task). In addition, participants memorized object locations within the environment (object location memory task), for which the arena was shown in two orientations (0° and 180°) to enhance spatial attention. Stimuli were presented using Psychophysics Toolbox version 3 (http://psychtoolbox.org/). Arena orientation alternated across trials of each condition, which were presented in counterbalanced order. The arena was surrounded by two landmarks (two white squares matched in brightness, one hollow and one filled, at 0° and 150° relative to horizontal axis in orientation 1). Landmark identity alternated between the two locations across participants. The experiment consisted of a total of 9 runs with 9 trials each. To ensure a balanced sampling of directions, the first trial of each run was discarded. For every run, 4 trials of the main experiment (∼1min each) and 4 trials of visual motion control (∼30s each) were included in the analysis. Condition sequence was counterbalanced. At the end of the experiment, participants

navigated the virtual environment in first-person view again (Supplementary Fig. 3.4c) and reported object locations by navigating to them and confirming the location via button press (25 trials, 6 trials per object, counterbalanced order, first trial was discarded).

Fixation task

Participants fixated at a fixation target (black cross on light gray circle, diameter: 0.8° visual angle) moving with a constant speed of 7.5° /s inside the arena, which was shown from a bird's eye view (Fig.3.1 and Supplementary Figs. 3.3 and 3.4). The total trajectory of the fixation target resembled a twelve-pointed star (diameter: 15° visual angle from one point to another) and incorporated 12 directions in steps of 30°. In every trial, this trajectory, and hence each of the 12 directions, was tested twice and in blocks of 2s each. To disrupt periodicity of the stimulus, we split the trajectory of each trial into 6 paths, whose presentation order was randomized. The path lengths varied from 3 to 5 path segments (movement of the fixation target from one point to the other and hence a 2-s long block of a given direction). After every 3-5 path segments, the fixation target stopped for 1s, jumped to another randomly chosen point, paused another 1.5s to facilitate successful fixation and then continued its movement from there (Fig.3.1b). Across runs, the trajectory was rotated either -10° or+10° in pseudorandomized fashion to get a total directional sampling resolution of 10° (Fig.3.1 and Supplementary Fig. 1). After every three runs (one run block), all trajectory rotations were sampled equally and in a pseudorandomized order (Fig.3.1c and Supplementary Fig. 3.3). These balanced run blocks served as data partitions for later analysis.

Object location memory task

To focus spatial attention on the virtual environment during the fixation task, participants additionally memorized locations of objects within the arena. Each object was a colored circle (green, red, blue or purple) that was visible only when the fixation target moved across it. After three pseudorandomly chosen trials per run, participants reported the locations of objects by moving a cursor to the memorized location via key presses (one object per cueing, uninformed two-back task, randomly chosen arena orientation), after which the true location was revealed (feedback). Additional feedback about task performance was given in written form at the end of each run to enhance motivation. Two of the objects were positioned at polar coordinates 120° and 300° (0 modulo 60° relative to screen horizontal) and two other objects were at 30° and 210° (30 modulo 60° relative to screen horizontal; Fig.3.1c and Supplementary Figs. 3.3 and 3.4). Object positions were chosen to ensure that the overlap with the trajectory was balanced

across trajectory rotations. Each two-piece set of objects was hence presented equally often and long (0.27s/trial) regardless of trajectory rotation and did not induce hexadirectional biases. In both viewing and navigation tasks, participants were able to report the locations of the objects (Supplementary Fig. 3.4d), and there were no differences in spatial memory performance (Euclidean distance between true and memorized location) between objects in either of the tasks (repeated-measures ANOVA results: viewing task: $F(3,81)=0.64$, $P=0.593$; navigation task: $F(3,81)=1.9$, $P=0.136$). To examine whether differences in spatial memory between objects were masked by differences in putative grid orientation between participants, we grouped objects into aligned and misaligned objects according to their position on each participant's putative grid pattern for each of the arena orientations. We then tested whether spatial memory performance differed between aligned and misaligned objects (averaged across arena orientations) but did not find evidence for it (two-sided t test results: $t(27)=0.097$, $P=0.924$). In addition, the average hexadirectional signal did not correlate with the participants' average spatial memory performance during subsequent first-person navigation ($r=0.122$, $P=0.537$). One participant did not finish the final navigation task and was excluded here.

Visual motion control

To test whether entorhinal cortex generally responded to eye movements in our task irrespective of direction, we additionally scanned a control condition in which the fixation target remained in the screen center while the arena moved instead (visual motion control). If entorhinal cortex responded to eye movements, we expected to find an increased activation in the main experiment relative to this control condition since the two conditions only differed in eye movements, not in directional or positional sampling or in otherwise confounding retinal motion. Pixels leaving the screen on one side entered the screen on the other side, keeping visual motion constant over time (Supplementary Fig. 3.4b) and allowing us to apply the same motion stimulus to the retina in both presence and absence of eye movements. To reserve the major part of scanning time for the main experiment, this control condition was tested only for ∼30s per trial (compared to ∼60s for the main experiment; the same directions were tested, but only once instead of twice). Data was analyzed using a separate first-level general linear model, modeling all onsets of eye or arena movements irrespective of direction or arena orientation. The numbers of onsets of eye movements was downsampled to match the numbers of onsets of the visual motion control. As done for the main analysis, each run was modeled separately and the same nuisance regressors

were included. Beta estimates for eye-movement and visual-motion regressors were contrasted for each participant and tested in an ROI-based analysis using two-sided Wilcoxon signed-rank tests (EC: z=1.5, P=0.133; early visual cortex: z=4.55, $P=5.3\times10^{-6}$; motor cortex: z=-0.08, P=0.940; frontal lobe: z=0.96, P=0.336; parietal cortex: z=-2.82, $P=4.8\times10^{-3}$). If the EC was hexadirectionally modulated in the visual motion control, we expected to find at least a statistical trend despite shorter scanning time. However, we did not find any evidence that the EC was hexadirectionally modulated in our visual motion control (one-sided Wilcoxon signed-rank test results: z=-0.39, P=0.651) and both the mean (mean: -8.5×10^{-4}, sem: 0.023) and the median (median: -5.2×10^{-3}, 25th percentile: -0.056, 75th percentile: 0.107) of the aligned versus misaligned contrast were slightly negative.

Data acquisition and preprocessing

Functional T2*-weighted gradient-echo echoplanar images were acquired on a Magnetom PrismaFit 3 Tesla magnetic resonance tomograph with the following parameters: repetition time (TR)=1,000ms, echo time (TE)=34ms, multiband acceleration factor=6, voxel size=2×2×2mm, flip angle=60°, field of view=210×210mm, 66 slices, base resolution 104×104. In addition, an anatomical scan was recorded using a T1-weighed MPRAGE sequence with voxel size 1×1×1mm. Functional images were preprocessed using SPM12. Images were corrected for head movements, co-registered to the structural T1 image, spatially normalized to MNI space and smoothed with a 4mm full-width-at-half-maximum Gaussian kernel.

Regions of interest (ROI) definition

Because the mediotemporal lobe is difficult to image due to magnetic field distortions and susceptibility artifacts, we delineated the EC bilaterally for each individual participant based on the normalized mean-EPI-image using itk-SNAP (www.itksnap.org) and carefully double-checked voxel selection with the structural T1 image. This ensured that only voxels belonging to the EC on the functional images were included in the analysis (Supplementary Fig. 3.5). On average, EC was centered at -18/-12/-25 and included 105 voxels in the left hemisphere and centered at+19/-12/-24 with 106 voxels in the right hemisphere. Hemispheres were combined to one bilateral entorhinal ROI. As control ROIs we chose early visual cortex and motor cortex, for which, to our knowledge, no hexadirectional coding has been reported; the frontal lobe, parts of which exhibit hexadirectional signals [158]; and the parietal cortex, due to its strong involvement in visuospatial processing14,15. ROIs were created by co-registering corresponding probability maps of the SPM anatomy toolbox to the structural scan of each participant and thresholding them at 50% probability. The resulting binary masks were of follow-

ing size: visual cortex: 2,731 voxels, motor cortex: 797 voxels, frontal lobe: 2,780 voxels, parietal lobe: 1,165 voxels. Out-of-brain voxels were neglected. To ensure that potential hexadirectional signals were not washed out due to bigger ROI sizes for control regions compared to the EC, we repeated our main analysis with higher probability thresholds for the control regions (Supplementary Fig. 3.8). Here thresholds were chosen so that the resulting ROI masks approximated the average size of the EC (211 voxels). Note that the higher-threshold masks comprised voxels with highest probability for belonging to the respective areas; they did not, however, capture the actual extent of the control areas we were interested in. Thresholds and average numbers of voxels were as follows: visual cortex: 76%, 219 voxels; motor cortex: 69%, 202 voxels; frontal lobe: 99%, 283 voxels; parietal cortex: 84%, 205 voxels. All ROI masks were resliced to match our functional image dimensions.

First-level analysis

Beta coefficients for every direction and voxel were estimated using one mass univariate general linear model per participant. We modeled each movement direction of the fixation target separately for each condition, arena orientation and run with a separate boxcar regressor. Four additional regressors per run modeled (i) presentation of objects, (ii) cueing, (iii) periods during which the fixation target/arena remained stationary and (iv) periods during which behavioral responses were given. Regressors were convolved with the hemodynamic response function. Realignment parameters (translations of X, Y and Z coordinates; pitch, roll and yaw) were modeled as nuisance regressors. Slow signal drifts were removed using a high-pass cutoff filter of 100Hz and only voxels in gray matter and white matter were included in the analysis.

Hexadirectional analysis

We extracted first-level beta estimates for three data partitions (run blocks), representing the first three, intermediate three and last three runs of the experiment. Within each data partition, all directions were sampled equally while condition sequence (main and control conditions) was counterbalanced, and object presentations, arena orientations and the temporal sequence of trajectory rotations across runs were pseudorandomized. We extracted beta estimates for voxels in the entorhinal cortex, normalized each beta estimate by subtracting the mean across all beta estimates obtained from the same run (to compensate for potential differences in overall beta-estimate amplitude for task regressors between runs) and averaged two of the three data partitions to increase the signal to noise ratio. We then fitted another linear model incorporating two regressors modeling

sine and cosine across directions (ϕ) with 60° periodicity (sin(6ϕ) and cos(6ϕ)) as well as one regressor for the grand mean to this 'training set'. The resulting beta estimates for sine (βsin) and cosine (βcos) of viewing direction were then used to estimate the putative grid orientation (Φ), or the phase of the oscillatory modulation across directions, as in Φ=arctan(βsin βcos)/6

The third data partition served as the 'test set' and was used to contrast directions aligned (0 modulo 60°) versus misaligned (30 modulo 60°) relative to the estimated putative grid orientation in 20° bins. This process was iterated until each data partition served as test set once and as training set twice. Since directions were sampled in steps of 10°, estimated putative grid orientations would always fall between two actually sampled directions. The 20° bin size was chosen by taking both directions closest to the estimated putative grid orientation (rounded up and down to closest sampled directions). The resulting contrast coefficients were then transformed to z-scores as follows. In each iteration, we bootstrapped the null distribution for each voxel and data partition by shuffling the direction labels of the test set 1,000 times, each time contrasting aligned versus misaligned directions as described before. The training set was not shuffled. We calculated the mean (χ) and s.d. (σ) of the null distribution and used it to transform the contrast coefficient (β) to a z-score (z) as in z=(β-χ)/σ

We then averaged the three resulting z-scores of each voxel across the three test sets to obtain one threefold cross-validated z-score per voxel (Supplementary Fig. 3.6). By averaging these z-scores across arena orientations and voxels of an ROI, we obtained one z-score per participant that was taken to group level. The distribution of z-scores was non-normal (one-sample Kolmogorov-Smirnov goodness-of-fit tests for EC and sixfold rotational symmetry: K=0.367, P=5.2×10^{-4}). To assess statistical significance, we therefore report Wilcoxon signed-rank tests throughout the imaging analysis. The putative grid orientation represents one of the directions for which the fMRI signal is strongest (one of the putative six high peaks across directions). We then contrasted all directions falling on high peaks against all directions falling on low peaks of the putative six-peaked modulation (Fig.3.1a). If there is hexadirectional modulation, this contrast should yield positive values independently from the absolute putative grid orientation observed. Note that our main hypothesis is therefore inherently one-sided. We next re-examined the main effect on group level using a nonparametric permutation-based one-sample t test using 10,000 permutations. For each permutation, the signs of the group-level means were switched randomly, each time computing the t score of the permuted data and finally computing the P value for the observed t score based

on the resulting t score distribution. The test confirmed previous tests in suggesting significant hexadirectional modulation of the EC ($t(28)=2.36$, $P=0.0126$). Note that even though our analysis should be more conservative when performed on z-scores, it yielded similar results when performed on cross-validated contrast coefficients averaged across voxels before bootstrapping (Fig.3.2c and Supplementary Fig. 3.7). Even though we did observe clustering of putative grid orientations across voxels, our analysis does not rely on such clustering. While earlier approaches tested putative grid orientations on the ROI level [158], we estimated and tested putative grid orientations for each voxel separately and averaged the resulting z-scores. This way, we maximized sensitivity and increased robustness compared to previous approaches investigating hexadirectional modulation [158]. Since previous studies report stronger hexadirectional signals in right-hemispheric compared to left-hemispheric EC [158, 168], we repeated the main analysis for the EC on individual hemispheres (Supplementary Fig. 3.8). In line with previous reports, right-hemispheric EC elicited strong hexadirectional signals ($z=2.85$, $P=2.2\times10^{-3}$), while the left-hemispheric EC did not ($z=1.06$, $P=0.145$, one-sided Wilcoxon signed-rank test).

Matched ROI sizes

To ensure that potential hexadirectional signals were not washed out due to bigger ROI sizes for control regions compared to the entorhinal cortex, we repeated the analysis for control regions that were matched in size (see above). Again, neither early visual cortex ($z=-0.39$, $P=0.651$), motor cortex ($z=-0.49$, $P=0.687$) nor parietal cortex ($z=1.09$, $P=0.137$) showed any significant modulation. However, we did observe a weak hexadirectional signal in the frontal lobe ($z=1.92$, $P=0.027$), which did not survive Bonferroni correction (Supplementary Fig. 3.8b). The higher-threshold mask for frontal lobe contained voxels of the prefrontal cortex (PFC; Supplementary Fig. 3.8a), in a subset of which hexadirectional modulation has been observed previously [158, 168]. The weak hexadirectional modulation we observe parallels previous reports of hexadirectional modulation of the PFC.

Examination of putative grid orientations

To see how much orientation estimates varied across voxels in our data, we averaged beta estimates for each direction across the three normalized data partitions (to increase the signal-to-noise ratio) and estimated one putative grid orientation for every voxel with 10° resolution in 60° space (36 bins). We then calculated the circular distance between the orientation estimated for each voxel and the mean orientation across voxels. If putative grid orientations were random across voxels, we would expect a uniform distribution of circular distances around the unit

circle. However, if there was a within-participant clustering of putative grid orientations, we should also be able to detect it here across participants as the average putative grid orientation of each participant was set to zero in this test. A V-test for nonuniformity revealed that circular distances were not random across voxels but did indeed cluster around each participant's mean orientation ($V=31.80$, $P=3.46\times10^{-6}$). Next, we examined whether there was a relationship between the amplitude of the hexadirectional signal of a voxel and the putative grid orientation that was estimated for it. We calculated the circular variance across putative grid orientations for two voxel selections, the strongest 10% and weakest 10% of hexadirectionally modulated voxels of each participant, and found that stronger voxels had a more similar putative grid orientations (smaller across-voxel variances in putative grid orientations) than weaker voxels (one-sided t test results: $t(28)=-6.67 P=1.5\times10^{-7}$, CI [-Inf, -0.19]; Supplementary Fig. 3.9). To ensure that this clustering did not reflect dependencies introduced by data smoothing, we repeated the analysis using unsmoothed data. Again, the strongest 10% of voxels were more similar in putative grid orientation than the weakest 10% ($t(28)=-2.74$, $P=5.3\times10^{-3}$, CI [-Inf, -0.02]; Supplementary Fig. 3.9). To examine whether voxels in one hemisphere were more similar in their putative grid orientation than voxels in different hemispheres, we then used unsmoothed data to compare the average absolute angular differences between putative grid orientations of each voxel to all voxels in the same hemisphere and to all voxels in the other hemisphere. We found that voxels in the same hemisphere had more similar putative grid orientations than voxels in different hemispheres (two-sided t test results: $t\ 28=-3.67$, $P=0.001$, CI [-0.03, -0.01]). To investigate a potential anchoring of putative grid orientations to perceptual features such as landmarks or screen axes, we then used the strongest 10% of voxels to compute one average putative grid orientation per participant, as described above, using smoothed data. Potential clustering was tested for each arena orientation separately using Rayleigh's tests for nonuniformity, which did not reveal any clustering of putative grid orientations across participants (arena orientation 1: $z=0.083$, $P=0.922$; orientation 2: $z=0.141$, $P=0.871$).

Eye tracking

Left eye position as well as pupil size was monitored at 1,000Hz using a video-based infrared eye tracker (Eyelink 1000) with MR-compatible long range optics (for five participants eye tracking calibration failed and no eye tracking was recorded due to technical problems; the imaging main effect (Fig.3.2a) remains significant without these participants: aligned vs. misaligned directions: $z=2.16$, $P=0.016$). Eye tracking data (Supplementary Fig. 3.10) was linearly detrended and

smoothed with a running average kernel of 100 ms. To additionally compensate for slow signal drifts, the across-trial median x and y positions were set to zero for both gaze and the fixation target. Data was downsampled to the monitor refresh rate of 60Hz. Noise due to lost tracking or blinks was discarded by removing any samples for which pupil size deviated more than one s.d. from the mean across the time series. We then calculated the across-trial average Euclidean distance between fixation target and gaze position for each direction (fixation error). Since a repeated-measures ANOVA revealed differences in fixation error between directions ($F(35), 805=8.57$, $P=3.5\times10^{-36}$), we used an approach similar to that used in the imaging data analysis to test for potential symmetries of this effect. We fitted regressors for sine and cosine of eye movement direction for four-, five-, six-, seven- and eightfold rotational symmetry (plus constant term) to the fixation error across directions of each participant. We then performed another repeated-measures ANOVA to test for differences in the R 2 statistics of these models and did not find any difference produced by different symmetries ($F(4,92)=1.7$, $P=0.156$). Differences in fixation accuracy therefore cannot explain our imaging results. The average fixation error of each participant differed between the main experiment and visual motion control (two-sided t test results: $t(23)=12.65$, $P=7.6\times10^{-12}$, CI [0.79, 1.10]), which was expected given that the fixation target moved in the main experiment but not in the visual motion control. Participants with higher fixation errors in the main experiment also showed higher fixation errors in the visual motion control, which might reflect general differences across participants in target tracking ability, but could also reflect differences in eye tracking quality between participants. For five participants, fixation errors in the main experiment were higher than 1s.d. from the mean across participants, indicating poor tracking performance. These participants were excluded from imaging analysis. In addition to fixation error, we calculated the average eye movement velocity for all 36 directions and tested for differences with repeated-measures ANOVA but did not find any differences ($F(35, 805)=1.1$, $P=0.313$). The average eye velocity matched that of the fixation target (average eye velocity across participants: 7.75±0.183 °/s, constant velocity of fixation target: 7.5°/s). Taken together, these results are not suggestive of any hexadirectional biases and show that participants performed the visual tracking task.

3.6 Supplementary material

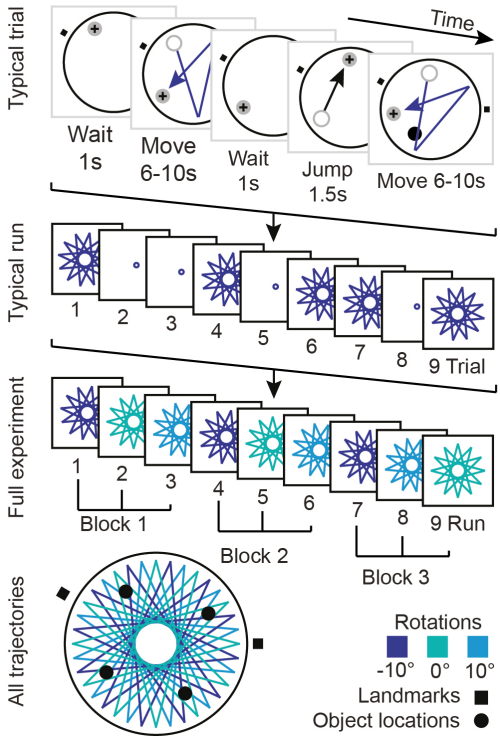


Figure 3.3: Visual tracking task. Typical trial: each trial started with the fixation target (black cross) being stationary (1s), then moving (6-10s) to sample 3-5 eye movement directions sequentially (in 2s blocks), followed by a waiting period (1s) and finally jumping to a random position and continuing moving from there. When the fixation target moved over predefined locations, objects were shown (represented by black circle). After 3 pseudo-randomly chosen trials per run, participants reported object locations by moving a cursor to the remembered position. Typical run: each run consisted of 9 trials, each testing 12 directions twice. Half of the trials tested a visual motion control. Condition sequence was counterbalanced. Full experiment: the full experiment consisted of 3 blocks with 3 runs each per participants. Each run tested a different set of 12 directions defined by trajectory rotation (-10°, 0°, 10°). Each run block tested all directions equally with 10° angular resolution. All trajectories tested in this experiment: The arena was presented in two orientations (0° and 180°) and was surrounded by two external landmarks depicted as black squares (0° and 150° in arena orientation 0°). Objects were cued at 4 predefined locations (black circles) that allowed a balanced sampling across trajectory rotations.

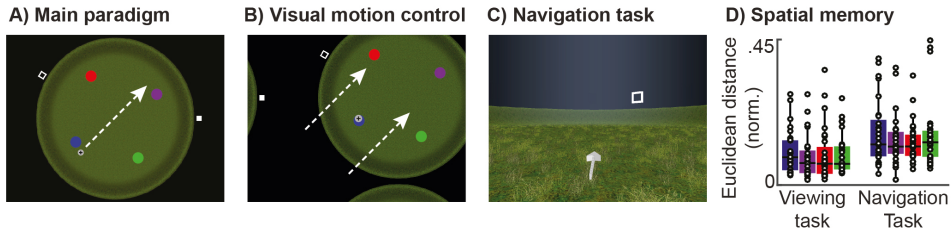


Figure 3.4: Stimuli and spatial memory performance. A) Main paradigm: virtual arena from bird's eye view, fixation target was moving. B) Visual motion control: fixation target remained at the screen center while arena was moving. Objects (colored circles) were shown only when the fixation target moved across them. C) First-person virtual navigation task. Participants navigated a virtual arena via button presses and reported object locations by navigating to them. White arrows indicate movement of either fixation disc (A), virtual arena (B) or the first-person agent (C) and were not shown during the experiment. D) Spatial memory performance for objects for both viewing-and first-person virtual navigation task. We plot the average Euclidean distance between memorized and true location normalized by the total size of the environment in the respective coordinate system (pixels vs. unreal coordinates) separately for the four objects across individual participants overlaid on box-whisker-plot (center: median, box: 25thto 75% percentile, whiskers: 1,5xIQR). No differences between objects were observed (repeated measures ANOVA, viewing task: $F(3, 81)=0.64$, $p=0.593$, navigation task: $F(3, 81)=1.9$, $p=0.136$, $n=28$).

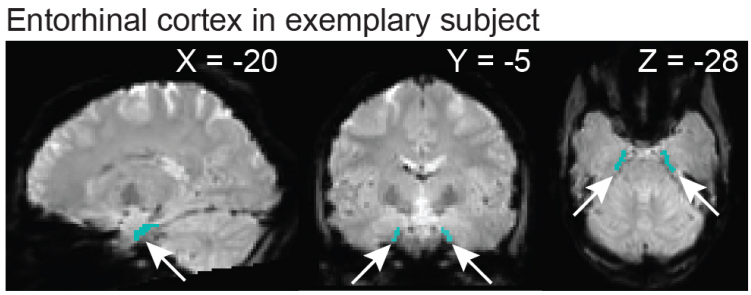


Figure 3.5: Entorhinal cortex. Binary regions of interest-mask for entorhinal cortex overlaid on single participant normalized mean-EPI image on which the ROI was drawn.

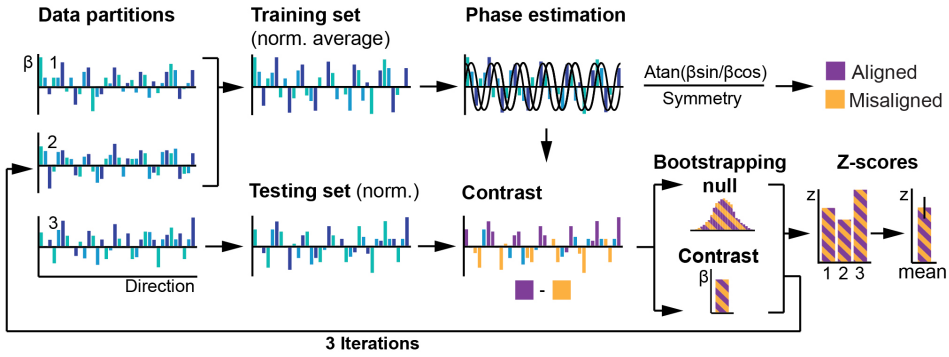


Figure 3.6: Analysis of hexadirectional activity. For each voxel we extracted beta estimates (β) for each direction for three data partitions within which directional sampling was balanced. From each beta estimate we subtracted the mean across all beta estimates obtained from the same run (normalization)and averaged it across two of the three partitions (training set). We then fitted regressors for sine and cosine of viewing direction with 60°-periodicity (and constant-regressor) to the training set and used the resulting beta estimates (βsin and βcos) to estimate the voxel's putative grid orientation. All directions in steps of 60° (0 modulo 60°) are considered aligned directions, those in-between (30 modulo 60°) are misaligned directions. In the third data partition (testing set), we then contrasted directions aligned to the putative grid orientation versus directions misaligned to it. This process was iterated until every data partition served as testing set onceand as training set twice. In each iteration, contrast coefficients were transformed into z-scores based on a bootstrapped null-distribution. By averaging across iterations and ROI-voxels we obtained one three-fold cross-validated z-score per participants that was taken to the group level.

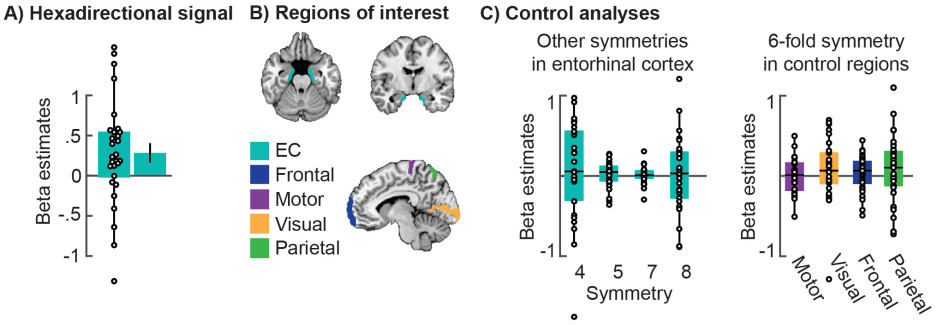


Figure 3.7: Visualization of main results when analysis is performed on threefold cross-validated contrast coefficients for Aligned versus Misaligned contrast instead of z-scores. A) Six-fold rotational symmetry in EC. We plot single participant data (n=29) overlaid on whisker-box-plot (center: median, box: 25thto 75% percentile, whiskers: 1,5xIQR) and mean and SEM across participants. B) Regions of interest on SPM single participant T1-template (MNI-coordinates: X=-8, Y=-11, Z=-20). C) Control symmetries for EC and six-fold rotational symmetry in control regions. We plot single participant data overlaid on whisker-box-plots.

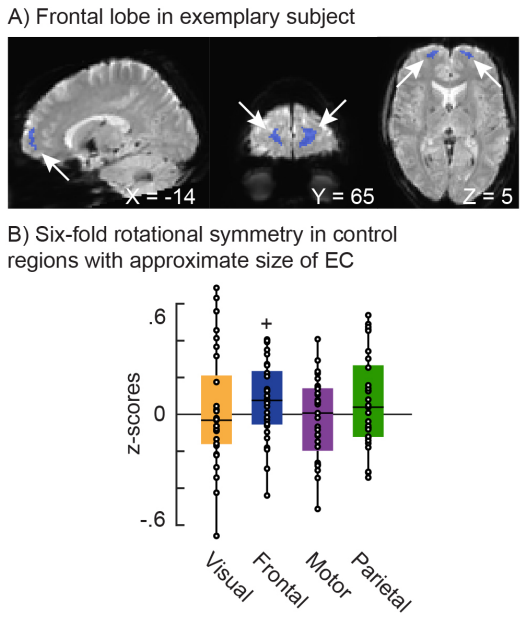


Figure 3.8: Results for ROIs matched in size to entorhinal cortex. A) Frontal lobe mask obtained by thresholding the corresponding probability map (SPM anatomy toolbox) at 99% probability to approximate the average size of the entorhinal cortex. Mask shown on single participant normalized mean-EPI image. B) Aligned versus misaligned contrast for six-fold rotational symmetryin control regions with sizes approximating the average size of entorhinal cortex. We plot single participant data overlaid on whisker-box-plots (center: median, box: 25thto 75 percentile, whiskers: 1,5xIQR). We found a weak hexadirectional signal in the prefrontal cortex (One-sided Wilcoxon signed rank test, z=1.92, +p=0.027, n=29) in line with previous reports [158, 168], which did not survive Bonferroni correction.

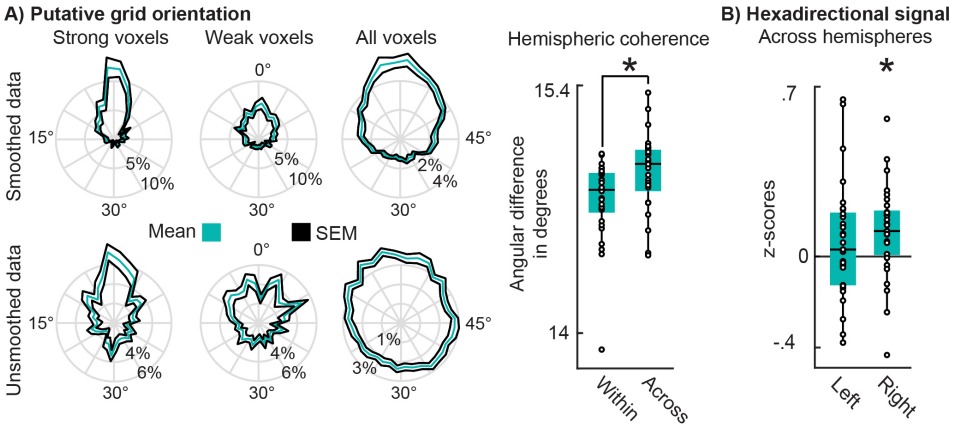


Figure 3.9: Putative grid orientation coherence and lateralization. A) Putative grid orientations across voxels and hemispheres. Smoothed (upper left panel) and unsmoothed data (lower left panel) depicted for all voxels as well as for the 10% strongest and 10% weakest hexadirectionally modulated voxels. Each plot depicts percentof voxels found for each possible putative grid orientation (10°-binning in 60°-space). The mean putative grid orientation of eachsubject was set to zero. The 10% strongest hexadirectionally modulated voxels had a more similar putative grid orientation (smaller variance) than the 10% weakest voxels for both smoothed (one-sided t-test, $t(28)=-6.67$ $p=1.5x10^{-7}$, CI [-Inf, -0.19]) and unsmoothed data (one-sided t-test, $t(28)=-2.74$, $p=5.3x10^{-3}$, CI [-Inf, -0.02]). Middle panel: Across-voxel coherence of putative grid orientations within each hemisphere compared to across hemispheres. We plot unsmoothed individual participant data overlaid on whisker-box-plots of the average absolute angular difference between putative grid orientations of each voxel to all voxels in the same hemisphere (within) and to all voxels in the other hemisphere (across). Voxels in the same hemisphere had a more similar putative grid orientation than voxels in different hemispheres (two-sided t-test, $t(28)=-3.67$, $p=0.001$, CI [-0.03, -0.01]). B) Main analysis for left and right hemisphere. Hexadirectional modulation was strong in the right hemisphere (one-sided Wilcoxonsigned rank test, $z=2.85$, $*p=2.2x10^{-3}$, n=29). Whiskerbox-plots show the median on the 25thto 75% percentile, whiskers represent 1,5xIQR.

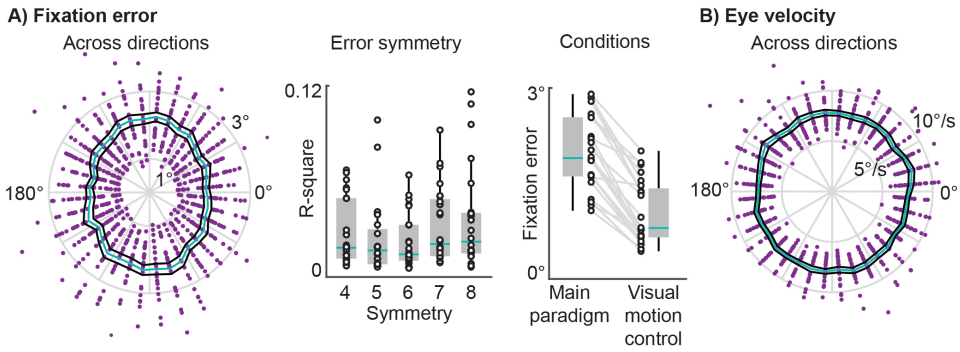
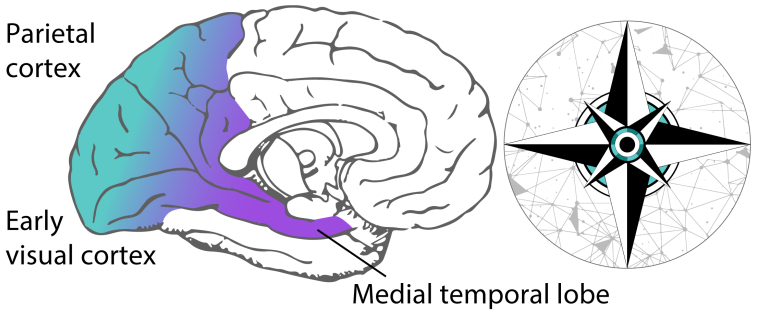


Figure 3.10: Gaze-dependent hexadirectional pattern cannot be explained by fixation error or eye velocity. A) Euclidean distance between fixation target and gaze (fixation error). Left: fixation error for all 36 directions and 24 participants in degree visual angle (radial axis). Across-participant mean in petrol green, SEM in black, single participant data in purple. Middle: R-square statistics for all participants and model symmetries tested. Data for individual participants (white dots) overlaid onwhisker-box-plot. We fitted linear models with regressors for sine and cosine of viewing direction with different periodicities (4, 5, 6, 7, 8-fold symmetry) to the fixation error (shown left) of each participant. If fixation accuracy was symmetrical, corresponding models were expected to produce higher R-square values andhence a higher goodness of fit. There was no difference in R-square statistics between model symmetries (repeated measures ANOVA, $F(4, 92)=1.7$, $p=0.156$). Right: fixation error for main paradigm and visualmotion control. We plot data for single participants next to whisker-box-plots (center: median, box: 25thto 75% percentile, whiskers: 1,5xIQR). B) Average velocity of eye movements in degree per second (radial axis) across all 36 directions. The fixation target movedwith a constant speed of 7.5°/s, average eye velocity across directions and participants was 7.75 ± 0.183. There was no difference in eye velocity across directions (repeated measures ANOVA, $F(35, 805)=1.1$, $p=0.313$). Across-participant mean in petrol green, SEM in black, single participant data in purple.

Chapter 4

Behavior-dependend directional tuning in the visual-navigation network



This chapter has been published as:
Behavior-dependend directional tuning in the human visual-navigation network. Nau, M., Navarro Schröder T., Frey M., Doeller, C.F. (2018). *bioRxiv 2019* https://doi.org/10.1101/765800

The data used in this work were collected at the Erwin L. Hahn Institute in Essen, Germany. All analyses and the writing of the manuscript were part of the doctoral work at NTNU.

4.1 Abstract

The brain derives cognitive maps from sensory experience to guide memory formation and behavior. Despite extensive efforts, it still remains unclear how the underlying population activity relates to active behavior and memory performance. Here, we combined 7T-fMRI with a kernel-based encoding model of virtual navigation to map world-centered directional tuning across the human cortex. First, we present an in-depth analysis of directional tuning in visual, retrosplenial and parahippocampal cortices as well as in the hippocampus. Second, we show that tuning strength, width and topology of the directional code during memory-guided navigation depend on successful encoding of the environment. Finally, we show that participants' locomotory state differentially influences this tuning in sensory and mnemonic regions such as the hippocampus. We demonstrate a direct link between neural population activity and cognition and show that high-level memory processing interacts with environmental coding in the service of behavior.

4.2 Introduction

Human scene processing and navigation regions interface between the lower-level sensory and higher-level cognitive domain. They gradually construct world-centered mnemonic representations of the environment, a cognitive mapping process thought to culminate in the medial temporal lobe (MTL) [96–99, 127, 315, 316]. Areas such as the retrosplenial cortex [96] and the parahippocampal cortex represent the spatial layout [104] and 3D structure [29] of the currently viewed scene, as well as its relative openness [105] or its boundaries [106]. Downstream regions like the entorhinal cortex and the hippocampus use this information to derive a stable representation of the world and one's own position, direction and speed in it [123, 124]. Together, such spatial representations are often referred to as 'cognitive map' and are thought to fundamentally shape our memories and guide behavior [129]. To understand this process, we believe it is critical to study the neural population activity of these regions in a naturalistic setting and in the light of the behavior they support.

A critical challenge the brain needs to solve to map the environment is keeping track of our own direction as we move. Previous studies revealed directional representations and activity related to heading perception in several areas including the medial parietal lobe and retrosplenial cortex [113, 114, 184, 185], the parahippocampal gyrus [114, 158, 162, 187–189], the entorhinal/subicular region [158, 185, 188, 317, 318], the thalamus [113] and the superior parietal cortex [192, 193]. Also, hippocampal and parahippocampal activity encodes heading in the horizontal plane [186]. Most of these studies used dedicated and constrained directional judgment- and mental imagery tasks and often examined direction in a self-centered frame of reference. To date, it remains unclear how cognitive mapping is mediated by the scene processing and navigation network and how active spatial behavior and memory relate to environmental processing in this pathway.

Here, we used 7T functional magnetic resonance imaging (fMRI) to monitor human brain activity during naturalistic virtual navigation in a spatial memory task (Fig.6.1). Inspired by prior successes of encoding models in characterizing fMRI responses in other domains [32, 33], we then developed an iterative kernel-based encoding model (Fig.6.2) of the participants' navigation behavior (Supplementary Fig.4.7) to map directional tuning across the human cortex. Importantly, our model treats direction in a world-centered frame of reference. We further analyzed the impact of spatial memory performance and locomotory states on this tuning, focusing on scene processing and navigation regions due to their proposed in-

volvement in cognitive mapping and known directional coding [98, 99, 315]. These regions include the early visual and retrosplenial cortex, the parahippocampal gyrus, the entorhinal cortex as well as the hippocampus.

Our objectives were twofold. First, we aimed to quantify and map directional tuning in the human scene processing and navigation network during active spatial behavior. Second, we examined how this tuning relates to the participant's behavior and memory.

4.3 Results

During fMRI scanning, participants freely navigated in a circular virtual reality (VR) arena via key presses while memorizing and reporting object locations within it (Fig.6.1A). Across different trials, participants indicated the locations of these hidden objects by navigating to them. After each trial, they received feedback about the true object location before the next trial started. We then tracked the improvement in memory performance over trials by assessing the 'memory error', i.e. the Euclidean distance between true and remembered location in each trial (Fig.6.1B, Supplementary Fig.4.8).

Our here developed encoding model analysis comprises multiple individual steps. We modelled world-centered virtual head direction (vHD) using basis sets of circul-ar-gaussian vHD-kernels (Fig.6.2A). Next, we estimated voxel-wise weights for each kernel with ridge regression using a training data set (Fig.6.2B, Supplementary Fig.4.9A). We then used these weights to predict the time course of each voxel in an independent test set (Fig.6.2C). We define directional tuning strength as the model performance, i.e. how well the model predicted the time course of a voxel or region. Finally, by iterating through multiple basis sets differing in the number and the full-width-at-half-maximum (FWHM) of the directional kernels (Supplementary Fig.4.9B) we also estimated the corresponding tuning width of each voxel. Tuning width was defined as the FWHM of the kernels leading to the best model performance. Finally, we examined the relationship between the tuning strength and width estimated for different regions, the participants' navigation behavior itself as well as their performance in the spatial memory task.

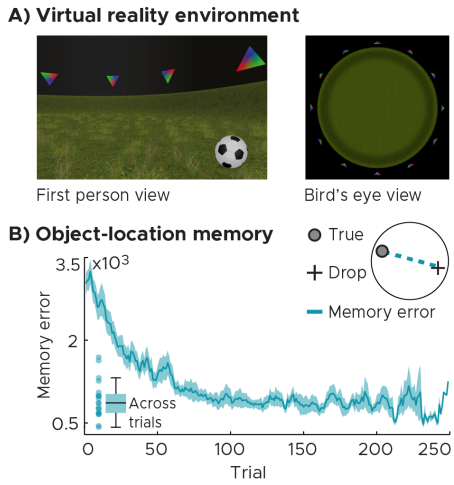


Figure 4.1: Spatial memory task in virtual reality (VR). A) First person and bird's eye view of the VR-environment in which participants navigated freely via key presses. The circular arena was surrounded by 12 landmarks matched in visual features (colored triangles). Across trials, participants memorized and reported object locations by navigating to them and pressing a 'drop' button followed by feedback. B) Object-location memory. Participants' memory performance improved as indicated by a decrease in memory error (Euclidean distance between drop and true location). The blue line and shaded area represent the mean and SEM of the memory error across participants. Data were smoothed with a moving average kernel of 5 trials. The inset depicts the median memory error across trials for single participants and as whisker-boxplot (center, median; box, 25th to 75th percentiles; whiskers, 1.5×interquartile range).

Our results are presented in three sections. First, we establish how the vHD-encoding model works by mapping directional tuning strength and width across the cortex. Second, we demonstrate that the strength, width and topology of this tuning depend on the participants' spatial memory performance. On a behavioral level, we show evidence that this likely relates to how well the environment has been encoded. Finally, we show that the tuning in both sensory and high-level mnemonic regions reflected the behavioral state of our participant's, i.e. whether they were moving or not. Notably, our results cannot be explained by biases in sampling (Supplementary Fig.4.7), model regularization or data quality (Supplementary Fig.4.9).

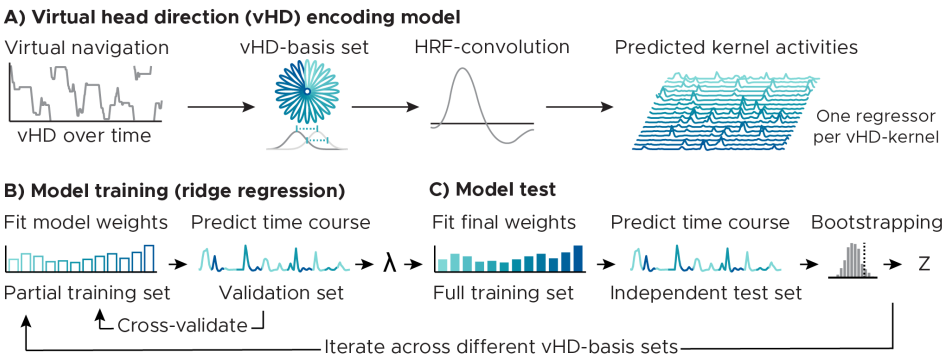


Figure 4.2: Analysis logic. A) Virtual head direction (vHD) encoding model. We modeled vHD using multiple basis sets of circular-gaussian kernels covering the full 360°. Given the observed vHD, we generated predicted time courses (regressors) for all kernels in each basis set. The basis sets differed in the (full-width-at-half-maximum) kernel width and number. Spacing and width were always matched to avoid over-representing certain directions. The resulting regressors were convolved with the hemodynamic response function (SPM12) to link the kernel activity over time to the fMRI-signal. B) Model training. We estimated voxel-wise weights for each regressor in a training data set (80% of all data of a participant or 4 runs, 10 minutes each) using ridge regression. To estimate the L2-regularization parameter (), we again split the training set into partial training (60% data, 3 runs) and validation sets (20% data, one run). Weights were estimated in the partial training set, and then used to predict the time course of the validation set via Pearson correlation. This was repeated for ten values of (log-spaced between 1-10.000.000) and cross-validated such that each training partition served as validation set once. We then used the that resulted in the highest average Pearson's R to fit the final model weights using the full training set. C) Model test: We used the final model weights to predict each voxel's time course in an independent test set (held-out 20% data, always the run halfway through the experiment) via Pearson correlation. These Pearson correlations were used to test model performance on a voxel-by-voxel level (Fig.6.3,6.4). For a regions of interest analysis (Fig.4.5,4.6), we additionally converted model performance into Z-scores via bootstrapping, ensuring that the results reflected effects of kernel width and not of number. The null distribution of each voxel was obtained by weight-shuffling (k = 500). Both model training and test were repeated for all basis sets.

Mapping directional tuning during spatial navigation

Participants navigated in a VR-environment, memorizing and reporting object locations within it. We used an iterative kernel-based voxel-wise encoding model of vHD to map directional tuning strength and width across the cortex (Fig.6.2). The model performance is the Pearson correlation between the voxel time course predicted by the model and the one observed in the test set. The model prediction builds on weights estimated for each kernel using an independent model training procedure.

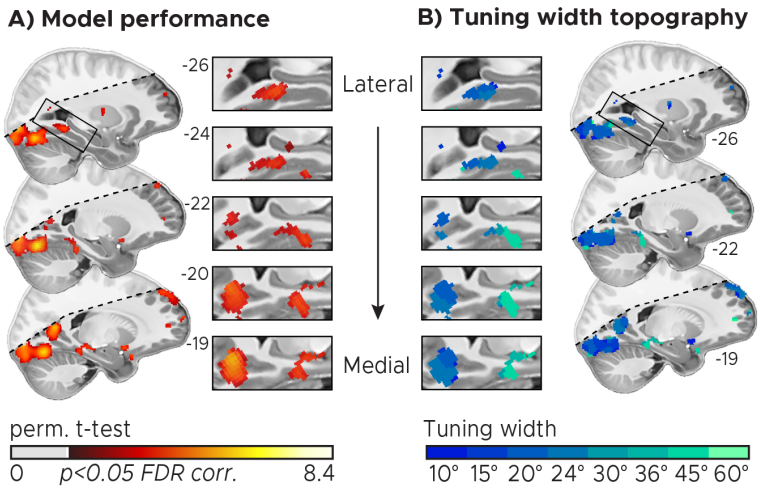


Figure 4.3: Mapping directional tuning across the human cortex (n = 20). A) Tuning map: directionally tuned voxels were determined by testing model performance against zero on group-level using a permutation-based one-sample t-test (Nichols and Holmes, 2002) with k = 10000 shuffles. We plot pseudo-T-maps thresholded at p<0.05, FDR-corrected for all basis sets at T1-resolution overlaid on the group average T1-scan. Approximate MNI coordinates were added. Inserts zoom in on the parahippocampal cortex and the medial parietal cortex. Multiple regions in the occipital lobe, the medial parietal and temporal lobes and the parahippocampal gyrus were directionally tuned. B) Tuning width: For each directionally tuned voxel, we color coded the median tuning width (full-width-at-half-maximum of the directional kernels in the optimal basis set) that led to the highest pseudo-T-value (depicted in A). The tuning width follows a narrow-to-broad topology along the parahippocampal long-axis. Shaded regions fell outside the scanning field of view in at least one participant.

Our model successfully predicted activity in multiple regions in the ventral occipital and medial parietal cortex as well as in the MTL (Fig.6.3A). These regions overlap with known scene processing and navigation regions such as the retrosplenial cortex [96] and the parahippocampal cortex as well as with the posterior hippocampal formation [99,315]. Along the parahippocampal long-axis, the tuning width followed a narrow-to-broad topology: narrow kernels best predicted activity in more posterior parts, wider kernels in anterior parts of the left-hemispheric parahippocampus (Fig.6.3B).

Directional tuning reflects spatial memory performance

After establishing that our vHD-encoding model did indeed predict activity in the visual-navigation network, we next asked whether the tuning was related to successful encoding of the environment and the object locations in it. We hypothesized that the tuning should be stronger in ventral visual stream and medial temporal regions in participants that performed well in the spatial memory task. This hypothesis built on the idea that stronger tuning should indicate enhanced retrieval of directional information from high-level mnemonic systems. To test this, we repeated the group level analysis depicted in Fig.6.3A, this time splitting the participants into two groups based on their across-trial median memory error (median split, Fig.6.1D).

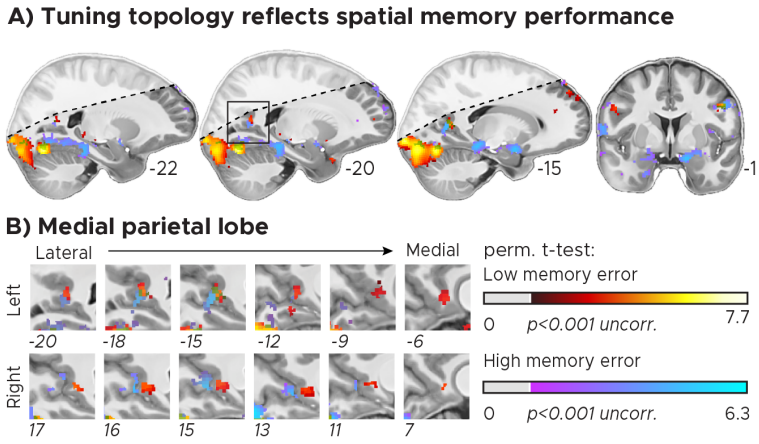


Figure 4.4: Directional tuning topology reflects spatial memory performance. Participants were split into two groups depending on their across-trial median memory error (2 x n = 10). A) Tuning maps: directionally tuned voxels tuning were determined by testing model performance against zero on group-level using a permutation-based t-test [319] with 1024 unique random possible shuffles. This procedure results in a minimal possible p-value of 0.00098, precluding FDR-correction. We therefore plot pseudo-T-maps thresholded at p<0.001 uncorrected for all basis sets at T1-resolution overlaid on the group-average T1-scan. Hot colors depict results for the low-memory-error group, cool colors for the high-memory-error group. Approximate MNI-coordinates added. B) Zoomed in depiction of the medial parietal lobe/retrosplenial cortex (RSC). There is an anterior-posterior distinction in directional tuning in RSC as a function of spatial memory performance. Shaded regions fell outside the scanning field of view in at least one participant.

We found that our model predicted activity in strikingly different networks in these two participant groups (Fig.6.4A), the direction of these effects however was opposite of what we had predicted. In participants with low memory error, i.e. good

memory performance, the model predicted activity in the medial and ventral occipital lobe. Strikingly, in participants with high memory error, the model predicted activity in the parahippocampal gyrus and in the MTL. A group-level permutation-based rank-correlation between memory error and model performance further indicated that these differences build on a systematic relationship between directional tuning and spatial memory (Supplementary Fig.4.10). In both groups, we observed bilateral clusters in the medial parietal lobe. These clusters however barely overlapped between groups (Fig.6.4B). In the low-memory-error group the model predicted activity in a more anterior part, in the high-memory-error group in a more posterior part of the medial parietal lobe, akin to previous reports of an anterior-posterior functional distinction in this region [25, 320].

To further characterize directional tuning explicitly in regions that derive world-centered representations of the visual environment, we next conducted a regions of interest (ROI) analysis focusing on the early visual cortex (EVC), the retrosplenial cortex (RSC), parahippocampal gyrus (PHG), the posteromedial entorhinal cortex (pmEC) and the hippocampus (HPC) (Fig.4.5A). We tested the pmEC subdivision of the entorhinal cortex because its rodent homologue region [321, 322] is known to encode direction [178–180]. We obtained the vHD-model performance for every voxel in our ROIs and every directional basis set as described before. For the ROI analysis, we added a bootstrapping procedure to ensure that our results reflected an effect of kernel width, and not of kernel number. For every basis set, we converted the Pearson correlations into Z-scores via weight-shuffling (Fig.6.2C) and then averaged across the 25% most reliable training voxels in each ROI to reduce noise (see methods for shuffling and voxel selection details). The resulting Z-scores expressed how well the model predicted the activity relative to the voxel's null distribution. For each participant group and ROI, we selected the basis set that led to the best model performance on average (Fig.4.5B). If a given region was not directionally tuned, the corresponding Z-scores should be zero. We tested this on group-level (one-tailed) as well as a difference between groups (two-tailed) using permutation-based t-tests (Fig.4.5C, see methods for details). We observed that EVC and RSC encoded direction in both participant groups (EVC: low memory error: $t(9) = 2.85$, $p = 0.014$, $pFDR = 0.048$; high memory error: $t(9) = 4.00$, $p = 0.004$, $pFDR = 0.027$; RSC, low: $t(9) = 3.04$, $p = 0.006$, $pFDR = 0.041$; high: $t(9) = 2.62$, $p = 0.015$, $pFDR = 0.040$). Importantly, in PHG and pmEC such tuning was observed only in participants that had a high memory error (PHG, low: $t(9) = 0.62$, $p = 0.320$; high: $t(9) = 2.16$, $p = 0.028$, $pFDR = 0.040$; pmEC, low: $t(9) = -0.42$, $p = 0.661$; high: $t(9) = 2.59$, $p = 0.020$, $pFDR = 0.040$, in line with the voxel-wise group results

(Fig.6.4). In pmEC, the tuning strength additionally differed between groups ($t(18) = 2.32$, $p = 0.036$). To again test whether this group difference reflected a systematic relationship between tuning strength and memory error, we conducted a post-hoc permutation-based rank-correlation between memory error and pmEC model performance on ROI level, which indeed seconded these results ($rho = 0.48$, $p = 0.035$, $k = 10000$, Supplementary Fig.4.10). Notably, while the rodent homologue of pmEC is known to encode world-centered direction [178–180], another entorhinal subregion, the anterolateral entorhinal cortex (alEC) is not. Consistently, we observed directional tuning only in pmEC, not in alEC (Supplementary Fig.4.12).

In addition to the tuning strength, our approach also allowed to estimate the tuning width for each ROI. For each individual participant and ROI, we selected the tuning width that led to the optimal model performance (Supplementary Fig.4.11) and compared it between groups. Strikingly, while above-mentioned model performance in RSC was matched, the tuning width differed between groups ($t(18) = 2.04$, $p = 0.044$). Participants with high memory error had a sharper tuning in RSC than participants with low memory error (Fig.4.5D).

In sum, while EVC, RSC, PHG and pmEC were directionally tuned in at least one of the participant groups, tuning strength in pmEC as well as tuning width in RSC strikingly reflected how well participants performed in the spatial memory task. This is in line with our hypothesis that the tuning should indicate whether the environment has been successfully encoded or not. However, we had hypothesized that stronger directional tuning in higher-level visual and MTL regions should be associated with better spatial memory performance. Our empirical test suggested the opposite. Why would mnemonic regions be more directionally tuned in participants that performed poorly in the spatial memory task?

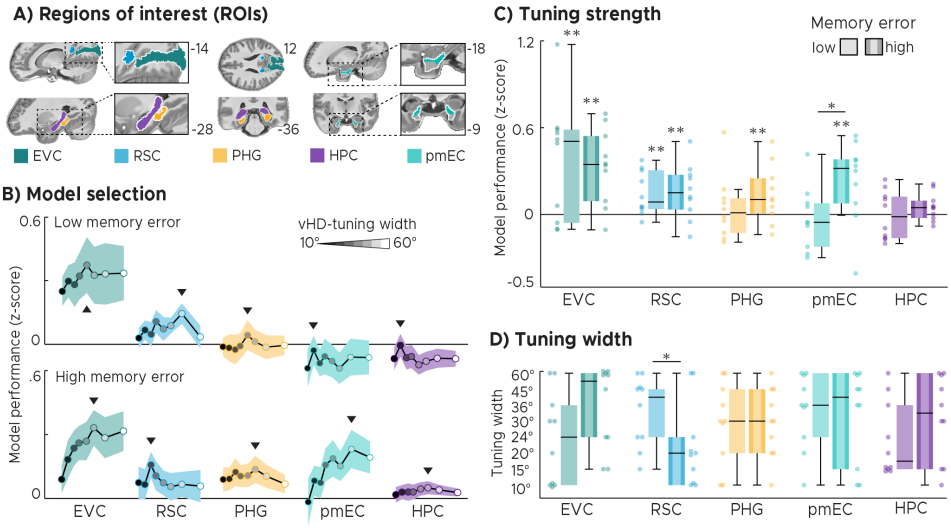


Figure 4.5: Regions of interest (ROI) analysis. A) Human scene processing and navigation regions tested in this study: Early visual cortex (EVC), Retrosplenial cortex (RSC), Parahippocampal gyrus (PHG), posteromedial entorhinal cortex (pmEC) and the hippocampus (HPC). B) Model selection: We plot the model performance (Z-score) on ROI-level for all basis sets. The black line and the shaded area represent the mean and SEM across participants. Each dot represents the group-average model performance for one basis set, with darker colors representing narrow kernels and lighter colors representing wider kernels. The following kernel widths were tested: 10°, 15°, 20°, 24°, 30°, 36°, 45°, 60°. The black triangles mark the basis set that lead to the optimal model performance. C) Optimal model performance for the two (high- and low-memory-error) participant groups. We plot single participant data and group-level whisker-boxplots (center, median; box, 25th to 75th percentiles; whiskers, 1.5×interquartile range). We observed directional tuning in EVC and RSC in both groups. In PHG and pmEC this tuning depended on spatial memory performance. D) Optimal tuning width. Similar to B,C we plot the tuning width that led to the highest memory performance selected on individual participant level. Participants with low memory error had wider tuning than the ones with high memory error. Hence, unlike tuning strength, the tuning width in RSC reflected spatial memory performance.

One explanation could be the following. Contrary to our initial hypothesis, the differences between groups could reflect the participants' ongoing effort in encoding rather than retrieving a map of the environment. If so, this predicted that participants with high median memory error were still in the process of mapping the environment and the object locations in it. In turn, they should be improving more in the spatial memory task than participants with an already low memory error. To test this post-hoc hypothesis, we analyzed the trial by trial memory error in more detail (Fig.6.1B, Supplementary Fig.4.8A). Indeed, we found that while both participant groups approached the same level of memory performance, the low-

memory-error participants reached ceiling earlier than the high-memory-error participants (Supplementary Fig.4.8A,B). We also estimated the learning rate for each participant, i.e. the slope of a linear regression line fit to the memory error across trials. The high-memory-error participants had steeper slopes ($t(18) = 2.65$, $p = 0.019$, $k = 10000$, Supplementary Fig.4.8C), again suggesting stronger improvements in memory performance compared to the low-memory-error participants.

Directional tuning reflects the behavioral state

To investigate the relationship between directional tuning and the participants' behavior further, we repeated above-described ROI-analysis twice, once only modeling periods in which the participants moved, and once in which they stood still. In both cases, the participants rotated (Supplementary Fig.4.7A,B). We compared these two scenarios by contrasting the respective model performances (Fig.4.6A), revealing a positive effect of locomotion on tuning strength in EVC and RSC (EVC: locomotion: $t(19) = 4.04$, $p = 0.0004$, $pFDR = 0.003$; stationary: $t(19) = 1.38$, $p = 0.091$; contrast: $t(19) = 2.06$, $p = 0.049$; RSC: locomotion: $t(19) = 3.68$, $p = 0.001$, $pFDR = 0.004$; stationary: $t(19) = -0.82$, $p = 0.79$; contrast: $t(19) = 2.20$, $p = 0.043$). Conversely, the activity in PHG and HPC could be better predicted while participants stood still (PHG: locomotion: $t(19) = -0.91$, $p = 0.799$; stationary: $t(19) = 2.37$, $p = 0.015$; contrast: $t(19) = -2.35$, $p = 0.024$; HPC: locomotion: $t(19) = -3.48$, $p = 0.999$; stationary: $t(19) = 1.77$, $p = 0.046$; contrast: $t(19) = -3.18$, $p = 0.004$, $pFDR = 0.028$, Fig.4.5C). Our results suggest that the directional tuning in human scene processing and navigation regions reflects not only spatial memory performance, but also the behavioral state of the participants (i.e. whether they move or not). Notably, the fact that the differences in tuning strength and width were region-specific (Fig.4.6A), suggest that our results cannot be explained by general differences in statistical power between the two behavioral states.

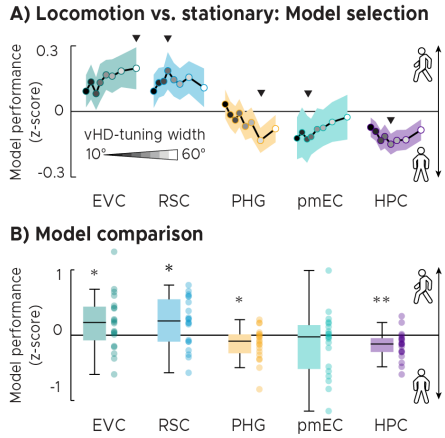


Figure 4.6: Behavioral state analysis. The analysis described in Fig.6.2 was repeated, this time separating periods when participants navigated and when they stood still. A) Model selection. We plot the difference in model performance (Z-score) between locomotion and stationary periods across tuning widths (grayscale dots represent tuning width: narrow:dark, wide:light., also see Fig.4.5). Positive values indicate that voxel time courses in an ROI could be better predicted when participants locomoted. Negative values indicate the opposite, with better model performance during stationary periods. Triangles mark the kernel width leading to the strongest difference between models. B) Model comparison. We plot the difference in model performance indicated in (A) as single participant data and group-level whisker-boxplots (center, median; box, 25th to 75th percentiles; whiskers, 1.5×interquartile range). EVC and RSC tended to be better predicted during locomotion, PHG and HPC could be better predicted during stationary periods. These results suggest that the tuning in visual and mnemonic regions depend on the locomotory state.

4.4 Discussion

The present study investigated fMRI-proxies of neural population activity reflecting directional coding during active spatial behavior. We put a focus on human scene processing and navigation regions due to their proposed involvement in cognitive mapping: they derive world-centered mnemonic representations of the environment from sensory experiences. We used 7T-fMRI to monitor brain activity of participants navigating in a virtual environment and performing a spatial memory task. We developed an iterative kernel-based encoding model of the navigation behavior to map directional tuning across the human cortex. In addition, examine its relationship to behavior and memory in detail. Visual, retrosplenial, parahippocampal, entorhinal and hippocampal regions showed distinct response profiles, with a narrow-to-broad tuning width topology along the left-hemispheric parahippocampal long-axis. Furthermore, we examined the relationship between

the tuning in each region, the participants' navigation behavior and the performance in the spatial memory task. We found that the tuning in the RSC and pmEC, and likely in the parahippocampal gyrus, reflected how accurately participants reported the location of objects in the environment. Strikingly, while it was the tuning strength in pmEC and PHG, it was the tuning width and topology in RSC that depended on memory. The strength, width and topology of directional tuning were therefore associated with the spatial memory performance of the participants, which demonstrates a direct link between neural population coding and cognition. The direction of this effect as well as behavioral evidence speaks to the idea that the underlying mechanism is related to encoding rather than retrieval of environmental information. Finally, the tuning in visual, retrosplenial and parahippocampal regions, but especially in the hippocampus, additionally signaled the behavioral state of the participants.

Directional representations in the human brain

Our observations emphasize the central role of scene processing and navigation regions in spatial cognition [96–99, 127, 315, 316, 323] and are consistent with previous work on directional representations in the human brain [113, 114, 158, 162, 184–188, 317, 318]. They are also consistent with lesion studies showing that damage to regions like the RSC can impair the ability to orient oneself relative to landmarks [116]. Importantly, the present study goes beyond previous reports in several aspects. Most studies used dedicated and constrained directional judgment- and imagery tasks to reveal directional representations in the brain. Here, we examined directional tuning during active naturalistic navigation (in VR) and additionally demonstrate that it depends on behavioral factors and on memory. Many of these studies also examined self-centered directional coding, whereas our approach examines vHD explicitly in a world-centered frame of reference. Furthermore, using an iterative encoding model we were also able to extract additional parameters such as the tuning width from fMRI responses. Similar approaches have been used for example to map the retinotopic [17, 18] and semantic organization of the cortex [30, 324], the 3D-depth tuning of scene processing regions [29], to identify perceived stimuli from brain activity [325, 326] and to reconstruct the online content of working memory [31, 327]. Unlike previous work, the encoding model developed here (Fig.6.2) does not build on information about a stimulus but importantly is informed by the behavior of our participants directly.

Visual information and head direction

An open question is whether the tuning observed here reflects processing of visual information or head direction (HD). Neurons representing world-centered HD, or HD-cells, are abundant in the brain and have been studied most intensively in rodents [175], but also in monkeys [181]. HD-cells reference facing direction relative to known landmarks [176] and are often compared to an 'internal compass' mediating our sense of direction [175, 328]. This HD-cell compass plays a central role in cognitive mapping and is thought to mediate homing, reorientation and path integration behavior [175,328–333]. We observed that multiple brain areas encoded direction, many of which overlap with regions known to contain HD-cells in rodents and monkey. These regions include the RSC [110–112], the postsubiculum (part of the hippocampal formation) [177] and the entorhinal cortex (EC) [178–180]. The latter consists of at least two subdivisions in rodents, the medial (MEC) and the lateral entorhinal cortex (LEC), likely corresponding to the pmEC and alEC in humans [321, 322]. HD-cells have been observed only in the MEC, not the LEC, paralleling our observations of world-centered directional tuning in the human pmEC, not in the alEC (Supplementary Fig.4.12).

We also observed directional tuning in the early visual cortex for which no HD-cells have been reported to date. This raises the question whether the effects reported here are due to lower-level sensory processing, not HD? We believe these two options are not exclusive and the underlying processes might strongly interact in our naturalistic task. Locomotion [334, 335] and world-centered location [336, 337] have for example been shown to modulate EVC activity in rodents. Here, we observed that locomotion had a positive effect on model performance in EVC and RSC, suggesting stronger directional tuning in these regions during running (Fig.4.6). Also, in monkeys [47] and humans [247], the EVC can represent the velocity of motion in a world-centered frame of reference, likely due to feedback from higher-level areas. Critically, high-level mnemonic regions such as the entorhinal cortex are not known to be visually responsive, yet we still observed a directional code there. Consistent with this are reports of HD-cells in MEC, the rodent homologue of the pmEC [178–180] as well as its well-connected position within the HD-circuit [152]. Also, locomotion had a positive effect on model performance in EVC and a negative effect in the hippocampus, again contrasting higher-level MTL function against early visual processing. Areas such as the PHG code direction even in the absence of visual experience [162]. In addition, shared information and activity covariations between the EVC and the hippocampal formation suggest that early perceptual processing could well be modulated by higher-level

cognitive processes [127, 217, 336–340].

Interestingly, the fact that the directional tuning we observed depends on behavior is at odds with the activity profile of 'classical' HD-cells. First, these cells provide a directional representation that is continuous (i.e. they always maintain similar activity on population-level following attractor-dynamics [175, 199, 329, 341]. Second, impairments or even lesions of the HD-cell system only have very moderate effects on behavior [329]. However, growing evidence suggests that there are at least two types of HD-cells in the brain: 'classical' continuous HD-cells, as well as non-continuous sensory HD-cells [329]. The latter have been shown to switch between active and inactive states, providing a directional representation that is controlled by visual landmarks instead of attractor-dynamics [178]. Our present results are in line with the function and the location of these sensory HD-cells found in the entorhinal and parahippocampal cortex [178] as well as potentially in the RSC [112]. Some of these cells also alternate between several preferred directions depending on context, potentially explaining the variation in RSC tuning width we observed (Fig.4.5D). On a population level, directional activity referenced to multiple landmarks would likely be interpreted as having broader tuning curves than the stable unidirectional counterparts. While classical HD-cells integrate vestibular inputs [175, 341], which were not available in our VR-task, visual information alone is sufficient to drive HD-cell-like coding in VR [342] and to determine its alignment to the environment [343]. Behavioral evidence further suggests that the visual environment, specifically its geometry, influences our sense of direction [344, 345]. In our task, the VR-arena was circular and all landmarks were carefully matched with respect to their visual appearance (Fig.6.1). In one half of the arena the triangle-shaped landmarks faced upwards, in the other half downwards (Fig.6.1B). The hypothetical axis in between did not contain any landmarks and did not bias vHD-sampling (Supplementary Fig.4.7).

Taken together, low-level sensory information alone can explain neither the tuning we observed, nor its relationship to behavioral performance. Instead, our results are well in line with previous work on higher-level spatial representations. In a naturalistic task such as ours, this necessarily includes sensory experiences. Our results support the notion that cognitive processes are closely intertwined with visual processing up to its earliest stages in the brain.

Influence of gaze direction

One limitation of the current study is that we cannot differentiate between vHD and the direction of gaze. Other studies using eye tracking indeed suggested that the landmarks capture most of the viewing [163]. Neurons that encode the allocentric location of gaze could hence in theory lead to a directionality on population level if location and gaze are correlated. Such cells exist in the primate hippocampus and encode allocentric view independent from head direction [216,219] or together with HD and self-location in a conjunctive code [223]. Computational models proposed that it is the conjunction between location and direction that mediates the environmental anchoring of HD in the RSC [109]. In monkeys [229] and humans [170–172], eye movement direction has also been shown to modulate activity in the entorhinal cortex in which we also observed a directional code. However, these studies focused on eye movements in the 2D plane (e.g. left-right, up-down) and the results are likely a manifestation of another grid-cell related mechanism [127, 225, 228]. Since gaze direction typically varies more broadly over time than vHD, estimated tuning curves in visually responsive areas might broaden. This should have a stronger influence on visual compared to HD-representations, potentially explaining the relatively broad tuning in EVC (Fig.4.5) as well as the influence of locomotion (Fig.4.6). If correct, this predicted that early visual cortex should show a sharper tuning when analyzed as a function of gaze direction compared to vHD. However, neurons might also represent information in different reference frames depending on task demand [89]. Future studies using eye tracking could address these questions directly. Acknowledging this ambiguity, we here refer to 'directional tuning' more generally.

Memory retrieval and learning

We hypothesized that participants with stronger directional tuning in higher-level visual and mnemonic regions should perform better in the spatial memory task. We based this prediction on the idea that stronger tuning indicated enhanced retrieval of directional information from memory. Our empirical test showed the opposite: participants with stronger tuning in higher-level mnemonic regions performed worse in the spatial memory task. Neither data quality or model parameters (Supplementary Fig.4.9), nor directional sampling (Supplementary Fig.4.7) could explain these results. While the underlying mechanism remains unclear, we can speculate about its nature. For that, we feel it is important to revisit what 'tuning' means in this context. The 'tuning strength' describes how well the time course of a voxel could be predicted. This does not necessarily constitute a net increase or decrease in activity, but instead how well we could predict the fluctu-

ations over time. A voxel that weakly but consistently follows a directional modulation could hence show stronger tuning than one that strongly codes direction at very few time points.

If the tuning reflected memory encoding instead of retrieval, the tuning should be stronger in participants that kept improving in the task compared to participants who already reached ceiling. Accordingly, we tested whether the high-memory-error group kept improving, and whether the low-memory-error group had already reached performance ceiling. We tested this post-hoc behaviorally, revealing that this was indeed the case. Both participant groups approached the same level of performance and participants with stronger tuning in higher-level regions showed steeper learning curves than those with weaker tuning (Supplementary Fig.4.8C). A similar pattern of results has been reported for the hippocampus, whose activity tracks the amount of knowledge obtained at a given time rather than the accumulated absolute knowledge [346].

In the present study, we did not investigate how directional representations develop over time. Yet, our observations are in line with previous work on the emergence of spatial representations in scene processing and navigation regions. For example, the RSC and PHG were shown to rapidly encode a novel environment by integrating information across different viewpoints and landmarks. Spatial representations emerging in the RSC were associated with a participant's ability to identify multiple scenes as belonging to the same or different location [347], as well as with wayfinding ability by registering landmark permanence [348]. Consistent with the present results, the RSC activity could hence signal whether an environment has been successfully encoded or not. We observed a broader tuning in RSC in well compared to poorly performing participants, potentially indicating that the RSC processed information with a broader field of view. This could aid the integration of different landmarks and viewpoints [347, 348]. Importantly, the medial parietal lobe, which comprises the RSC, is known for both its perceptual and mnemonic capacities [349], which might be topologically distributed. Previous work revealed an anterior-posterior distinction for scene construction and scene perception [25, 99, 320]. Our results (Fig.6.4B) are consistent with this observation in two aspects. First, we also observed an anterior-posterior split on group-level in bilateral medial parietal lobe as a function of memory performance (Fig.6.4B). Second, the participant group with more anterior tuning also performed better in the spatial memory task. Given above-mentioned reports, we speculate that this topology could indicate enhanced scene construction in participants with good memory performance. Again, this would also be consistent with the idea that

these participants have already successfully encoded the environment and the object locations in it.

Notably, not only the RSC (Fig.4.5C), but also the pmEC tuning (Fig.4.5B) depended on the participants behavior, suggesting a different but related effect in pmEC. In line with this, rodent reorientation behavior depends on the stability of entorhinal HD-cell firing relative to the environment [333]. This alignment is also tightly coupled with the one of other spatial codes such as the hippocampal place field map [350], which in turn predicts goal-oriented navigation behavior [351]. Since the pmEC is the key mediator of hippocampal-cortical communication [152] and since the fMRI signal reflects input and local processing rather than output spiking [352], the present effects could therefore also reflect a related mechanism orchestrated by the hippocampus. This might include a reduction in hippocampal activity as environments become more familiar [353] or the successful encoding of object locations relative to the boundary [354]. Finally, the effects could also reflect individual preferences for egocentric versus allocentric navigation strategies [355].

Future studies employing more dynamic model tests such as inverted encoding modeling (proof of principle in Supplementary Fig.4.13) could address these questions directly by reconstructing the modeled behavioral features more dynamically [32, 33, 356–359]. In addition, high temporal resolution measures such as magnetoencephalography (MEG) or intracranial recordings in combination with VR-tasks [165] could monitor trial by trial changes in tuning as a function of behavioral performance.

Conclusion

Using virtual reality, a novel behavioral encoding model and 7T-fMRI, we examined world-centered directional tuning during active spatial behavior in humans. We demonstrated such tuning in visual, retrosplenial and parahippocampal cortices and the hippocampus. By mapping the tuning width across the cortex, we revealed a narrow-to-broad organization along the parahippocampal long-axis. Entorhinal and parahippocampal tuning strength as well as retrosplenial tuning width and topology reflected how well participants performed in a spatial memory task. We provide evidence that these effects likely depend on the encoding of the environment and the object locations within it. Finally, we show that the tuning in visual, retrosplenial and parahippocampal cortices as well as the hippocampus reflects the locomotory state. These results show the efficacy of encoding models for studying neural population dynamics during naturalistic navigation in human

fMRI. They demonstrate a direct link between neural population coding and cognition and show that high-level cognitive processes modulate directional tuning in the service of behavior.

4.5 Extended methods

Participants

We recruited 26 participants for this study (11 females, 19-36 years old). Four subjects were excluded because of excessive head motion, i.e. the number of instantaneous movements larger than 0.5mm [360] exceeded the across-participant average for more than one standard deviation. Another 2 participants were excluded because they finished fewer than four scanning runs. A total of 20 participants entered the analysis. The study was approved by the local research ethics committees (ethics committee University Duisburg-Essen, Germany and CMO region Arnhem-Nijmegen, NL) and participants gave written consent prior to scanning.

Virtual reality task

Participants performed a self-paced object-location memory task in virtual reality (Fig.6.1A) adapted from Doeller and colleagues [158]. The circular virtual arena was created using the UnrealEngine2 Runtime software and was surrounded by 12 distinct landmarks positioned in steps of 30° and matched in visual similarity (triangles either tilted up- or downwards, with red, green and blue colored corners). Participants could freely navigate in this arena via key-presses. The smallest instantaneous rotational movement possible was 10° and translational movement speed was constant after a 500ms ramp. In the beginning of the experiment, six everyday objects had to be collected, which were scattered across the arena. Across different trials and without the objects being present, participants were prompted to navigate to the location of a previously cued object. After indicating the remembered location via key press (drop), the respective object appeared at the correct location to give feedback and participant collected the object again before the next trial began. After an average of 3 trials (range 2-4), a fixation cross was presented on a gray background for 4 seconds. An average of 179 trials were performed (range: 94-253 trials due to the self-paced nature of the task, Supplementary Fig.4.8D) and object locations were randomized across participants. In order to explain the task and to familiarize participants with it, they performed a similar task with different objects in a different virtual environment prior to scanning. We tracked the improvement in memory performance over trials by assessing the 'memory error', i.e. the Euclidean distance between true and remembered

location in each trial measured in virtual vertices (arbitrary units).

Behavioral analysis

To ensure that there were no prominent or distinct directional cues that biased navigation behavior, we ruled out differences in the time spent facing into different directions. Supplementary Fig.4.7 depicts the results of these analyses for all participants and time points, split into locomotion and stationary periods as well as into high- and low-memory-error participant groups. We accounted for individual differences in absolute time spent in the experiment by expressing time spent as percent of the total experimental duration. We binned vHD in steps of 10° and performed a repeated-measure (rm) ANOVA across directions, which did not reveal any biases in directional sampling ($F(35, 665) = 0.77$, $p = 0.834$), also not when testing the high-memory error group ($F(35, 315) = 0.55$, $p = 0.984$) or the low-memory error group ($F(35, 315) = 0.87$, $p = 0.675$) individually. In addition, we did not observe biases in directional sampling across the experiment when splitting the data into locomotion ($F(35, 665) = 0.79$, $p = 0.806$) and stationary ($F(35, 665) = 0.87$, $p = 0.681$) periods (Supplementary Fig.4.7A).

In addition to the directional sampling in the course of the experiment, we analyzed the distribution of vHD within each TR. We again binned vHD into 10° steps and converted it into percent of total viewing time. We then circular-shifted each of the resulting histograms such that the most sampled direction lined up across TRs (), revealing that participants spent 52% of the time within each TR facing into a single direction (Fig.6.1B, Supplementary Fig.4.7A). The distribution of vHD within each TR was therefore non-uniform and centered on one predominant direction. We used a two-tailed permutation-based unpaired t-test to compare the time spent facing towards this predominant direction within each TR across across groups, which did not reveal a difference ($t(18) = -1.26$, $p = 0.224$, $k = 10000$, Supplementary Fig.4.7C). Using a paired version of this test, we did observed a difference between stationary and locomotion periods ($t(19) = 6.12$, $p = 0.0001$, $k = 10000$). Importantly, this had a differential effect for different regions of interests (Fig.4.6), suggesting that there was no general positive or negative effect on model performance. During this task, participants spent around 54% of their time navigating, with shorter time spent rotating during locomotion compared to stationary periods (two-tailed permutation-based paired ttest: $t(38) = 7.95$, $p = 0.0001$, $k = 10000$). There were no differences in the time spent translating ($t(18) = 0.41$, $p = 0.690$, $k = 10000$) or rotating between participant groups (during locomotion: $t(18) = 0.55$, $p = 0.625$, $k = 10000$ and during stationary periods: $t(18) = 1.00$, $p = 0.338$, $k = 10000$).

MRI acquisition

During the object-location memory task in VR we acquired T2*-weighted functional images on a 7T Siemens MAGNETOM scanner using a 3D-EPI pulse sequence [361], a 32-channel head coil and following parameters: TR = 2756 ms, TE = 20 ms, flip angle = 14°, voxel size = 0.9 mm x 0.9 mm, slice thickness = 0.92 mm, slice oversampling = 8.3%, 96 slices with a 210 mm x 210 mm field of view, phase encoding acceleration factor = 4, 3D acceleration factor = 2. The first 5 volumes of each run were discarded. Functional images were acquired across 5 scanning runs of 210 TR's or approximately 10 minutes each. In addition, we acquired T1-weighted structural images (MP2RAGE; voxel size: 0.63 mm isotropic) and a B0-field map (gradient echo; voxel size: 1.8 x 1.8 x 2.2 mm^3) for each participant.

Preprocessing

The data used here were used in two previous reports [163, 322]. Data were preprocessed using the automatic analysis library [362], utilizing functions of several analysis packages. For each participant, functional images were realigned using SPM8, followed by independent component analysis (ICA) denoising using FIX artifact removal implemented in FSL 5.0.4. To improve signal-to-noise ratio, and with it the ICA-detection of noise components, data were smoothed with a Gaussian full-width-at-half-maximum kernel of 2.5 mm. Images were then non-linearly normalized to a group-average EPI-template using the Advanced Neuroimaging Toolbox (ANTs) [363] and high-pass filtered with a 128-s cutoff using FSL. Voxelwise variance explained by the six realignment parameters (x,y,z, pitch, roll, yaw) as well as by spikes (sudden deviations in signal intensity of more than two temporal standard deviations) was removed via nuisance regression. Out-of-brain voxels were excluded.

Regions of interests (ROIs)

In an ROI analysis we tested human scene processing and navigation regions (Fig.4.5A) that were previously proposed to support cognitive mapping [97]. The hippocampal (HPC), anterolateral entorhinal (alEC) and posteromedial entorhinal (pmEC) ROIs were defined manually using ItK-SNAP (www.itksnap.org) based on the high-resolution group average EPI-template. The entorhinal masks were based on previous reports [322], in which the entorhinal mask was divided into anterolateral (alEC) and posteromedial entorhinal cortex (pmEC). The ROIs for the parahippocampal gyrus (PHG) as well as the retrosplenial cortex (RSC) were based on the reverse inference meta - analysis for "Retrosplenial cortex" and "Parahippocampal cortex" using Neurosynth [364]. We took the top 5% highest probability voxels from each respective Neurosynth map and removed isolated voxels from

the resulting binary masks. This procedure resulted in coherent bilateral clusters in the medial parietal cortex and parahippocampal gyrus respectively. The early visual cortex (EVC) ROI was created by thresholding the corresponding probability map 'Visual_hOc1' of the SPM anatomy toolbox at 50% and co-registering it non-linearly to our group-average template space. To do so, we used SPM to first segment the group template and then to 'normalize' the ROI into our template-space using the resulting tissue maps and nearest-neighbor interpolation. The resulting ROI masks were located at following average MNI coordinates [X,Y,Z] and were of following size. EVC: left hemisphere [-4, -88, 0], right hemisphere [14, -86, 0], n voxels: 16299, RSC: [-14, -56, 12] and [18, -54, 14], n voxels: 1926, PHG: [-26, -38, -12] and [26, -34, -16], n voxels: 2392, HPC: [-24, -24, -14] and [28, -22, -14], n voxels: 9781, alEC: [-20, 0, -34] and [22, 0, -34], n voxels: 1693, pmEC: [-20, -10, -28] and [20, -8, -28], n voxels: 1692.

Analysis overview

Our analysis estimated the directional tuning of a voxel in several steps. First, we built a vHD-encoding model by incorporating the participant's navigation behavior into basis sets of circular-gaussian von-Mises distributions, which we call vHD-kernels. Each individual direction was modeled with a different vHD-kernel, each representing a smooth directional tuning without discretizing the data into bins. Second, we estimated voxel-wise weights for each of these kernels, together representing a voxel's tuning curve. We refer to this step as model training. Third, we used these weights to predict activity in held out data which constituted the model test. This way we obtained a measure of model performance for the given vHD basis set. Finally, by iteratively varying the full-width-at-half-maximum of the vHD-kernels in the basis set and repeating above mentioned steps, we not only tested one vHD-basis set, but multiple ones. This approach allowed us to also estimate the tuning width of each voxel (the kernel width that maximized prediction accuracy). All processing steps mentioned comprised several individual sub-steps, which are described in detail below.

Building the virtual head direction (vHD) encoding model

We modeled vHD using a basis set of circular-Gaussian von Mises distributions as implemented in the Circular Statistics Toolbox for Matlab [365]. Each kernel in this basis set covered the full 360° with an angular resolution of 1°. Across different iterations of our analysis, we varied the full-width-at-half-maximum of these kernels, each time testing how well the resulting model weights allowed to predict activity in held-out data. To balance directional sensitivity across iterations, the spacing between kernels always matched the kernel width (i.e. the

broader each individual kernel, the fewer kernels were used). We tested following kernel widths, all representing divisibles of 360°: 10°, 15°, 20°, 24°, 30°, 36°, 45°, 60°. Note that our regions of interest analysis builds on weight-shuffling to rule out that the number of kernels influenced model accuracies. For each kernel and participant, we computed the predicted kernel activities based on the vHD over time. To give an example, for a given vHD of 30° a kernel centered on 30° was assigned high activity, a kernel centered on 40° slightly lower activity and a kernel centered on 200° was assigned very low activity. By doing this for all time points and kernels, we built regressors representing the predicted activity given a directional tuning and the specific vHD over time. Since vHD was sampled at higher temporal resolution than the imaging data, we then computed the within-TR activity of each kernel as the median activity across all time points within the TR. Finally, the resulting regressors were scaled from 0 to 1 and convolved with the hemodynamic response function (HRF) as implemented in SPM12 using default settings (kernel length: 32s, time to peak: 6s). Each regressor represented a predicted activity profile over time as modeled by the respective kernel. To follow the example above, the activity of a voxel encoding the direction 30° should be more similar to the activity predicted by the kernel centered on 30°, than to the one predicted by the kernel centered on 200°.

Model training

We estimated voxel-wise weights for all vHD-kernels in the corresponding basis set using l2-regularized (ridge) regression. To improve the directionality of this model, we added a covariate modeling movement independent of direction whose weight was discarded. Because vHD is not independent at two successive time points, the resulting design was multicollinear. Ridge regression avoids potential biases in the resulting weights by penalizing high coefficients, which could otherwise affect model accuracies. Since the regularization parameter () cannot be known a priori, our model training builds on leave-one-out cross-validation to find the optimal and with it the optimal model weights. As training set, we used the first two and the last two scanning runs, leaving the third run as the final and independent test set. A scan run typically took around 10 minutes. Since run 3 was acquired in the middle of the experiment, our results are invariant to the duration of the scanning session. If a participant did not complete all five runs, we always used the third valid run as a test set and all others as training set. Within the training set, we used all runs except one to fit voxel-wise weights for ten different regularization parameters log-spaced between 1 to 10,000,000, each time testing how well these weights predicted the activity in the left-out

validation run. To assess prediction performance, we used Pearson correlation between the real time course of a voxel in the validation set and the time course predicted by vHD weighted by the estimated model weights. We cross-validated this prediction such that each run within the training set served as validation set once. The regularization parameter that led to the best prediction performance on average was then determined for each voxel. If no clear best-performing could be determined (i.e. Pearson's R negatively approximated zero with increasing regularization), the respective voxel was excluded. We then averaged across voxels within each participant and used it to estimate the final model weights using the full training set. These model weights serve as basis for all further model tests described below.

Model test

All model tests were performed on the held-out and independent test set (one scan run of around 10 minutes in the middle of the experiment). First, we predicted voxel-wise activity (similar to the model training) in a univariate 'forward-model' approach. For each voxel, we generated a predicted time course by weighting the design matrix of the test run after HRF-convolution by the model weights obtained for this voxel during model training. The resulting predicted time course was then compared to the observed time course using Pearson correlation. Please note that all steps described below were repeated for multiple vHD-basis sets differing in the number and width of the corresponding vHD kernels.

We mapped directional tuning across the cortex using the statistical non-parametric mapping (SnPM) toolbox ([319], http://warwick.ac.uk/snpm). We performed a permutation-based one-sample t-test of model performance (Pearson correlations) against zero (k = 10000 shuffles, input image and variance smoothing: 7.2mm). To reduce computational costs the preprocessed data were downsampled from 0.9 mm to 1.8 mm isotropic for this step. This was repeated for each vHD basis set. We then used the SnPM toolbox to threshold the resulting pseudo-T-maps at a FDR-corrected $p<0.05$. For each voxel we then selected the across-participant median tuning width of the vHD basis set that maximized the pseudo-T. Figure 3 depicts these results for all participants. For visualization we plot the results overlaid on the group-average T1 template at T1-resolution obtained via nearest-neighbor interpolation. We repeated this analysis split into high-memory-error and low-memory error participants (i.e. median-split of memory error, Fig.6.4). Note that each participant group comprised n = 10 participants, resulting in 1024 possible shuffles and a minimal possible p = 0.000977. Because this precludes FDR-correction we use uncorrected p = 0.001 to visualize these sub-group effects.

In addition to the voxel-wise group analysis we also conducted a regions of interest analysis for areas involved in scene processing and navigation such as the early visual cortex, retrosplenial cortex, parahippocampal cortex, the posteromedial entorhinal cortex and the hippocampus (Fig.4.5A). We again performed model training and test, this time focusing specifically on voxels in our ROIs. This greatly reduced the number of voxels and hence computational cost. To avoid potential influences of the number of kernels on these ROI results, we instead performed voxel-wise bootstrapping to convert the Pearson correlations into z-scores. The necessary null distribution of each voxel was obtained by shuffling the training weights 500 times, each time computing Pearson's R between the predicted and the observed time course of the test set. All shuffles were unique. Since not all voxels in our (probabilistic) ROIs were expected to carry vHD-information and to increase robustness of our effects, we performed voxel selection within each ROI, limiting the model test to voxels with high predictability in the model training (top 25% highest prediction accuracy in the training). If a voxel was not directionally tuned in the training, we did not expect it to be directionally tuned in the test. Finally, the z-scores of the remaining voxels were averaged within each ROI. Again, we iterated this analysis for all basis sets (Fig.4.5B), yielding the tuning strength (model performance) and the tuning width (the width of the vHD-kernels in the basis set) of each ROI. Statistical inference was performed for each ROI and participant group (high- and low-memory-error participants) using permutation-based one-sample t-tests on group level (all possible 1024 shuffles, n = 2 x 10) as implemented in the mult4comppermt1 function distributed by Mathworks. To test for differences across participant groups, we used permutation-based unpaired two-sample two-tailed t-tests (k = 10000) as implemented in statcond distributed via the EEGLab Matlab toolbox.

In addition to the forward model test, we also inverted the encoding model to reconstruct, or 'decode', vHD multivarietly from the population of voxels within each ROI (Supplementary Fig.4.13). We multiplied the Moore-Penrose pseudoinverse of all voxel-wise weights in an ROI (m voxels x k weights)-1 with the multivoxel-pattern at each image acquisition (m voxels) to obtain the estimated vHD-kernel activities at each for volume (k weights). By doing this for all volumes, we reconstructed the vHD-kernel activities for the entire test set of each participant. To assess reconstruction performance, we used 2D correlation between the reconstructed kernel activities (k weights x n TR's) and the design matrix of the test run (also k weights x n TR's). Since again the number of kernels and hence reconstructed weights differed across iterations and basis sets, we used weight-

shuffling to convert the resulting correlations into z-scores. As in the forward model, we shuffled the model weights (500 random unique shuffles), each time going through the full vHD-reconstruction procedure. As expected, the results obtained by inverting the encoding model (Supplementary Fig.4.13) resemble the ones obtained by the voxel-wise forward-model procedure (Fig.4.5).

Temporal signal-to-noise ratio (tSNR) and model regularization do not explain the results

The estimated optimal regularization parameter (Supplementary Fig.4.9A) depended on the basis set (Supplementary Fig.4.9B) (rmANOVA results: $F(7, 126) = 11.80$, $p = 1.7 \times 10^{-11}$) as expected, but not on participant group ($F(1, 18) = 0.0037$, $p = 0.952$) and there was no interaction between the two ($F(7,126) = 0.37$, $p = 0.920$). There were differences in tSNR across ROIs (rmANOVA results: $F(5, 90) = 258.29$, $p = 9.1 \times 10^{-52}$), but not across participant groups ($F(1, 18) = 0.082$, $p = 0.777$) and there was no interaction between the two ($F(5, 90) = 1.34$, $p = 0.255$) (Supplementary Fig.4.9C). Neither model performance (Spearman correlation: $rho = 0.076$, $p = 0.409$), nor tuning width ($rho = -0.059$, $p = 0.525$) correlated with tSNR Supplementary Fig.4.9D.

4.6 Supplementary material

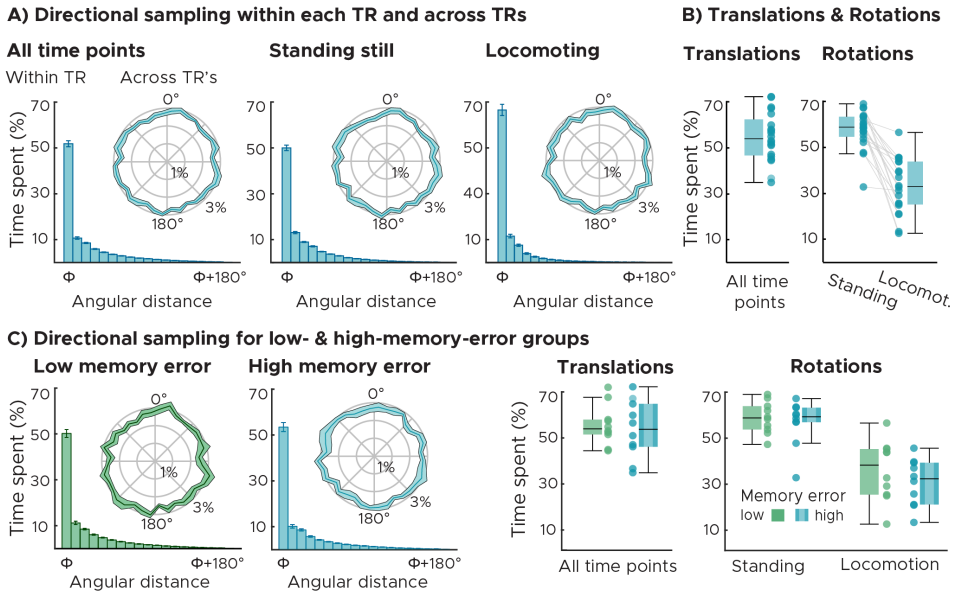


Figure 4.7: Navigation behavior (n = 20). A) Directional sampling across all time points (left), stationary periods only (middle) and locomotion periods only (right). Directional sampling expresses the time spent facing into each direction in percent of total experiment time. The bar plots depict the across-participant mean and SEM within each fMRI acquisition (TR). Data were pooled according to the angular distance to the predominant direction (, first bar). Within each TR, participants' faced into one predominant direction for around 50% of the total experimental time. The time spent facing into the predominant direction differed between stationary and locomotion periods (two-tailed permutation-based paired ttest results: t(19) = 6.12, p = 0.0001, k = 10000). Polar plots depict the directional sampling across the entire experiment (line: mean, shaded area: SEM). Across the experiment, sampling was matched across directions, showing that there were no asymmetric spatial cues that biased navigation behavior (rmANOVA results for all time points: F(35, 665) = 0.77 , p = 0.834, while standing still: F(35, 665) = 0.87, p = 0.681 and while locomoting: F(35, 665) = 0.79, p = 0.806). B) Translations and rotations. Participants locomoted around 50% of the total experimental time (left panel). While locomoting, participants spent less time rotating than while standing still (right panel) (two-tailed permutation-based paired ttest results: t(38) = 7.95, p = 0.0001, k = 10000). We plot single-participant data and group-level whisker-and-box plots (center, median; box, 25th to 75th percentiles; whiskers, 1.5×interquartile range). C) Directional sampling across participant groups. There were no differences in directional sampling in neither of the two participant groups across TRs (rmANOVA results: low-memory-error group: F(35, 315) = 0.87, p = 0.675, high-memory-error group: F(35, 315) = 0.55, p = 0.984). The two participant groups spent equal amount of time locomoting (t(18) = 0.41, p = 0.690), and rotating (while locomoting: t(18) = 0.55, p = 0.625 and while standing still: t(18) = 1.00, p = 0.338) and facing towards one predominant direction within each TR (t(18) = -1.26, p = 0.224, k = 10000).

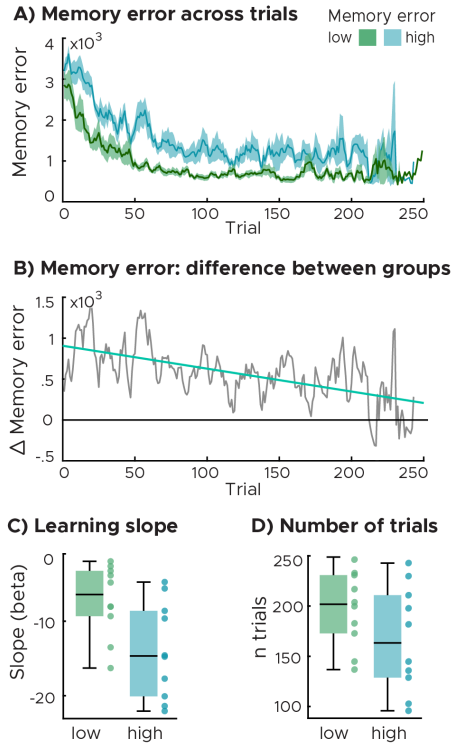


Figure 4.8: Memory performance during the object-location memory task. A) Memory error across trials for the two participant groups differing in the median memory error. The memory error is the Euclidean distance between true and remembered object location expressed in virtual vertices. Data were smoothed using a running average kernel of 5 trials. B) Differences in memory error between groups. Zero constitutes no difference. Both groups converge on the same level of performance in the course of the experiment. C) Learning slopes. Regression slopes of a linear line fitted to the raw memory error scores across trials of each participant. Participants with higher median memory error showed steeper slopes (two-tailed permutation-based unpaired ttest results: $t(18) = 2.65$, $p = 0.019$), indicating that the memory error improved faster than in the low-memory-error group. D) Number of trials. Participants with low memory error performed slightly more trials than the ones with high memory errors, the difference between groups was however small ($t(18) = 1.46$, $p = 0.137$). C,D) We plot single-participant data and group-level whisker-and-box plots (center, median; box, 25th to 75th percentiles; whiskers, 1.5×interquartile range).

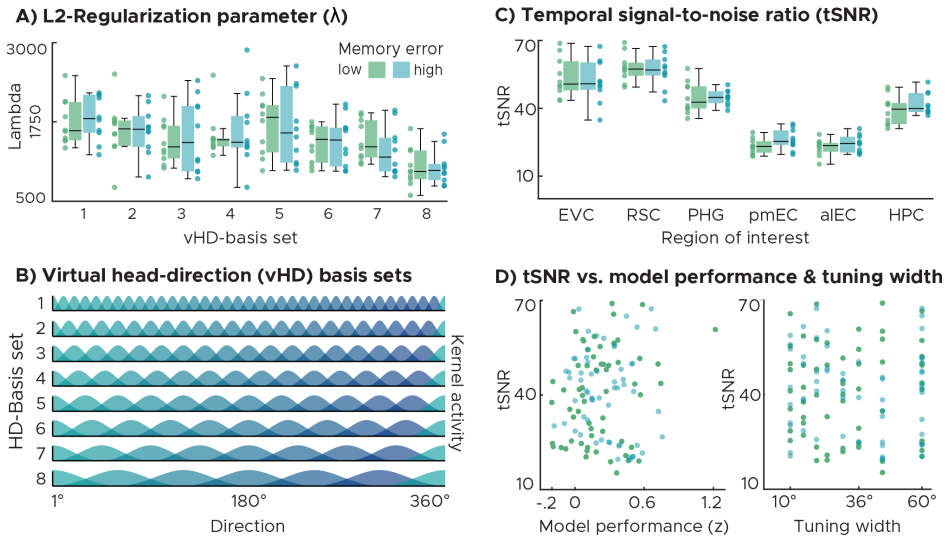


Figure 4.9: Model parameters and data quality. A) L2-regularization parameters () for all basis sets and the two participant groups (differing in memory error). Each dot represents of a single participant and basis set averaged across all ROIs (Fig.6.4). Group-level whisker-boxplots of the same data were added (center, median; box, 25th to 75th percentiles; whiskers, 1.5×interquartile range). Lambda depended on the basis set (rmANOVA results: F(7, 126) = 11.80, p = 1.7 x 10-11) as expected, but not on participant group (F(1, 18) = 0.0037, p = 0.952). There was no interaction between the two (F(7,126) = 0.37, p = 0.920). B) Histogram depiction of all basis sets used. Each Gaussian represents one directional kernel covering the full 360° with 1° resolution. Basis sets differed in kernel width and spacing. Resulting regressors were scaled between 0 and 1. C) Temporal signal-to-noise ratio (tSNR) across ROIs and participant groups. We plot group-level whisker-boxplots and single participant data of the average tSNR of each region. There were differences in tSNR across ROIs (rmANOVA results: F(5, 90) = 258.29, p = 9.1 x 10-52), but not across participant groups (F(1, 18) = 0.082, p = 0.777) and there was no interaction between the two (F(5, 90) = 1.34, p = 0.255). D) Scatter plots for tSNR over model performance (left) and estimated tuning width (right). Dots represent the average tSNR of each region and participant. Neither model performance (Spearman correlation: rho = 0.076, p = 0.409), nor tuning width (rho = -0.059, p = 0.525) correlated with tSNR.

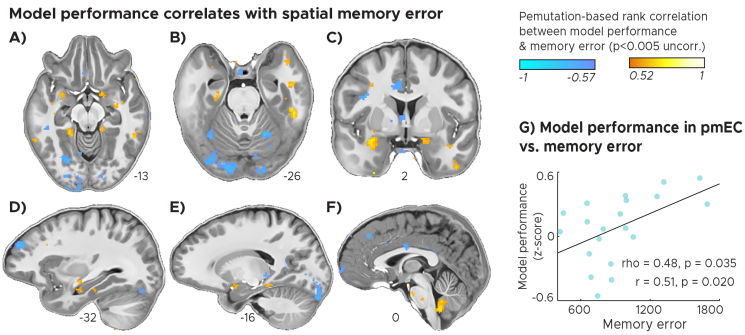


Figure 4.10: Across-participant permutation-based rank-correlation between the model performance (see Fig.6.3) and across trial median memory error (Fig.6.1). A-F) We plot Spearman's correlation coefficients (Rho) thresholded at p<0.005 uncorrected for visualization and overlaid on the group-average T1 template at T1 resolution. We observed a positive correlation between model performance and memory error in the parahippocampal gyrus (A,B,D,E) and anterior medial temporal lobe (A,B,C,D,E) including posterior entorhinal cortex (B) and anterior hippocampus (D), as well as in the ventral cerebellum (F). Negative correlations could be observed in the ventral occipital lobe (A,B,D,E), frontal lobe (D,F), the medial cingulate cortex (C) and the posterior thalamus (E). G) Scatter plot of posteromedial entorhinal cortex (pmEC) model performance (Fig.4.5) over across-trial median memory error (Fig.6.1). Least-square line as well as permutation-based rank and linear-correlation coefficients were added. There is a correlation between model performance and memory error in pmEC.

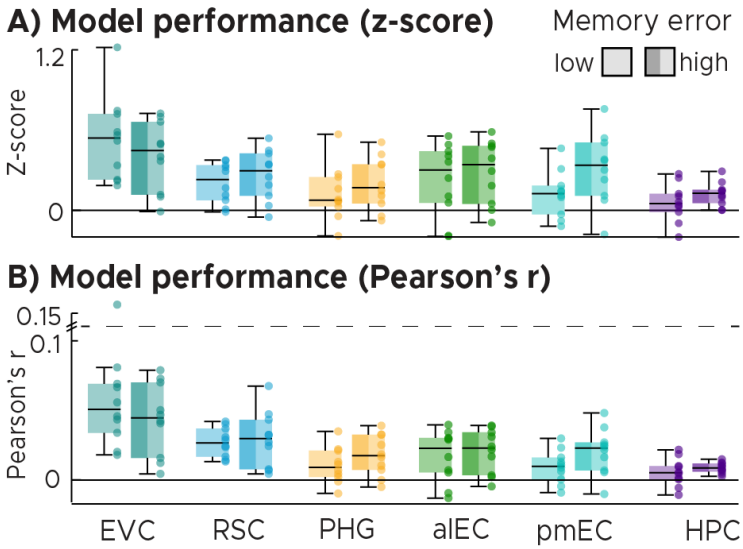


Figure 4.11: Optimal model performance expressed in Z-scores (A) and Pearson's R (B) selected for each individual participant instead of group level (Fig.4.5B,C). We plot single-participant data overlaid on group-level whisker-and-box plots (center, median; box, 25th to 75th percentiles; whiskers, 1.5×interquartile range) for low and high-memory-error groups. B) Y-axis was cut for visualization (dashed line).

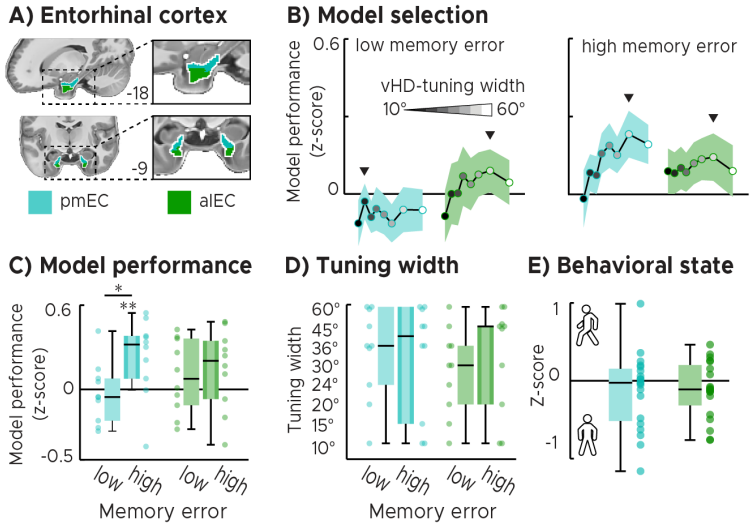


Figure 4.12: Comparison between entorhinal cortex subdivisions. A) Regions of interests: posteromedial (pmEC) vs. anterolateral entorhinal cortex (alEC). B) Model selection: We plot the model performance (Z-score) for all basis sets and the two participant groups. The black line and the shaded area represent the mean and SEM across participants. Each dot represents the group-average model performance for one basis set, with darker colors representing narrow kernels and lighter colors representing wider kernels. The following kernel widths were tested: 10°, 15°, 20°, 24°, 30°, 36°, 45°, 60°. For each ROI on group-level, we selected the best performing basis set as the optimal model to be tested. Also see Fig.4.5B. C) Optimal model performance selected in B for the two (high- and low-memory-error) participant groups. We plot single participant data and group-level whisker-boxplots (center, median; box, 25th to 75th percentiles; whiskers, 1.5×interquartile range). We observed directional tuning in pmEC in participants with high memory error and a difference between groups (low memory error: $t(9) = -0.42$, $p = 0.661$; high memory error: $t(9) = 2.59$, $p = 0.020$, $pFDR = 0.040$, $t(19) = 2.32$, $p = 0.036$). In alEC, neither directional tuning nor the difference between groups could be observed (Low : $t(9) = 1.09$, $p = 0.146$; high: $t(9) = 1.58$, $p = 0.075$; contrast: $t(19) = 0.44$, $p = 0.655$). Also see Fig.4.5C. D) Optimal tuning width. Similar to B,C, we plot the tuning width that led to the highest memory performance selected on individual participant level (see Fig.4.5D). E) Behavioral state analysis. Model performance during locomotion and stationary periods. Positive values indicate that voxel time courses in the ROI could be better predicted when participants locomoted. We plot the difference in model performance as single participant data and group-level whisker-boxplots (center, median; box, 25th to 75th percentiles; whiskers, 1.5×interquartile range). There is no effect of locomotion per se on the tuning in pmEC (and alEC which was not directionally tuned in our task).

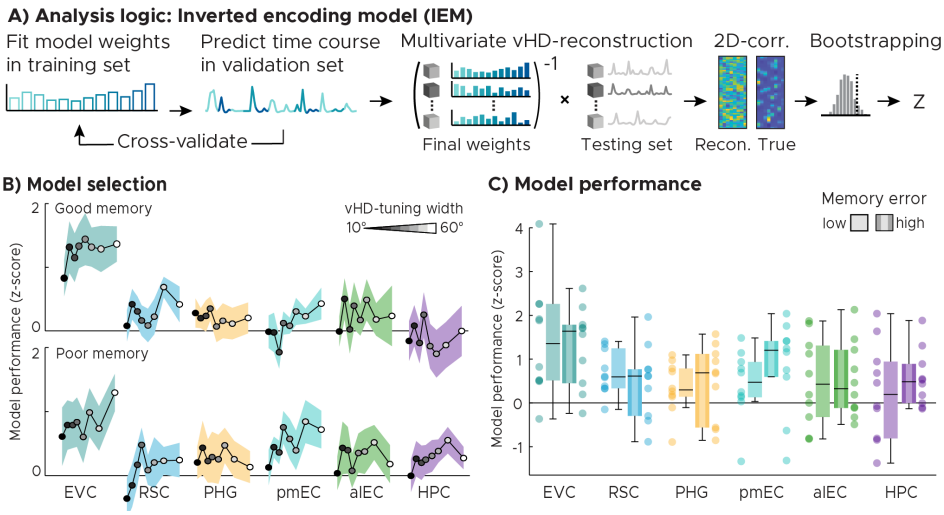


Figure 4.13: Multivariate inverted encoding model (IEM). A) Analysis logic. We estimated voxel-wise weights following the same training procedure used for the forward model (see methods). We then multiplied the Moore-Penrose pseudoinverse of all voxel-wise weights in an ROI (m voxels x k weights)-1 with the multivoxel-pattern (m voxels) at each image acquisition (TR) to obtain the estimated vHD-kernel activities at each TR (k weights). We then used 2D-correlation between the reconstructed kernel activities (k weights x n TR's) and the design matrix of the test run (also k weights x n TR's) to compare reconstructed and observed vHD. Finally, the resulting correlation coefficient was converted to a z-score using the bootstrapped null distribution of each ROI (k = 500 unique shuffles). B) Model selection: We plot the IEM-model performance (Z-score) for all basis sets. The black line and the shaded area represent the mean and SEM across participants. Each dot represents the group-average model performance for one basis set, with darker colors representing narrow kernels and lighter colors representing wider kernels. The following kernel widths were tested: 10°, 15°, 20°, 24°, 30°, 36°, 45°, 60°. For each ROI on group-level, we selected the best performing basis set as the optimal model to be tested. C) Optimal model performance selected in B for the two (high- and low-memory-error) participant groups. We plot single participant data and group-level whisker-boxplots (center, median; box, 25th to 75th percentiles; whiskers, 1.5×interquartile range).

Chapter 5

General discussion

How does the brain derive stable representations of the world from sensory inputs and how do they shape the way we perceive and interact with the world? In my doctoral work, I addressed these questions by examining the cognitive processes underlying a stable visual perception and spatial mapping in the human brain. I approached this question from multiple angles by focusing on different computational levels along the cortical hierarchy. I used fMRI to monitor human brain activity during tightly controlled visual-tracking tasks and naturalistic navigation in virtual reality, and combined it with predictive computational modeling to study the neural systems underlying human behavior. I focused on viewing and vision as a model system for exploration behavior in general.

For cognitive maps to be formed, visual information needs to undergo a series of transformations that ultimately construct non-sensory and world-centered representations of the environment (Fig.5.1). First, retinotopic representations of space are being processed and maintained in visual cortices. Second, these retinotopic representations are being integrated with self-motion and body position signals to reconstruct non-retinotopic representations in mid-level and higher-level visual cortices. This includes the integration of efference copies with visual input, retinotopic updating, modulation of visual activity by gaze-related gain fields, proprioceptive compensation of movements and transsaccadic memory (see Chapter 1). Third, once visual receptive fields are stablized and a stable perception is ensured, associations between the location of the receptive fields and environmental features such as landmarks can be learned. This engages scene processing

regions such as the retrosplenial and parahippocampal cortex and likely builds on the conjunctive representation of direction and place. This anchors receptive fields to these landmarks in a world-centered frame of reference and allows to track them over time. Finally, these world-centered representations provide input to attractor networks in the medial temporal lobe, which mediate a non-sensory, unified and continuous representation of space. These are updated by path integration mechanisms (e.g. vestibular inputs), but require visual anchoring to avoid error accumulation. This anchoring is achieved by visual inputs from scene processing regions, which reference the intrinsic medial temporal activity dynamics to the external world.

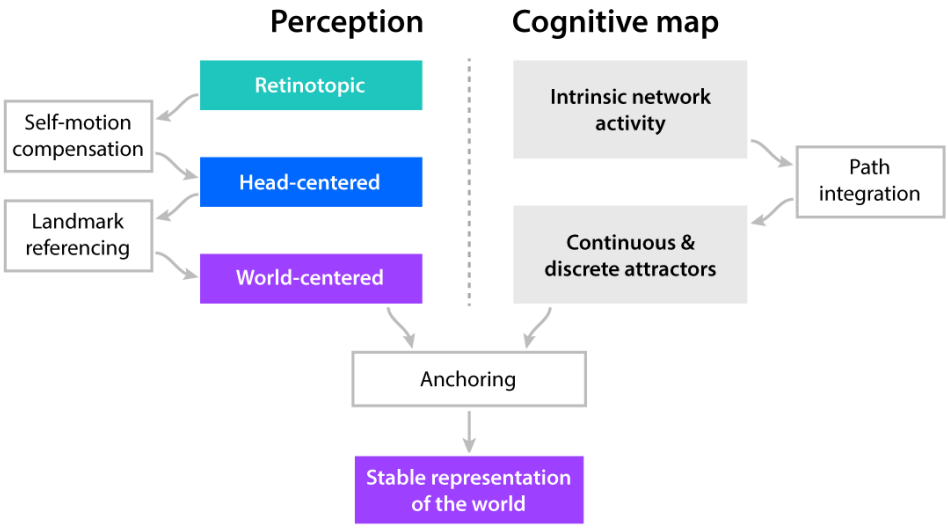


Figure 5.1: Scheme of how perception and the cognitive map combine. The brain derives a stable representation of the world from visual inputs using two interconnected mechanisms. Retinotopic input is being acquired by our senses and processed in visual areas. Non-retinotopic (e.g. head-centered) receptive fields are being reconstructed from retinotopic ones in parietal areas by compensating for self-motion and maintaining memory traces of the visual field. Once receptive fields are stable, they are associated with visual features such as landmarks. This generates world-centered (possibly vectorial) visual representations of the environment in scene processing regions. In the medial temporal lobe, intrinsic network activity gives rise to attractor dynamics representing locations and directions. Path integration mechanisms update the activity state in the attractor network. With every update however, they accumulate updating errors. Ultimately, the world-centered representations in scene processing regions are being used to anchor the non-sensory medial temporal attractor dynamics to the external world (See Fig.1.3 for how this could be achieved).

The introduction mainly focused on how cognitive maps are formed. This thesis however also examined the behavioral and perceptual consequences of this process. We discussed many of these implications in our a review article [127], which is added here as 'Extended discussion' as Chapter 6.

Review paper 4 [127] reviews the growing body of literature that implicates the mammalian hippocampal formation in mapping visual space using a non-sensory representational format. We propose that medial temporal spatial representations provide the optimal solution to multiple computational challenges that are shared between navigation and vision. They represent visual space in a world-centered frame of reference, a mnemonic representation of the visual field that might modulate activity in upstream sensory and motor areas. By doing so, they support critical cognitive functions such as behavioral planning and provide a relational metric for memory formation and imagery. Finally, hippocampal remapping might provide the basis for contextual memory retrieval to aid visual search and recognition memory across different contexts. We propose that vision and navigation are two expressions of one common medial temporal mechanism supporting exploration in general. Vision is an ideal domain to study lower and higher cognitive functions simultaneously in the human brain.

Because these (and more) points are discussed in detail in the 'Extended discussion' Chapter, I will instead elaborate on points that were beyond the scope of this review in the following. I will begin by summarizing the empirical work presented here, then discuss why cognitive mapping and world-centered coordinates are useful, and finally present selected examples that demonstrate how cognitive maps shape perceptual processing and behavior.

5.1 Synopsis

Paper 1: Real-motion signals in human early visual cortex

Consider following scenario. You are standing at the Trondheim train station, watching a train that is about to leave. Whenever you move your eyes, the visual image of the train will change its location on your retina. This in turn induces visual motion, which will be signaled to the brain. However, if you do not move your eyes, but the train moves instead, this will also result in visual motion. In both cases, there will be movements on the retina (retinal motion), yet the environmental source of that motion is drastically different. How can the brain tell

whether retinal motion was self-induced or not?

Electrophysiological studies suggested that neurons in the monkey visual system integrate retinal motion with the motor commands that tell the eyes to move. The computations these neurons perform seem simple: adding the anticipated eye velocity to the incoming retinal motion signal reconstructs how fast the train moved in the outside world. If the eyes move faster than the train, these neurons perform a subtraction of the two factors. The neurons performing this computation are called 'real-motion neurons' [47]. By teasing apart 'real' from self-induced motion, these neurons are thought to stabilize visual perception in space. In monkeys, real-motion neurons were found in several brain areas including V1 [47], V2 [250], V3A [62] and V6 [366].

Complementary work in human fMRI showed that similar processes might support human visual perception [61, 65, 258]. Using planar-motion dot stimuli during pursuit eye movements, they showed that human areas V3A ad V6 signal real-motion velocity independent from eye movement speed. These earlier reports however had one shortcoming: the motion conditions used differed not only in eye movements and real-motion, but also in retinal-motion. It therefore remained unclear how much of the observed activity could be attributed to retinal motion, and how much was due to real-motion as such.

We aimed to answer this question and to provide a detailed overview on real-motion encoding across the human visual system. To do so, we again used planar-dot motion stimuli and a visual tracking task similar to earlier work [61] while monitoring brain activity and eye movements using fMRI and eye tracking. Going beyond earlier work, we here designed motion conditions that differed in eye movement speed and real-motion, but importantly not in retinal motion. Using independent stimuli and data, we localized 16 visual areas and revealed that a whole network of brain regions engaged in this process. In addition to confirming earlier reports about real-motion selectivity in V3A and V6, our carefully controlled stimulation procedure revealed that even the earliest visual cortices in the human brain encoded visual motion in a non-retinotopic, potentially world-centered frame of reference. This parallels reports of real-motion neurons in monkey V1 [47].

This work illustrates how very simple computations enable to reconstruct a non-retinotopic frame of reference [43, 44]. It also suggests that extra-retinal signals modulate activity even in the earliest stages of the human visual system to stabilize our perception during self-motion.

Paper 2: Hexadirectional coding of visual space in human entorhinal cortex

Seminal work in rodents showed that the brain evolved a world-centered spatial mapping system mediated by the hippocampal and entorhinal circuits [123, 124]. These spatially tuned cells typically represent position, direction and distance relative to environmental boundaries or landmarks. Among this 'zoo' of cells are entorhinal grid cells, which represent a map of navigable space [151]. They fire as a function of self-location and tessellate the environment with hexagonally arranged firing fields. Strikingly, akin to rodent grid cells during navigation, neurons in the monkey entorhinal cortex represent a grid-like map of visual space [225].

Studies of entorhinal activity during virtual reality suggested that similar grid-cell mechanisms support navigation in humans [158, 173]. As described in Chapter 1, the respective fMRI analysis probes 'hexadirectional signals', a six-fold rotationally symmetric modulation of entorhinal activity as a function of running direction. Critically, whether the human entorhinal cortex represents a grid-like map of visual space similar to its monkey homologue remained unknown.

The second study addressed this question and examined hexadirectional signals during a visual tracking and spatial memory task. Similar to project 1, participants fixated at a fixation dot that moved within a virtual arena seen from bird's eye view. While following the dot, participants memorized object locations within the arena. We used eye tracking to ensure tight control of our participants' viewing behavior and analyzed entorhinal activity as a function of eye movement direction. To do so, I developed a cross-validated voxel-wise version of the hexadirectional analysis proposed earlier [158]. If grid cells in the human EC represented visual space, we expected to see a six-fold rotational modulation of activity as a function of eye movement direction.

Indeed, we found that EC activity exhibited this predicted six-fold rotational symmetry and was hence hexadirectionally modulated in our task. We neither observed such modulation in control regions such as motor cortex, visual cortex and frontal lobe, nor did the EC activity exhibit any other rotational symmetry but six-fold. A control condition further excluded the possibility that optic flow could have influenced our results and ensured that activity in the EC was not generally modulated by eye movements or visual motion per se.

These results provided the first evidence that the human entorhinal cortex represents a non-sensory map of visual space using a grid-like code. This suggests that

the spatial mapping systems for navigation and vision are strongly intertwined and that entorhinal coding principles play a broad and general role in cognitive functions beyond navigation. Most beautifully, these results were published together with additional independent evidence that directly supported the same conclusions [171, 218].

Paper 3: Behavior-dependent directional tuning in the human visual-navigation network

To encode and maintain a cognitive map of the environment, the brain needs to continuously monitor our viewing direction as we move [341]. This engages a large network of brain regions, spanning from early visual cortex over scene processing regions to high-level mnemonic areas such as the hippocampal formation [97, 98, 127, 315]. Despite decades of intensive research, the neural underpinnings of this process remain poorly understood in humans, especially in the light of the behavior it supports. This is partly due to methodological constraints when studying human brain activity during active behavior.

Here, we overcame some of these limitations by developing a predictive behavioral modeling framework to study the neural population tuning underlying cognitive mapping in fMRI. We re-analyzed ultrahigh-field 7T-fMRI data of participants freely navigating in a virtual reality (VR) arena and performing an object-location memory task [322]. As a window into cognitive mapping, we examined directional tuning across key processing stages along the visual cortical hierarchy. This included the early visual, retrosplenial, parahippocampal and entorhinal cortices as well as the hippocampus.

We analyzed these data using a newly developed voxel-wise encoding model of the navigation behavior, an approach that was heavily inspired by prior successes of similar techniques in the visual domain [18, 32, 33]. First, we modeled world-centered virtual head direction using basis sets of directional kernels differing in kernel spacing and width. Using a cross-validated training-test procedure, we then assessed how well these models predicted the activity of a voxel in independent data. By iterating through all basis sets, we obtained the directional tuning strength and width of each voxel.

Our results presented multiple conceptual and methodological advances. The model successfully predicted activity in visual, retrosplenial, parahippocampal and entorhinal cortices, demonstrating directional tuning in the human scene

processing and navigation network for the first time during active spatial behavior. Moreover, mapping the tuning width across the brain revealed a narrow-to-broad tuning width topology in the parahippocampal cortex.

Most interestingly however, analyzing the tuning as a function of spatial memory performance showed that tuning strength, width and its cortical topology all depended on how well participants remembered the object locations within the arena. The direction of these effects as well as additional behavioral evidence suggested that these tuning differences reflected how well the environment has been encoded. Finally and following rodent work [335], we analyzed the effect of locomotion on the present results, showing stronger tuning in visual and retrosplenial areas during movement, and stronger tuning in hippocampal and parahippocampal regions while standing still. Multiple control analyzes along with the careful design of the VR-environment ensured that our results could neither be explained by directional sampling biases, nor by model parametrization or data quality.

In sum, these results suggest that our behavioral and mnemonic state influences directional tuning in the scene processing and navigation network during active memory-guided navigation. By doing so, they demonstrate that a high-level cognitive mapping process interacts with network-wide environmental processing and show the power of encoding models to study human brain activity during naturalistic behavior. The following section will speculate about reasons why the brain might have developed cognitive mapping in the first place and why a world-centered reference frame is a useful representational format.

5.2 Why cognitive mapping & world-centered coordinates?

Following the discovery of place cells in the hippocampus [128], O'Keefe and Nadel proposed that place cells are the neural implementation of the cognitive map [129]. By showing that place cells encode locations when rats navigated through a box, they demonstrated that in this context place cells represented a map of the environment in a world-centered frame of reference. Since then, multiple other cell types have been found in the hippocampal-entorhinal circuit that support similar functions as well [123, 124]. Importantly, Tolman conceptualized the cognitive map as a general organizational principle of how information is represented in the brain that is not necessarily limited to navigation and the representation of space [137]. In primates, growing evidence suggests that the same neural systems that represent a non-sensory map of space during navigation might also

represent a map of visual space as well [127, 216, 217].

Why is it important to represent space in a map-like format and in a world-centered reference frame? The critical function of spatial representations in the hippocampal formation might be to organize information in a way that enables flexible expression of behavior and knowledge [137]. World-centered coordinates and cognitive maps might be optimal to support these functions for the following reasons. First, one practical answer is *efficiency*. Representing all spatial relationships exclusively referenced to oneself would require the encoding of an enormous amount of vectorial representations, which needed to be updated perpetually as we move. Second, a closely related argument is *robustness*. Because path integration and spatial updating accumulates errors, this would ultimately lead to a drift in spatial representations, wich renders them useless for the guidance of behavior. By anchoring them to the environment, they encode a robust, stable and reliable map of the environment. Third, world-centered representations allow to pre-compute trajectories and hence support the *planning* of our actions in space. Pre-computing saccade sequences for example would be difficult for a purely retinotopic system, because all vectors will have changed after the execution of the first saccade. This is not the case if the trajectory is pre-computed in world-centered coordinates. Intriguingly, empirical evidence suggests that saccade sequences are indeed pre-computed in movement-invariant, possibly world-centered coordinates [367–369]. Hippocampal-entorhinal codes are optimal to perform these computations, because they likely allow to generate vectorial representations from positional ones when needed [97, 98, 127, 154, 228, 370]. Forth, map-based encoding of information allows greater *flexibility* to infer new relationships between places or events that were not originally experienced together. For example, this means that shortcuts can be derived from prior knowledge without having to experience these shortcuts before. This enables to respond more flexibly to new challenges and changes in the environment.

Importantly, the above-mentioned functions depend on one critical aspect of cognitive maps: they maintain information over time. For some functions like planning saccade trajectories a short-term storage of the spatial scene would suffice, others however require long-term storage of information in the form of episodic *memories* [371–373]. This likely relies on an interplay between the spatially tuned cell types discussed above with others encoding temporal relationships or the presence of objects in environment. Together, these mediate representations of 'where' and 'when' a certain event took place and 'what' exactly happened (recently reviewed by [374]). This allows to encode and later retrieve experiences

and events and perform mental simulation [97, 98, 373, 374].

Notably, this process again builds on the intrinsic attractor network dynamics within the hippocampal formation [375]. Time and space are continuous, but events need to be memorized discretely and independently of each other to avoid confusion during retrieval. The combination of continuous and discrete attractor networks might provide the solution to this challenge. They help to memorize events using orthogonalized population codes (pattern separation) while still providing smooth transitions between states. During retrieval, the same network dynamics might also help to retrieve an event from memory using partial inputs only (pattern completion) [375]. From this perspective, place cells might encode information with high dimensionality in discrete states, with grid cells providing the efficient and low dimensional basis set that encodes the state transition structure [238, 239, 376]. This idea receives growing empirical support in the past years [168, 169, 237, 377].

In sum, representing information in a map-like, non-sensory and world-centered format enables efficient, robust and flexible computations to plan and guide our actions in space. In addition, it provides the foundation for the formation of memories by allowing to encode and retrieve experiences independent of our current location.

5.3 Cortico-hippocampal interactions are bidirectional

The above-discussed computational efficiency, robustness and flexibility of hippocampal-entorhinal coding puts it in a key position to guide our behavior and to shape perceptual processing in upstream areas. In the following, I will briefly discuss a few selected examples that demonstrate that hippocampal-cortical interactions are bidirectional and that hippocampal spatial representations do indeed shape perceptual processing and behavior.

5.3.1 Episodic recall

One important role of the hippocampal-entorhinal attractor networks discussed in Chapter 1 is pattern completion, the reconstruction of complete stored memory representations from incomplete inputs [375, 378]. Importantly, 're-experiencing' past events from memory engages not only the hippocampus, but the same sensory-specific cortices that were active during perception [379]. Consolidation models

propose that memory encoding and retrieval generally depend on hippocampal-cortical interactions and that the long-term storage of memories lies in cortical networks [380]. Consistently, the vividness of recalled stimuli depends on how similar sensory cortex activity is compared to when these stimuli where perceived [381]. Potentially, this reflects a code that is common for perception, imagination and episodic recall for which especially the anterior portion of the hippocampus seems critical [382]. Indeed, human fMRI studies showed that the ability to decode reinstated associative memories from visual cortex correlates with activity in the hippocampus [338]. Likewise, parahippocampal memory reinstatement as measured by pattern similarity between perception and recall correlates with hippocampal activity [383]. While such studies show that pattern reactivations in sensory cortices covary with hippocampal activity, the direction of this link remains unclear in these studies. In rodents however, inactivations of hippocampal CA1 neurons showed a direct top-down relationship between hippocampal activity and reinstatement in sensory cortices [384]. In sum, pattern reinstatement in sensory cortices during recall is likely orcestrated by hippocampal pattern completion mechanisms, which are also a fundamental part of the cognitive map [375]. Furthermore, an important but more implicit memory function is the retrieval of contextual information. This point will be discussed in detail in the 'Extended discussion' in the 'Context selectivity' section.

5.3.2 Predictions

By allowing to form memories and to learn from past experiences the hippocampal formation is a critical component of the 'prediction machine' that arguably forms the most essential cognitive function of all [385]. Learning from past experiences is useful only if one can form expectations about future events and adapt one's behavior accordingly. In fact, this predictive principle has been proposed as a general framework to understand brain function, in which the brain is considered to be a hierarchical and recursive generative network that constantly tries to match top-down expectations to sensory signals to minimize prediction errors [385]. This relies on a constant bidirectional cascade of both bottom-up inputs and top-down feedback [299] that modulates sensory processing on a large scale (recently reviewed by [386]). While not all predictions are necessarily hippocampus-dependent, a growing number of studies shows that the hippocampal formation is indeed involved in predictive processing. In rodents for example, evidence for this has been observed in the form of phase precession [387]. Here,

place cells fire in progressively earlier theta phases of the local field potential as the animal approaches the neuron's receptive field. Phase precession can be interpreted as the prediction of the sequence of upcoming positions or events [388]. Similar coding has been observed in grid cells [389], and strikingly also in V1 neurons [336], which are modulated by position [337]. These spatial representations in V1 are correlated with the ones in the hippocampus even in the absence of visual cues [336], suggesting that predictive positional signals might permeate from the hippocampus all the way to V1. In humans, fMRI studies using sequence-learning paradigms showed that the medial temporal lobe is involved in learning statistical regularities [390] and that the ability to decode stimuli from V1 activity correlates with the ability to decode the stimulus sequence from the hippocampus [339]. Similarly, after participants learned to predict visual stimuli based on auditory cues, the shape of the predicted stimulus can be reconstructed from hippocampal activity [340]. While the exact role of the hippocampus in this process might change over time, along with its background connectivity to sensory regions [391], such reports show that it does contribute actively to forming predictions.

Notably, the mnemonic, predictive and spatial functions of the entorhinal cortex were recently incorporated into a computational model of recognition memory. Here, putative visual grid cells encode the location of visual features in the scene and guide the oculomotor systems to test predictions about stimulus identity [228]. Importantly, my doctoral work presented in this thesis provided the first evidence for visual grid cell activity in the human entorhinal cortex [170]. Together, such findings and models suggest that the hippocampal formation likely contributes greatly to the generation of predictions, which then also modulate processing in sensory regions.

5.3.3 Attention

Processing in sensory and motor regions is strongly modulated by attention in various ways [392]. One possible influence is the sharpening of the representations of attended stimuli, which are then also transmitted more strongly to downstream regions such as the hippocampal formation. Attention is controlled by a fronto-parietal brain network in which neural processing is dominated by a theta rhythm [393, 394], which can also be observed on a perceptual level [395]. This has led to the proposal that 'the attentional spotlight blinks' [396, 397] to organize perceptual and motor functions in time [397, 398]. Interestingly, the hippocampal formation is driven by similar theta rhythmicity [370, 399–401], which is phase-

reset at the time of gaze fixations [402, 403]. This phase reset has been proposed to provide the time window for new sensory information to enter the hippocampus [404]. While it remains unclear whether these processes are related, one can speculate about potential interactions. As discussed above, attractor dynamics in the hippocampal-entorhinal circuit tend to drift and need to be anchored to the external world using visual information. Potentially, the hippocampal formation strengthens the recruitment of visual inputs via an attentional mechanism for example during reorientation. This could also explain why non-retinotopic representations are modulated by attention [38,70] and need time to build up [75]. They might be recruited by attention and forwarded to downstream regions such as the hippocampal formation whenever needed. This would predict that hippocampal activity is modulated by attention specifically when the required information is spatial, which is indeed the case [405, 406]. The hippocampus contributes to attention to spatial layout and relational information, but not to feature based attention [407, 408]. Consistently, visual grid cell representations in the entorhinal cortex reflects the locus of attention independent of gaze position [227]. While the here proposed relationships remain speculative, these examples demonstrate a close link between hippocampal-entorhinal processing and attention and point towards the possibility that this link is bidirectional [406].

5.3.4 Oculomotor control

Controlling our viewing behavior is related to all three points mentioned above. Eye movements support memory retrieval [409] and where we look and attend can be an expression of our predictions about the environment [410]. While oculomotor behavior and memory processing are known to depend on each other [314, 411], the direction of this relationship is often less clear. Does our gaze behavior determine subsequent memory encoding or is the former a direct expression of the latter? The extended discussion section in Chapter 6 will elaborate on related points in detail. Here, I will instead focus on a phenomenon called 'mnemonic gaze reinstatement', which shows a clear top-down relationship between memory and oculomotor output (recently reviewed by [409]). Tracking the eyes during picture viewing and subsequent recall showed that gaze movements during encoding are partially reinstated during retrieval and that preventing these movements leads to less accurate memory [412, 413]. This suggests a direct functional contribution of the oculomotor system in memory retrieval. Gaze reinstatement depends on hippocampal function [414], which is also in line with the observa-

tion that hippocampal activity predicts how indicative eye movements are of relational memory [415]. Spatial representations in the medial temporal lobe might be critical in this process as an interface between the mnemonic and sensorymotor domains [127, 217, 228]. We proposed that grid cells compute vectors between visual locations similar to their proposed role in navigation [127]. These vectors could guide our gaze movements to currently relevant aspects in the visual scene, which might hold true also during memory retrieval.

5.3.5 Medial temporal lesions: a counter-argument

The work presented in this thesis proposes that the human hippocampal formation mediates a non-sensory representation of space that shapes network-wide perceptual processing and guides viewing behavior. Is there empirical evidence that speaks against this idea? One counterargument could be the lack of strong visual and behavioral deficits following hippocampal damage [329]. Lesions do impair performance in recognition memory [416,417] and working-memory tasks [418], the mechanistic origin of these deficits however is often unclear. A big challenge for empirical tests is differentiating perceptual and behavioral deficits from the ones explained by general memory deficits in such tasks. An optimal experiment would hence test all of these factors independent from the others. This however remains challenging. Importantly, the fact that perception remains unaffected by hippocampal lesions is not at odds with the present proposal. The hippocampal formation is not a perceptual area per se (see Box 1 in Chapter 6). Damaging it would impair the non-sensory representations mediated by the hippocampal formation, but not all world-centered representations in the brain in general. Many here discussed perceptual and behavioral functions benefit from cognitive mapping, but could also be achieved with more effort by other cortical systems and even in retinotopic or head-centered reference frames. One example for such a systems-level re-weighting has been observed in participants with reduced grid-cell like coding in the entorhinal cortex, who tend to show stronger hippocampal activity [164].

Which behavioral or perceptual deficits would one expect following hippocampal-entorhinal damage? It might lead to impairments in transsaccadic integration [79], because the stable and mnemonic hippocampal representation of the visual field is missing. This should lead to a stronger accumulation of retinotopic updating errors in participants with hippocampal damage even though visual constancy can still be achieved. Interestingly, such patients are impaired in attending to spatial

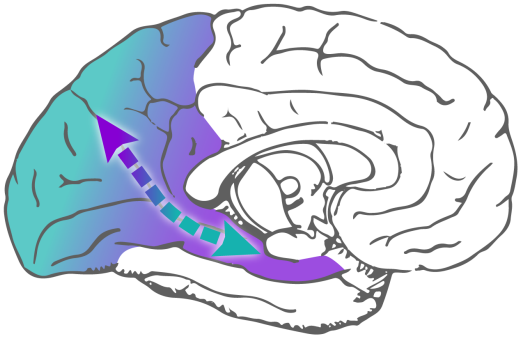
relations (e.g. the spatial layout of the visual stimulus) [408], but are still capable of learning statistical regularities of visual stimuli and to control feature-based attention [407]. Consistently, the expected impairments in transsaccadic integration should be stronger for spatial than for feature information. Future studies could test this proposal empirically.

5.3.6 Concluding remarks

Collectively, the work presented in my doctoral thesis bridges across multiple long-standing lines of research. It ranges from the fields of vision science to navigation and memory research and aims to provide a holistic perspective of the processes involved. The empirical work presented here improves our understanding of spatial perception and cognitive mapping in the human brain and highlights the domain-general nature of our brain's spatial mapping system. It is complemented by a theoretical framework proposing that our visual and navigational mapping systems are strongly intertwined and that vision is an optimal domain to study human cognition on all levels. By demonstrating that the computations along the cortical hierarchy are interdependent, this work also emphasizes the importance of studying human cognition with a broad network-level perspective and in the light of the behavior it supports.

Chapter 6

Extended discussion: How the brain's navigation system shapes our visual experience



This chapter has been published as:
How the Brain's Navigation System Shapes Our Visual Experience.
Nau, M. & Julian J.B., Doeller, C.F. (2018).
Trends in Cognitive Sciences 22, 810-825.
https://doi.org/10.1016/j.tics.2018.06.008

6.1 Abstract

We explore the environment not only by navigating, but also by viewing our surroundings with our eyes. Here we review growing evidence that the mammalian hippocampal formation, extensively studied in the context of navigation and memory, mediates a representation of visual space that is stably anchored to the external world. This visual representation puts the hippocampal formation in a central position to guide viewing behavior and to modulate visual processing beyond the medial temporal lobe (MTL). We suggest that vision and navigation share several key computational challenges that are solved by overlapping and potentially common neural systems, making vision an optimal domain to explore whether and how the MTL supports cognitive operations beyond navigation.

6.2 All eyes on the hippocampal formation

Navigation and vision are two fundamental strategies used to explore the world, and how we use one often directly affects how we use the other [419, 420]. Like navigation—defined here as physical body-based movement through the environment—, eye movements make it possible to acquire new information rapidly. However, they also introduce a critical problem that the brain must solve: each time the eyes move, all features in the visual scene change their position relative to the retina (self-centered reference, see Glossary), and yet our subjective experience is that their location remains stably defined in external world-based coordinates (world-centered reference). Such world-centered coordinates are useful; they not only stabilize perception, but also are critical to efficiently accomplish many actions, such as visual search or reaching. Despite decades of research revealing several interconnected mechanisms, how the brain generates and maintains world-centered representations of visual space remains unclear. Here we suggest a possible solution to this problem, based on how the brain solves a parallel problem in the navigational domain.

There are two primary ways navigators keep track of their position as they move. First, navigators can use path integration to keep track of their displacement, a process that involves the use of internal self-motion cues (e.g., vestibular or proprioceptive signals) without reference to the external world [421, 422]. A limitation of this strategy, however, is that tracking errors inevitably accumulate over time. Thus, an alternate strategy is world-centered navigation, which involves the use of external sensory cues to maintain a representation of navigable space that is invariant to one's own movements [423]. A key mediator of world-centered coding during navigation is the hippocampal formation (HF) [123, 315]. Here we review growing evidence that the HF not only supports world-centered navigation, but also represents world-centered visual space as well. This observation implies that this article's very own title is somewhat misleading. Referring to the HF as 'the brain's navigation system' does not capture the full range of cognitive functions it supports. Visual exploration and navigation might in fact not be two parallel operations, but may rather be two expressions of a common mechanism for exploring the world.

6.3 Medial temporal codes support navigation and vision

The hippocampus contains place cells that fire whenever a navigator occupies particular environmental locations [128]. Different place cells fire in different locations and thus, as a population, represent a world-centered map of navigable space. Place cells use self-motion information to maintain spatial representations in darkness for a short time, but eventually accumulate errors over longer timescales that are corrected using external sensory cues [140, 424]. Intriguingly, place cells also have an entirely visual analogue. Specifically, the primate hippocampus contains neurons that increase firing rate when specific environmental locations are viewed, irrespective of the location of the body, eye position in the orbit, or facing direction (Figure 6.1A, Key Figure) [220, 221]. These spatial view cells are updated based on eye movements when the preferred view is hidden in darkness or obscured, but tend to drift and become less sharply tuned [222]. Therefore, spatial view cells have similar properties as place cells, except tuned to world-centered gaze location instead of self-location. Place cells are not topographically organized in the rodent hippocampus [425] (i.e., cells tuned to nearby environmental locations are not more anatomically proximal than cells tuned to far apart locations), which may explain why previous studies did not observe a topographically organized representation of visual space in the hippocampus [87]. Notably, the presence of a world-centered map of visual space in the hippocampus does not necessarily imply that the hippocampus is involved in visual perception per se beyond providing a coordinate system (see Box 1 for further discussion).

One of the primary inputs to the hippocampus is the entorhinal cortex (EC). During navigation, entorhinal grid cells fire in a regular hexagonal lattice of locations that tile the floor of the environment [151]. Grid cells are thought to combine self-motion and environmental information to provide a stable metric for the place cell map [157]. Similar to grid cells, recordings of EC neurons in head-fixed monkeys revealed strikingly similar hexagonal firing patterns encoding the location of gaze during free viewing of visual scenes (Figure 6.1B) [225]. A proportion of these visual grid cells shift their firing fields in concert with translation of the visual scene, showing that some represent visual space in world-centered coordinates (Figure 6.1C) [226]. In two recent studies, we extended these findings by showing that also human EC supports a world-centered grid-like representation of visual space [170, 171].

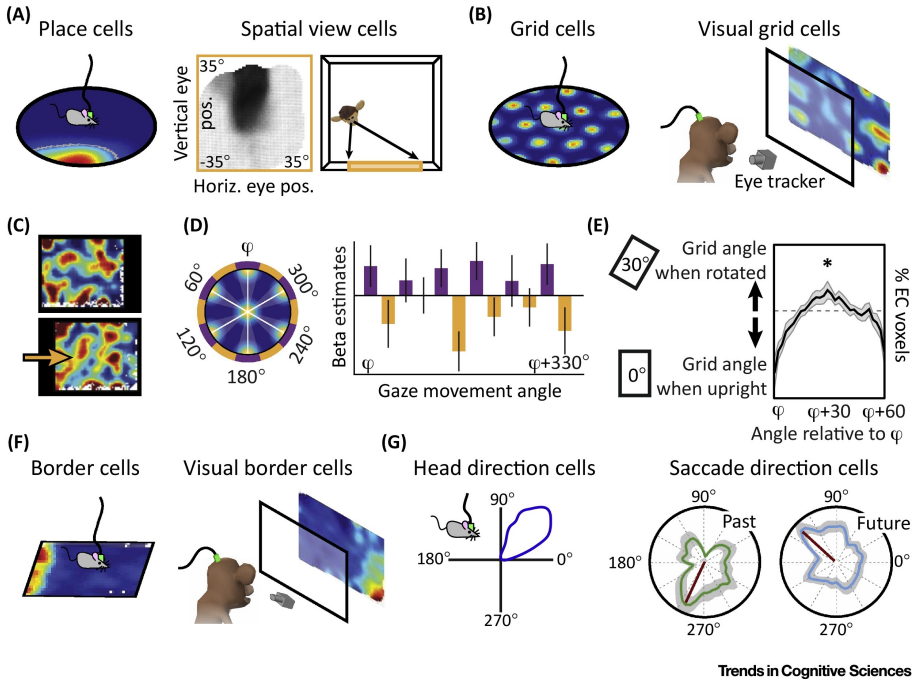


Figure 6.1: Key Figure: Spatial Codes in Vision and Navigation. (A) Place cells and spatial view cells in hippocampus. Left: rodent place cells encode self-location in the world (firing rate color-coded). Right: monkey spatial view cells encode world-centered gaze location (firing rate in dark color); example cell from [220]. (B) Grid cells and visual grid cells in entorhinal cortex (EC). Left: grid cells fire at several locations arranged in hexagonal lattice tiling the environment. Right: monkey visual grid cells show the same hexagonal firing but encode gaze position in the visual scene (adapted from [225]). (C) Grid pattern anchors to boundary. Firing pattern of some visual grid cells shifts (yellow arrow) in concert with shifts in the stimulus (adapted from [226])). (D) Human fMRI-grid signature. Left: visual grid cell model predicts higher activity for gaze directions parallel to the grid axes (and steps of 60°, white lines) than for directions in between. Right: fMRI activity in EC depends on gaze direction showing predicted sixfold rotational symmetry. (E) Visual boundary anchoring of human fMRI-grid signature. When search display rotates during visual search, fMRI grid signature in most voxels changes orientation, mirroring the search display rotation. (F) Border cells and visual border cells in EC. Left: rodent border cells encode proximity to navigational boundaries. Right: visual border cells encode proximity of the monkey's gaze to the edges of a visual stimulus (example cell from [225]). (G) Head direction cells and saccade direction cells in EC. Left: rodent head direction cells encode facing direction. Right: monkey saccade direction cells encode direction of future and past saccades (adapted from [229]).

Participants had their gaze position monitored while they performed an object tracking or visual search task. fMRI responses in EC exhibited a sixfold symmetry as a function of gaze movement direction, which is indicative of grid cell activ-

ity (Figure 6.1D) and akin to the fMRI grid-signature found during navigation [158]. This visual grid representation was absent when participants fixated but the visual scene was moved instead, suggesting that active movements of gaze or attention are required to drive visual grid cells. Passive transport likewise abolishes grid coding of navigable space in mice [426]. Critically, we also found that the EC visual grid signal is anchored to the visual scene in the same way as rodent grid cells representing navigable spaces [427,428], adopting reliable alignments to the borders of the visual stimulus and rotating in concert with rotation of these borders (Figure 6.1E).

Place and grid cell firing fields are likely anchored to the external world [351,429] by border cells in the EC and the subiculum that fire when navigators are a particular distance and direction from spatial boundaries [194,196,430]. Border cells are tuned not only to navigational obstacles like walls, but also to vertical cliffs, raising the possibility that such cells represent the edges of the local environment broadly [196]. Akin to border cells, monkey EC contains neurons that increase their firing rates when gaze is close to one or more of the borders of a visual stimulus [225] (Figure 6.1F). Visual border cells respond to the outer edges of a visual stimulus irrespective of the stimulus content. The borders of visual space have long been known to influence search efficiency in visual search tasks [431]. Importantly, unlike neurons in primary visual cortex tuned to orientated edges in retinotopic coordinates, visual border cells do not respond to edges within the visual stimulus itself.

Another critical component of the world-centered navigation system are head direction (HD) cells found in several subcortical and cortical structures, including EC [341]. HD cells fire based on the orientation of the head in the navigational plane, independent of body location, and are updated by a combination of vestibular and external sensory inputs, including visual information that references the HD signal to the environment [109]. Different HD cells have different preferred orientations and thus, as a population, are akin to a neural compass [341], likely acting in concert with border cells to orient place and grid cell firing fields relative to the external world. Similar to HD cells, neurons selective for eye movement direction were recently observed in monkey EC [229]. These saccade direction (SD) cells are tuned to the direction of an upcoming saccade or a previously completed saccade, or both, independent of the position the monkey is currently looking at in visual space (Figure 6.1G). Whether SD cells code direction in world-centered coordinates is unknown, but one possibility is that they represent SD relative to an axis defined by gravity.

Box 1. Is the hippocampal formation involved in visual perception?

Perceptual impairments following MTL damage, and MTL activation to visual stimuli, have raised the question whether the HF is involved in visual perception, independent of memory [382, 432, 433]. Some of the strongest evidence for this idea comes from studies examining hippocampal involvement in perception of visual scenes. In fMRI studies, hippocampal responses are stronger for visual scenes than for objects, even if no explicit task was performed [434], and several studies found that the ability to differentiate visual scenes depends on healthy hippocampal function [416–418, 435]. Such results have led to the proposal that the hippocampus is an integral part of the visual scene perception network [436].

An alternate possibility is that hippocampal responses to visual stimuli reflect the extent to which perception requires coding of relational information. Visual scenes are defined not only by their local features, but also by the specific global arrangement of those features. Such global feature arrangements engage the HF more strongly than details in a visual scene [405], and increasing overlap between images interferes with recognition to a greater extent in individuals with MTL damage compared with healthy controls [416, 417]. While the implications of such findings are still debated (e.g., [432, 437]), they suggest that the role of the HF in visual perception may be the processing of relational information, rather than perception of visual content per se, a process not restricted solely to perception or memory. This also resonates well with reports that hippocampal patients are impaired at imagining scenes with high spatial coherence [438].

Such results may explain why MTL perceptual effects are most pronounced for visual scenes defined not only by specific features (e.g., vase and table), but also by the spatial arrangement of those scene features (e.g., the vase stands on the table versus the vase stands under the table). Relational coding requires a coordinate system relative to which the locations of visual features can be specified. The MTL might provide just such a world-centered coordinate system that is invariant to the specific content of the scene. This implicates the HF in visual perception insofar as relational processing is required. In addition, as we discuss in the Recall and Planning section, the MTL likely also plays a role in perception by guiding overt [225] and covert [227] perceptual sampling, possibly by forming predictions [376] that also modulate visual processing [339].

In support of this idea, the orientation of the head relative to the gravity axis affects perception of the orientation of visual stimuli [439, 440]. The monkey anterior thalamus contains neurons that carry a gravity-anchored head orientation signal that could provide input to SD cells [441], similar to how thalamic HD cells serve as input to parahippocampal HD neurons in rodents [341]. These functional

and structural similarities suggest that primate SD cells may have an evolutionary origin similar to the rodent HD circuit [229].

In sum, each of the key neural mechanisms that represent a world-centered map (place cells, grid cells, border cells) and compass (HD cells) for navigation have purely visual analogues that encode a world-centered map of visual space (spatial view cells, visual grid cells, visual border cells) and the direction of eye movements (SD cells). The HF thus represents world-centered visual space using similar mechanisms as it does to encode world-centered navigable space, putting it at a key position to support computations in both domains. We suggest that the HF may provide the optimal solution to three key computational challenges shared by navigation and vision: reference frame transformations, recall and planning, and context specificity (which we will discuss below).

6.4 Reference frame transformations

All sensory information is self-centered, and yet our perceptual experience of the world is stable during movements. This remarkable phenomenon requires the brain to reconstruct self-motion invariant coordinates from noisy self-centered inputs [309], a computation critical for both navigation and vision. How we experience the world as stable despite variability in retinal input across eye movements (i.e., how the brain performs this transformation between reference frames) continues to be a matter of debate [40, 41, 442]. Consideration of the neural basis of such reference frame transformations during navigation suggests a possible solution by which self-centered visual representations are transformed into world-centered ones as well.

During navigation, self-centered coordinates are transformed into world-centered ones by a brain network consisting of the posterior parietal cortex (PPC), the medial parietal retrosplenial complex (RSC) that includes the retrosplenial and posterior cingulate cortices, and the HF [93, 97, 316] (Box 2). Each stage of this network contributes differentially to the two reference frames, with PPC mainly processing self-centered (often body-based) information, the HF encoding world-centered information, and RSC serving as the key transformation stage between the two [96, 97, 316]. This model suggests that the HF receives self-centered input from neocortex, and in turn projects world-centered coordinates back to guide navigational behavior [93, 97]. Coordinates are likely converted between the two reference frames by integrating external sensory information with proprioceptive

and vestibular signals related to self-motion and path integration [443].

Retinotopic representations of the visual field [13] are also likely transformed to world-centered ones by the PPC-RSC-MTL pathway. In PPC, some visual receptive fields are invariant to eye movements [50] (Figure 6.2A) and retinotopic representations are updated to compensate for eye displacement before eye movements are executed [51] (Figure 6.2B). This coordinate transformation is driven by integration of visual inputs with corollary discharges about impending eye movements [40], a process known as retinotopic updating [41,444]. Retinotopic updating is also observed in a number of other brain regions [41] and similar mechanisms support non-retinotopic encoding of visual motion in several monkey [64] and human [38,247] brain areas as well. It is often unclear which extra-retinal coordinate system these regions use to represent visual space, but head- and body-based coordinates have been observed. Moreover, some PPC neurons dynamically switch between reference frames depending on whether the direction of gaze is fixed relative to the head or body [89]. By integrating retinal information with self-motion signals, these regions compensate for self-induced changes to the visual scene and anticipate the consequences of future saccades [51, 444, 445]. However, retinotopic updating alone cannot account for a range of findings related to visual memory, such as head-centered (possibly world-centered) memory traces in retinotopic receptive fields [78] (Figure 6.2C). Moreover, like path integration, retinotopic updating is prone to error accumulation over time [446], as it requires constant updating of locations in the visual field, and thus a further corrective and stabilizing mechanism is needed.

One possibility is that visual representations in PPC, transformed to non-retinotopic coordinates by retinotopic updating and self-motion integration [40, 51, 447], serve as input to the HF, which then forms a robust world-centered map of visual space through predictive statistical learning [376]. The PPC-RSC pathway is well positioned for this signal transmission to the MTL [448] (Box 2). During navigation, RSC neurons in rodents encode turn direction, path position, and direction-dependent locations in route-based coordinates, as well as body orientation and location in multiple world-centered reference frames [112, 449–451] (Figure 6.2D), in line with human imaging results [113, 114]. Analogously, the posterior portion of RSC represents visual space in retinotopic coordinates [25], and the posterior cingulate represents visual space in world-centered coordinates [452] (Figure 6.2E). Rodent RSC also signals head movement information directly to early visual regions, referencing visual motion processing to the current status of the observer's head [453].

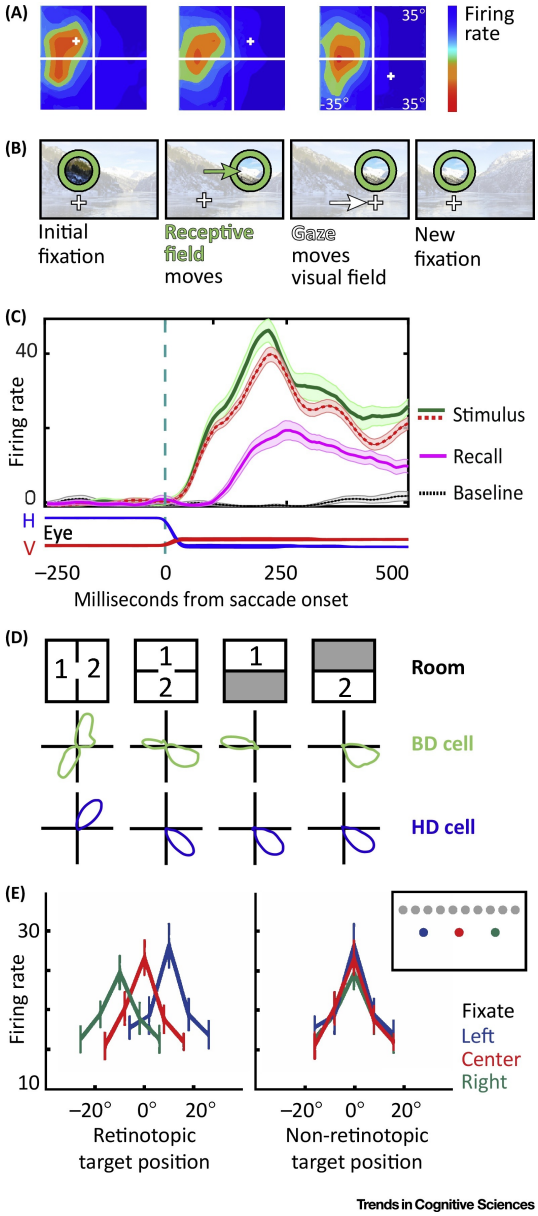


Figure 6.2: Reference Frame Transformation. (A) Non-retinotopic receptive fields (RF) in monkey parietal cortex. Single-cell rate map of the visual field for three gaze locations (white crosses). RF remains at the same location independent from gaze location (adapted from [50]). (B) Retinotopic updating. Retinotopic RF moves to postsaccadic visual field location before saccade is executed. Adapted from [51]. (C) Non-retinotopic memory signal in retinotopic neurons in monkey parietal cortex. Trial-averaged single-cell response for trials in which a saccade brings stimulus location into RF (green & red). Neuron responds also when stimulus is not shown in the particular trial (pink). Horizontal (H) and vertical (V) gaze movements depicted. Adapted from [78]. (D) Mixed reference frames in rodent retrosplenial cortex. Polar plots of firing rate versus head direction (HD) in different rooms for exemplary bidirectional (BD) and HD cells. Adapted from [112]. (E) Single-cell responses to visual stimuli (grey dots, upper right panel) at three gaze locations (color-coded). Responses plotted in retinotopic (left panel) and non-retinotopic locations (right panel). Adapted from [452].

The PPC-RSC-MTL pathway is bidirectional, suggesting that the MTL also communicates its world-centered visual representation back to neocortical areas. These top-down signals might modulate neocortical processing and contribute to eye-movement invariant receptive fields and visual stability. One possibility is that the MTL provides gaze controlling areas with spatial information about the visual

field before saccades are executed, helping to guide the shift of receptive fields during retinotopic updating [40, 51]. If so, the MTL could act in concert with other brain areas, such as the superior colliculus, which are known to send saccadic corollary discharge signals to cortical regions via the thalamus [40, 444]. However, deficits in visual stability following MTL lesions have not been previously reported, suggesting that the MTL is unlikely to participate in retinotopic updating directly.

Alternatively (or in addition), the MTL may play an important role in transsaccadic memory, an idea at the center of a longstanding debate [80, 454, 455] in which the MTL has been largely overlooked. Specifically, the MTL could give rise to a world-centered memory signal that modulates retinotopic neurons [78, 456] (Figure 6.2C) even after a saccade landed (see e.g., [78,80,454,455] for related discussion). Unlike presaccadic retinotopic updating, this MTL world-centered memory trace could be integrated with postsaccadic visual input, providing an efficient (i.e., no active compensation for self-motion required) and noise-resilient (i.e., externally anchored and robust to error accumulation) recalibrating mechanism that could be employed as needed. Areas such as PPC [78] or frontal eye fields [456] could match postsaccadic visual input to the eye-movement invariant visual field representation maintained in the MTL. For example, for each new fixation onset, the self-centered location at which salient visual information occurs could be compared with the MTL's world-centered map of visual space. Once a correspondence between the self- and world-centered locations is found, the vector that encodes the receptive field shift required to recalibrate the self-centered representation could be fed back to the cortex. The entorhinal grid system has been implicated in this type of vector computation [154, 457]. This proposal does not predict that individuals with MTL lesions have visual stability deficits, since retinotopic updating remains intact, but rather predicts that such individuals should have an increased sensitivity to visual localization error accumulation across saccades. To be useful for such a recalibration process, a maplike memory of the visual field must be maintained in the MTL for at least the time intervals relevant for working memory, a time scale on which the HF does indeed maintain visual memory, as demonstrated by several hippocampal lesion studies (e.g., [432, 458–460]).

Box 2. Anatomy of neocortical–hippocampal interactions in vision

A concept guiding vision science for several decades is the dichotomous organization of the visual system into two major multisynaptic pathways, the dorsal and ventral visual stream. Here, we want to emphasize that these pathways, thought to mediate different aspects of vision, both converge on the MTL in the primate brain [12].

A key processing and relay station of the dorsal occipitoparietal stream is the inferior parietal lobule [12]. It connects not only many parietal, temporal, and occipital regions involved in visuospatial processing, sensory-motor integration, and action planning (e.g., lateral and ventral intraparietal areas, areas V6 and V6A, and the MT+ motion complex [64], but also projects to the gaze-controlling frontal eye fields (FEF) [461] as well as directly and via RSC [122] to CA1 [95] and subicular [462], parahippocampal, and entorhinal cortices [152, 463]. Parietal cortex is involved in integrating visual input with gaze movements [50, 51] in concert with the FEF [461] and is directly, as well as via FEF and the thalamus, connected to the superior colliculus [40], which is thought to be the prime source or relay station for gaze-related efference copies [40, 444]. Thus, this dorsal pathway is most suited to provide the MTL with gaze- and self-motion information.

By contrast, ventral visual areas support visual representations that are invariant to eye movements [11]. The ventral visual stream is a strongly recursively connected network, spanning from early visual cortex via areas V4 and the MT+ [464], inferotemporal areas TEO and TE [465], and RSC [122] to the most anterior parts of the inferior temporal lobe and the HF [466]. It processes information predominantly related to object quality and is thought to extract perceptually relevant features from visual scenes [467], irrespective of their location. However, recent work has found information about locations of visual objects in higher-order ventral visual regions [468], raising the possibility that the ventral pathway provides visual positional information to the MTL as well.

The MTL is hence a convergence zone for visual information in the brain. How might the MTL in turn shape neocortical processing and guide behavior? The key mediator of cortico-hippocampal interactions is the EC [152, 463], receiving strong hippocampal input via subicular cortices [462] and projecting to many regions on frontal, temporal, and parietal cortices [469], as well as the RSC [121, 122]. The latter is a likely mediator between visual and mediotemporal processing given its strong connectivity to the neocortex. In sum, connectivity suggests strong interactions between visual and mediotemporal systems, putting the hippocampal formation at a key position to shape vision.

6.5 Recall and planning

Like navigation, a key function of visual exploration is to acquire new information about the world. The relationship between viewing and memory (reviewed by [217, 314, 411]) is typically examined in image recognition tasks, in which sequences of pictures are presented while humans or monkeys indicate whether they have or have not seen these pictures before. Viewing behavior differs between later remembered and forgotten images, with the number [470,471] and duration [470] of fixations linked to successful memory encoding. Hippocampal activity directly depends on and predicts visual sampling, even for images that were not consciously remembered [415, 472] (Figure 6.3A), suggesting that the timing of hippocampal mnemonic processing is tightly linked to visual exploration. Indeed, the timing of saccades themselves is not random but phase-locked to neural oscillations in visual and MTL areas during successful memory encoding [473] (Figure 6.3B).

Gaze movements interrupt the flow of incoming visual information, prompting the HF to switch between active and inactive encoding states around the time of saccades. The switch between encoding states may be driven by a phase-reset of the hippocampal theta rhythm [402, 403] (Figure 6.3C), anticipating new incoming sensory information whose encoding depends on a precise interplay between hippocampal spikes and theta phase [401]. Theta rhythms are also critical for grid cell activity during navigation, which lose spatial periodicity when theta is inhibited [399, 400], and for place cells during navigation that fire at specific phases of a theta cycle [370]. While the precise functional role of neural oscillations in the MTL is not yet fully understood [474], they clearly play a critical role in both visual and navigational domains.

Importantly, the relationship between viewing and memory is not unidirectional; what we remember also directly influences how we explore the world (for review see e.g., [411]). Once an image has been memorized, it is less visually explored when it is repeated than when it is novel [225, 475, 476], a change in viewing behavior that depends on the hippocampus [475, 477]. Further, when participants were asked to recall images they had seen before, their eyes re-enacted the same movements they showed during encoding, an effect causally connected to the quality of the recalled memory [412].

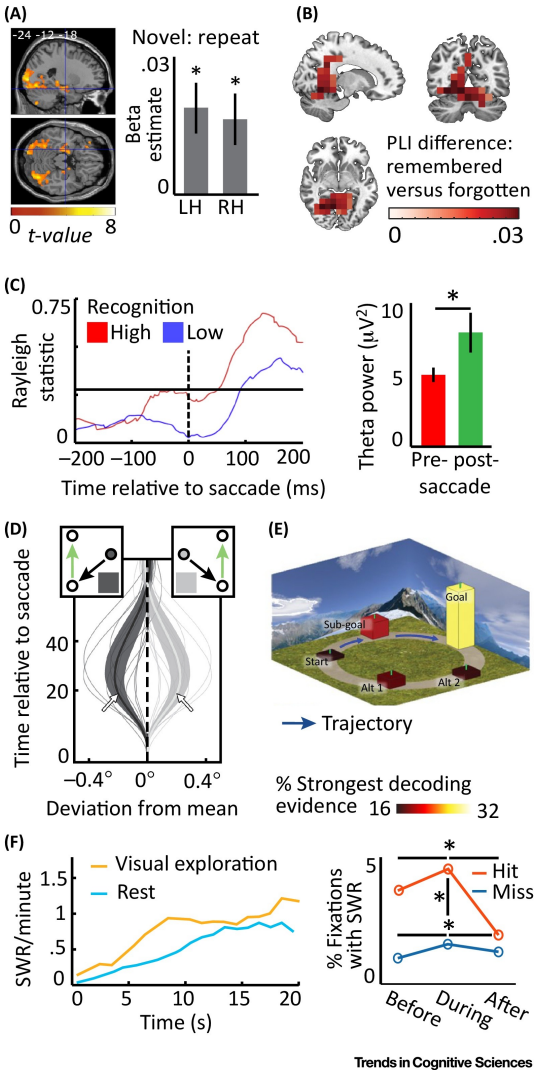


Figure 6.3: Memory-Guided Planning. (A) Visual sampling of novel images predicts human hippocampal activity. Left: linear parametric modulation (LPM) of number of fixations (t-map threshold P = 0.005, 10-voxel extension, uncorrected). Right: LPM is stronger for novel versus repeated images in left (LH) and right (RH) hemispheric hippocampus. Adapted from [472]. (B) Saccades are phase locked to alpha oscillations during successful memory encoding. Phase locking index [across-trial local field potential (LFP), phase-locking index (PLI)] differs significantly between later-remembered and later-forgotten images in human medial temporal lobe. Adapted from [473]. (C) Saccades reset theta phase in monkey hippocampus. Left panel: increase in LFP phase coherence after saccade is higher for high versus low recognition trials. Right panel: theta power increases significantly after saccade. Adapted from [402]. (D) Non-retinotopic carryover effects in saccade sequences in humans. Second saccade (green arrow in task insets) curves away from first fixation location (adapted from [369]). (E) Navigational route planning in human hippocampus. Decoder trained on goal location and tested on prenavigation planning periods favors subgoals on taken path over those on alternative path. Adapted from [478]. (F) Sharp wave ripples (SWR) during visual exploration in monkey. Left panel: SWR rate increases over time after trial onset. Right panel: SWRs occur more frequently in successful versus unsuccessful visual search trials. Figures adapted from [479].

Such memory-guided viewing has been intensively studied using the 'inhibition of return' (IOR) phenomenon, which broadly refers to the fact that after attending to a given visual location, reaction times for returning gaze to that location increase. IOR has been proposed as a novelty-seeking mechanism that maximizes efficiency when exploring a visual scene (for review see [480]). IOR for example has been shown to rely on non-retinotopic visual coordinates [76]. Moreover, when two saccades are performed in a sequence, the non-retinotopic (possibly world-centered) location of the first fixation spot influences the curvature of the subsequent saccade [369] (Figure 6.3D), suggesting that a top-down non-retinotopic signal influences saccade execution or planning. Such saccade sequences also often have latencies too short to plan and initiate each saccade separately [367,368], requiring the sequence to be preplanned in coordinates invariant to eye movements because all precomputed SDs will have changed after the first saccade [217]. Thus, planning and controlling oculomotor behavior likely requires coordinates referenced to the external world. Since the MTL contributes to goal detection during visual search [481] and contains both a world-centered visual map and SD cells, it is well positioned to perform such computations. In fact, the required computations are similar to those of goal-directed navigational route planning for which the HF is critical [478] (Figure 6.3E). In rodents, memory-guided route planning during navigation has further been linked to hippocampal sharp-wave ripples [482], which have also been observed in monkeys during visual exploration [479] (Figure 6.3F), especially when the presented images are repeated [483]. Growing evidence also shows that the same MTL mechanisms guide (or are guided by) the position of visual attention independent of gaze [227,406]. Together, these results suggest that the MTL drives viewing behavior to efficiently acquire new information about the visual environment [484], in line with the idea that gaze is a behavioral expression of visual predictions [410] generated by the HF [376].

6.6 Context specificity

In addition to encoding a map of the local spatial environment, the hippocampus stores multiple world-centered maps of navigable space (in the 'cognitive atlas'), allowing it to represent locations in multiple navigational contexts [142]. The ability of the hippocampus to distinguish between contexts during navigation is indexed by remapping, in which contextual changes cause all simultaneously recorded neurons to shift place fields to new locations or stop firing altogether, quickly resulting in a new and distinct spatial representation [141,485] (Figure 6.4A).

The emergence of remapping depends on several factors, including a navigator's experience with a context [143,486], and can be eliminated by inhibiting hippocampal plasticity [487]. These mnemonic components indicate that remapping facilitates contextual memory during navigation, rather than perceptual processing.

Context also plays an important role in visual tasks by guiding visual search and recognition. In a now standard demonstration of this idea [488], the spatial configuration of an array of distractors in a visual search display provides a unique context that reliably determines the location of a search target. Participants typically find the visual search target faster when they have had prior exposure to the visual context, an effect referred to as 'contextual cueing' [489]. Hippocampal remapping may provide a critical mechanism underlying such visual context effects, by storing multiple maps of visual space (in a 'cognitive picture book') for multiple visual contexts. fMRI studies have shown that contextual cueing is mediated by the hippocampus [490, 491] (Figure 6.4B), and patients with MTL damage do not show a search benefit for repeated arrays [492–494]. Hippocampal volume correlates with the magnitude of contextual cueing in typical older adults and adults diagnosed with mild cognitive impairment associated with MTL atrophy [495] (Figure 6.4C). The strength of contextual cueing is modulated by a viewer's experience with the visual context [496, 497], similar to the dependence of remapping on navigational experience [143, 486]. Notably, visual context also modulates fMRI activity in PPC and RSC during search tasks, suggesting that not only world-centered coordinates as reviewed above but also hippocampal context representations may feedback to the broader visuospatial mapping network to guide context-dependent viewing behavior [498].

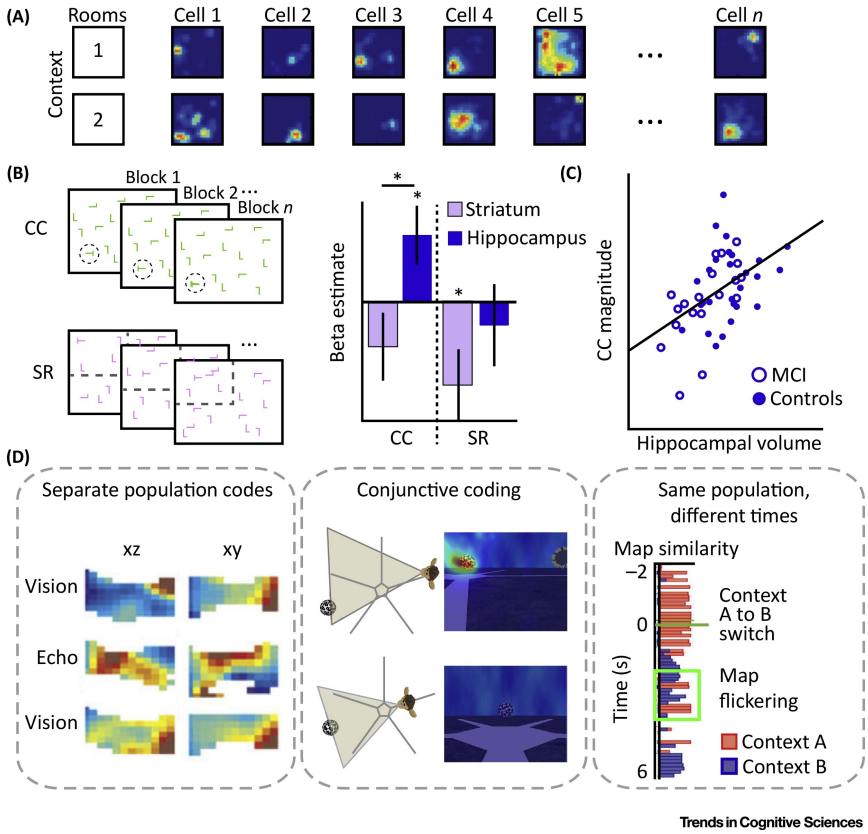


Figure 6.4: Context Specificity. (A) Hippocampal remapping. Recruitment of a new place cell maps specifically to each navigational context, such as different rooms (adapted from [485]). (B) The hippocampus mediates visual context. Left: participants performed a visual search task during fMRI. Target location was cued either via specific distractor arrangement [contextual cueing (CC)] or probabilistically by stimulus color [stimulus response (SR)]. Right: hippocampus predicted search benefits in CC, the striatum in SR conditions (adapted from [491]). (C) Hippocampal volume correlates with the magnitude of visual contextual cueing (mean response time on novel minus repeated search configurations) in typical humans and patients with mild cognitive impairment (MCI) (adapted from [495]). (D) Three possible hippocampal codes for vision and navigation. Left: same population, different population code. Example: place cells in flying bats remap depending on whether they use vision or echolocation to navigate in the same context (xz is horizontal plane, yz is vertical plane; cell adapted from [499]). Middle: conjunctive coding. Example: monkey hippocampal neurons conjunctively represent gaze- and self-location (example cell adapted from [223, 500]). Right: same population, different time. Example: after one navigational context is rapidly switched to another, rodent hippocampal place cell population spontaneously flickers back-and-forth between representations of the two contexts (map similarity is the correlation between the current population response vector and the average population response vector in context A and B, respectively; adapted from [501]).

The existence of multiple maplike representations in the hippocampus raises three alternatives for how the hippocampus supports maps of both navigational and visual spaces (Figure 6.4D). First, separate hippocampal populations may mediate representations for the two domains and remap between them depending on their behavioral relevance. In the bat hippocampus, place fields remap depending on whether bats employ vision in light or echolocation in darkness to navigate [499], indicating that, in this case at least, space perceived using different sensory modalities is mapped using separable neural representations. Second, some neurons in the monkey and human hippocampus are sensitive to both body and view location during navigation [135, 223]. Thus, the same hippocampal populations may conjunctively represent navigational and visual spaces. Finally, the same neuronal population may represent both navigational and visual spaces, but at different times, and alternate between visual and navigational maps during simultaneous eye- and body-based movements. Consistent with this idea, the hippocampus 'flickers' between representations of two distinct navigational contexts at different theta cycles when there are abrupt shifts in context [501]. Future research using behavioral tasks that require simultaneous monitoring of locations in both visual and navigational spaces are needed to dissociate these alternatives.

6.7 Concluding remarks

It has long been known that vision is important for navigation [213]. However, the converse is less well appreciated. Here we have reviewed mechanisms in the HF that provide the fundamental resource to support several critical computations shared by navigation and vision. We have proposed that the HF, traditionally believed to support navigation, also mediates a world-centered representation of visual space and guides viewing behavior. Since primates are particularly visual creatures, the neural mechanisms that evolved to support navigation in rodents may have been co-opted to support visual exploration as well, leading to strongly intertwined navigational and visual mapping systems. In support of this view, we have drawn on data from both primates and rodents; hence, an important caveat is that we may have elided relevant species differences. The rodent HF mediates representations of non-navigational spaces, such as spaces defined purely by auditory information [237], but whether it mediates a map of visual space is unknown. In primates, visual and navigational mapping systems overlap on a systems level, but whether the two domains are supported by the same neurons remains as yet unknown (see Outstanding Questions).

As a coda, we wish to note our belief that the MTL computations reviewed here likely have broad applications beyond both navigation and vision. The idea that the HF performs domain general computations is not new. Indeed, Tolman originally conceptualized the cognitive map as a heuristic for flexibly guiding behavior in general [137]. Yet, it is only recently that empirical research has begun to take this idea beyond metaphor. We encourage researchers of cognitive domains other than navigation and memory to 'look at' the HF, and further suggest that vision may prove to be the ideal domain for future explorations of how MTL computations support cognition broadly, because visual representations can be characterized concretely in terms of distances and directions in the same way as navigational spaces. By testing whether the same principles govern spatial representations for vision and navigation, and whether the MTL plays the same role in solving analogous problems in both domains, we will have made a key step toward illuminating the function of the MTL in general.

Box 3. Outstanding questions

How, and to what environmental features, are world-centered maps of visual space anchored in a dynamically changing world (e.g., during real-world navigation)?

Do head direction and saccade direction cells play analogous orienting roles in navigation and vision, respectively?

Since visual space does not contain eye movement obstacles in the same way as navigational space, what information drives visual border cells?

Do spatial view and visual grid cells code visual space in depth (e.g., along the ground plane), or only in the two-dimensional visual plane?

Does the MTL world-centered map of visual space support visual constancy (by guiding retinotopic updating or by mediating memory traces of the visual field)?

Is contextual cueing mediated by hippocampal remapping?

Do MTL visual representations drive or are they driven by shifts in visual attention?

Do the same MTL neural populations mediate world-centered representations of visual and navigational spaces (Figure 6.4C) and how do they interact during navigation?

Do MTL world-centered visual representations emerge in typical development at the same time as world-centered representations of navigational space? Relatedly, do visual representations in the MTL break down along similar trajectories in typical aging and disease as the coding of navigational spaces?

Glossary 1

Border cell: entorhinal or subicular neuron that is active when the animal occupies locations along navigational boundaries, such as walls or cliffs.

Cognitive map: first proposed as mental representation of locations and the relationship between them, cognitive maps are discussed as general map-based coding principle for information in the brain.

Context: spatial, temporal, and situational setting associated with particular behavioral or mnemonic outputs.

Corollary discharge: also known as 'efference copy', it is a copy of a motor command that is sent to the muscles to produce a movement. Sensory areas use corollary discharges to anticipate self-motion related sensory change.

Grid cell: entorhinal neuron that is active when the animal occupies certain locations arranged in a hexagonal lattice, tessellating the environment. Grid cells encode self-location during navigation.

Head direction (HD) cell: neuron that is active when the animal faces into a certain direction in the environment. Head direction cells were found in several brain areas, including the hippocampal formation.

Hippocampal formation: compound of brain regions in the temporal lobe, including the hippocampus proper with its subfields CA1, CA2, and CA3; the dentate gyrus; subicular cortices; as well as the entorhinal cortex.

Non-retinotopic: umbrella term describing representations of visual space that are not retinotopic (see definition 'Retinotopic'). Includes all movement-invariant reference frames, such as world-centered (spatiotopic) and head-centered (craniotopic) reference.

Place cell: hippocampal neuron that is active when an animal occupies a certain location in the environment. Place cells encode self-location during navigation.

Proprioceptive signal: neural representation of mechanoreceptive information about tendon status and muscle tone. Together with vestibular signals (see definition 'Vestibular signal'), proprioception provides information about position of body parts and movements.

Glossary 2

Retinotopic: retinotopy is the mapping of visual space from the retina to neurons in visual cortex. Here, neighboring visual locations drive neighboring cells. Retinotopic maps are self-centered (see definition 'Self-centered reference').

Saccade direction (SD) cell: saccade direction cell; neuron that is active when the eyes move into a certain direction.

Self-centered reference: coordinate system referenced to one's own body, or parts of it, such as the eye (see definition 'Retinotopy'). Also known as 'egocentric'.

Spatial view cell: hippocampal neuron that is active when an animal looks at a certain location in the environment. Spatial view cells encode gaze location during visual exploration.

Transsaccadic memory: visual short-term memory representation of the presaccadic visual field that influences postsaccadic processing. Transsaccadic memory is discussed as one of multiple mechanisms mediating perceptual stability.

Vestibular signal: neural representation of sensory information about head/body movements relative to the axis of gravity. Vestibular information constitutes an important cue about self-motion and body position in space.

Visual border cells: entorhinal neuron that is active when the animal looks at locations close to the edge of a visual stimulus display. Does not respond to edges within the display.

Visual grid cell: entorhinal neuron that is active when the animal looks at or attends to certain locations in the visual scene. Its receptive fields are arranged in a hexagonal lattice, tessellating visual space. Visual grid cells encode gaze location during visual exploration.

Visual exploration: behavioral strategy to explore the environment by means of eye movements and viewing without the need of navigating, hence without physically moving through the environment.

World-centered reference: self-motion invariant coordinates referenced to external cues such as landmarks or visual features in a scene. Also known as 'allocentric'. Also see the definition of 'Non-retinotopic'.

References

[1] H. Munk. *Ueber die Functionen der Grosshirnrinde : gesammelte Mittheilungen aus den Jahren 1877-80. Mit Einleitung und Anmerkungen*. Berlin, Verlag von August Hirschwald.

[2] S. E. Henschen. On the visual path and centre. *Brain*, 16(1):170–180.

[3] T. Inouye. *Die Sehstoerungen bei Schussverletzungen der kortikalen Sehsphaere nach Beobachtungen an Versundeten der letzten Japanische Kriege*. Leipzig, W. Engelmann.

[4] G. Holmes. Disturbances of vision by cerebral lesions. *The British journal of ophthalmology*, 47(9):472–5, 1918-09.

[5] S. A. Talbot and W. H. Marshall. Physiological studies on neural mechanisms of visual localization and discrimination. *American Journal of Ophthalmology*, 24(11):1255–1264, 1941.

[6] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3):574–591.

[7] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106–154, 1962-01.

[8] D. C. Van Essen and J. H. R. Maunsell. Hierarchical organization and functional streams in the visual cortex. *Trends in Neurosciences*, 6:370–375, 1983-01.

[9] M. A. Goodale and A. D. Milner. Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1):20–25.

[10] M. Mishkin, L. G. Ungerleider, and K. A. Macko. Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*, 6:414–417, 1983-01.

[11] D. J. Kravitz, K. S. Saleem, C. I. Baker, L. G. Ungerleider, and M. Mishkin. The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, 17(1):26–49, 2013.

[12] D. J. Kravitz, K. S. Saleem, C. I. Baker, and M. Mishkin. A new neural framework for visuospatial processing. *Journal of Vision*, 11(11):319–319, 2011.

[13] B. A. Wandell, S. O. Dumoulin, and A. A. Brewer. Visual field maps in human cortex. *Neuron*, 56(2):366–383, 2007-10.

[14] M. I. Sereno, A. M. Dale, J. B. Reppas, K. K. Kwong, J. W. Belliveau, T. J. Brady, B. R. Rosen, and R. B. Tootell. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, 268(5212):889–93, 1995-05.

[15] E. A. DeYoe, G. J. Carman, P. Bandettini, S. Glickman, J. Wieser, R. Cox, D. Miller, and J. Neitz. Mapping striate and extrastriate visual areas in human cerebral cortex. *Proceedings of the National Academy of Sciences*, 93(6):2382–2386.

[16] C. Tyler, L. Likova, C. C. Chen, L. Kontsevich, M. Schira, and A. Wade. Extended concepts of occipital retinotopy. *Current Medical Imaging Reviews*, 1(3):319–329.

[17] B. A. Wandell and J. Winawer. Imaging retinotopic maps in the human brain. *Vision research*, 51(7):718–37, 2011-04.

[18] S. O. Dumoulin and B. A. Wandell. Population receptive field estimates in human visual cortex. *NeuroImage*, 39(2):647–60, 2007-01.

[19] B. A. Wandell and J. Winawer. Computational neuroimaging and population receptive fields. *Trends in Cognitive Sciences*, 19(6):349–357, 2015-06.

[20] R. B. Tootell, J. D. Mendola, N. K. Hadjikhani, P. J. Ledden, A. K. Liu, J. B. Reppas, M. I. Sereno, and A. M. Dale. Functional analysis of v3a and related areas in human visual cortex. *The Journal of neuroscience*, 17(18):7060–78, 1997-09.

[21] M. J. Arcaro, S. McMains, B. D. Singer, and S. Kastner. Retinotopic organization of human ventral visual cortex. *The Journal of neuroscience*, 29(34):10638–52, 2009-08.

[22] J. D. Swisher, M. A. Halko, L. B. Merabet, S. A. McMains, and D. C. Somers. Visual topography of human intraparietal sulcus. *The Journal of neuroscience*, 27(20):5326–5337, 2007-05.

[23] J. Larsson and D. J. Heeger. Two retinotopic visual areas in human lateral occipital cortex. *The Journal of neuroscience*, 26(51):13128–42, 2006-12.

[24] S. Pitzalis, C. Galletti, R. S. Huang, F. Patria, G. Committeri, G. Galati, P. Fattori, and M. I. Sereno. Wide-field retinotopy defines human cortical visual area v6. *The Journal of neuroscience*, 26(30):7962–73, 2006-07.

[25] E. Silson, A. Steel, and C. I. Baker. Scene selectivity and retinotopy in medial parietal cortex. *Journal of Vision*, 16(12):528, 2016.

[26] W. Chen, X. H. Zhu, K. R. Thulborn, and K. Ugurbil. Retinotopic mapping of lateral geniculate nucleus in humans using functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences*, 96(5):2430–2434.

[27] D. M. van Es, W. van der Zwaag, and T. Knapen. Topographic maps of visual space in the human cerebellum. *Current Biology*, 29(10):1689–1694.e3, 2019-05.

[28] G. A. Keliris, Q. Li, A. Papanikolaou, N. K. Logothetis, and S. M. Smirnakis. Estimating average single-neuron visual receptive field sizes by fMRI. *Proceedings of the National Academy of Sciences*, 116(13):6425–6434, 2019.

[29] M. D. Lescroart and J. L. Gallant. Human scene-selective areas represent 3d configurations of surfaces. *Neuron*, 101(1):178–192.e7, 2019.

[30] A. G. Huth, W. A. De Heer, T. L. Griffiths, F. E. Theunissen, and J. L. Gallant. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600):453–458, 2016.

[31] T. C. Sprague, E. F. Ester, and J. T. Serences. Restoring latent visual working memory representations in human cortex. *Neuron*, 91(3):694–707, 2016.

[32] N. Kriegeskorte and D. K. Pamela. Interpreting encoding and decoding models. *Current Opinion in Neurobiology*, 55:167–179, 2019.

[33] T. Naselaris, K. N. Kay, S. Nishimoto, and J. L. Gallant. Encoding and decoding in fMRI. *NeuroImage*, 56(2):400–410, 2011.

[34] D. Melcher and M. C. Morrone. Nonretinotopic visual processing in the brain. *Visual Neuroscience*, 32.

[35] W. A. Freiwald. Attention to objects made of features. *Trends in Cognitive Sciences*, 11(11):453–454.

[36] Z. Lin and S. He. Seeing the invisible: The scope and limits of unconscious processing in binocular rivalry. *Progress in Neurobiology*, 87(4):195–211.

[37] J. M. Henderson and A. Hollingworth. Eye movements during scene viewing. In *Eye Guidance in Reading and Scene Perception*, pages 269–293. Elsevier.

[38] D. C. Burr and M. C. Morrone. Spatiotopic coding and remapping in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1564):504–515, 2011.

[39] M. A. Sommer and R. H. Wurtz. A pathway in primate brain for internal monitoring of movements. *Science*, 296(5572):1480–1482, 2002.

[40] M. A. Sommer and R. H. Wurtz. Brain circuits for the internal monitoring of movements. *Annual Review of Neuroscience*, 31(1):317–338, 2008-07.

[41] L. D. Sun and M. E. Goldberg. Corollary discharge and oculomotor proprioception: Cortical mechanisms for spatially accurate vision. *Annual Review of Vision Science*, 2(1):61–84, 2016.

[42] R. H. Wurtz. Neuronal mechanisms of visual stability. *Vision research*, 48:2070–2089, 2008.

[43] H. von Helmholtz. *Handbuch der physiologischen optik, vol. 3*. Leopold V., 1860.

[44] E. von Holst and H. Mittelstaedt. Das reafferenzprinzip. *Naturwissenschaften*, 37(20):464–476, 2004.

[45] C. Y. Chen, L. Sonnenberg, S. Weller, T. Witschel, and Z. M. Hafed. Spatial frequency sensitivity in macaque midbrain. *Nature Communications*, 9(1).

[46] M. C. Morrone, J. Ross, and D. C. Burr. Saccadic eye movements cause compression of time as well as space. *Nature Neuroscience*, 8(7):950–954, 2005-07.

[47] C. Galletti, S. Squatrito, Piero Paolo B., and M. G. Maioli. 'real-motion' cells in the primary visual cortex of macaque monkeys. *Brain Research*, 301(1):95–110, 1984.

[48] R. G. Erickson and P. Thier. A neuronal correlate of spatial stability during periods of self-induced visual motion. *Experimental Brain Research*, 86(3):608–16, 1991-09.

[49] E. P. Merriam and C. L. Colby. Active vision in parietal and extrastriate cortex. *The Neuroscientist*, 11(5):484–493.

[50] J. R. Duhamel, F. Bremmer, S. B. Hamed, and W. Graf. Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature*, 389(6653):845, 1997.

[51] J. R. Duhamel, C. Colby, and M. E. Goldberg. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, 255(5040):90–92, 1992.

[52] A. S. Tolias, T. Moore, S. M. Smirnakis, E. J. Tehovnik, A. G. Siapas, and P. H. Schiller. Eye movements modulate visual receptive fields of v4 neurons. *Neuron*, 29(3):757–67, 2001-03.

[53] B. Bridgeman. Efference copy and its limitations. *Computers in biology and medicine*, 37(7):924–9, 2007-07.

[54] B. Bridgeman and L. Stark. Ocular proprioception and efference copy in registering visual direction. *Vision Research*, 31(11):1903–1913.

[55] C. Galletti, P. P. Battaglini, and P. Fattori. Eye position influence on the parieto-occipital area PO (v6) of the macaque monkey. *European Journal of Neuroscience*, 7(12):2486–2501.

[56] F. Bremmer, A. Schlack, J. R. Duhamel, W. Graf, and G. R. Fink. Space coding in primate posterior parietal cortex. *NeuroImage*, 14(1):S46–S51, 2001.

[57] R. Andersen and V. B. Mountcastle. The influence of the angle of gaze upon the excitability of the light- sensitive neurons of the posterior parietal cortex. *The Journal of Neuroscience*, 3(3):532–548.

[58] D. Zipser and R. A. Andersen. A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, 331(6158):679–684.

[59] Y. Trotter and S. Celebrini. Gaze direction controls response gain in primary visual-cortex neurons. *Nature*, 398(6724):239–242.

[60] N. Daddaoua, P. W. Dicke, and P. Thier. Eye position information is used to compensate the consequences of ocular torsion on v1 receptive fields. *Nature communications*, 5:3047, 2014.

[61] E. Fischer, H. H. Buelthoff, N. K. Logothetis, and A. Bartels. Human areas v3a and v6 compensate for self-induced planar visual motion. *Neuron*, 73(6):1228–1240, 2012-03.

[62] C. Galletti, P.P. Battaglini, and P. Fattori. Real-motion cells in area v3a of macaque visual cortex. *Experimental Brain Research*, 82(1):67–76, 1990-08.

[63] C. Galletti, P. P. Battaglini, and P. Fattori. Functional properties of neurons in the anterior bank of the parieto-occipital sulcus of the macaque monkey. *The European journal of neuroscience*, 3(5):452–461, 1991-01.

[64] C. Galletti and P. Fattori. The dorsal visual stream revisited: Stable circuits or dynamic pathways? *Cortex*, 98:203–217, 2018.

[65] E. Fischer, H. H. Buelthoff, N. K. Logothetis, and A. Bartels. Visual motion responses in the posterior cingulate sulcus: A comparison to v5/MT and MST. *Cerebral Cortex*, 22(4):865–876, 2011.

[66] D. M. Arnoldussen, J. Goossens, and A. V. van den Berg. Adjacent visual representations of self-motion in different reference frames. *Proceedings of the National Academy of Sciences of the United States of America*, 108(28):11668–73, 2011-07.

[67] J. Goossens, S. P. Dukelow, R. S. Menon, T. Vilis, and A. V. van den Berg. Representation of head-centric flow in the human motion complex. *The Journal of neuroscience*, 26(21):5616–27, 2006-05.

[68] G. D'Avossa, M. Tosetti, S. Crespi, L. Biagi, D. C. Burr, and M. C. Morrone. Spatiotopic selectivity of BOLD responses to visual motion in human area MT. *Nature Neuroscience*, 10(2):249–255, 2007.

[69] E. P. Merriam, C. R. Genovese, and C. L. Colby. Spatial updating in human parietal cortex. *Neuron*, 39(2):361–373, 2003.

[70] S. Crespi, L. Biagi, G. D'Avossa, D. C. Burr, M. Tosetti, and M. C. Morrone. Spatiotopic coding of BOLD signal in human visual cortex depends on spatial attention. *PLoS ONE*, 6(7), 2011.

[71] F. Strappini, S. Pitzalis, A. Z. Snyder, M. P. McAvoy, M. I. Sereno, M. Corbetta, and G. L. Shulman. Eye position modulates retinotopic responses in early visual areas: a bias for the straight-ahead direction. *Brain Structure and Function*, 220(5):2587–2601.

[72] H. C. Barron, M. M. Garvert, and T. E. J. Behrens. Repetition suppression: a means to index neural representations using BOLD? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1705):20150355.

[73] A. McKyton and E. Zohary. Beyond retinotopic mapping: The spatial representation of objects in the human lateral occipital complex. *Cerebral Cortex*, 17(5):1164–1172.

[74] J. D. Golomb, A. R. Albrecht, S. Park, and M. M. Chun. Eye movements help link different views in scene-selective cortex. *Cerebral Cortex*, 21(9):2094–2102.

[75] E. Zimmermann, M. C. Morrone, and D. C. Burr. Buildup of spatial information over time and across eye-movements. *Behavioural Brain Research*, 275:281–287.

[76] M. D. Hilchey, R. M. Klein, J. Satel, and Z. Wang. Oculomotor inhibition of return how soon is it recoded.pdf. *Attention, Perception, & Psychophysics*, 74(6):1145–1153, 2012.

[77] S. Nishimoto, A. G. Huth, N. Y. Bilenko, and J. L. Gallant. Eye movement-invariant

representations in the human visual system. *Journal of Vision*, 17(1):11, 2017.

[78] M. Semework, S. C. Steenrod, and M. E. Goldberg. A spatial memory signal shows that the parietal cortex has access to a craniotopic representation of space. *eLife*, 7, 2018-02-16.

[79] D. Melcher and M. C. Morrone. Transsaccadic memory. In *Eye Movements*, pages 213–233. Elsevier.

[80] D. Melcher and C. L. Colby. Trans-saccadic perception. *Trends in Cognitive Sciences*, 12(12):466–473, 2008.

[81] N. J. Kleene and M. M. Michel. The capacity of trans-saccadic memory in visual search. *Psychological Review*, 125(3):391–408.

[82] D. A. Cronin and D. E. Irwin. Visual working memory supports perceptual stability across saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 44(11):1739–1759.

[83] D. Aagten-Murphy and P. M. Bays. Functions of memory across saccadic eye movements. Springer Berlin Heidelberg.

[84] G. Edwards, R. VanRullen, and P. Cavanagh. Decoding trans-saccadic memory. *The Journal of Neuroscience*, 38(5):1114–1123, 2018.

[85] D. E. Irwin. Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(2):307–317.

[86] A. F. Ten Brink, J. H. Fabius, N. A. Weaver, T. C. W. Nijboer, and S. Van der Stigchel. Trans-saccadic memory after right parietal brain damage. *Cortex*, 120:284–297.

[87] J. D. Golomb and N. Kanwisher. Higher level visual cortex represents retinotopic, not spatiotopic, object location. *Cerebral Cortex*, 22(12):2794–2810, 2012.

[88] J. L. Gardner, E. P. Merriam, J. A. Movshon, and D. J. Heeger. Maps of visual space in human occipital cortex are retinotopic, not spatiotopic. *Journal of Neuroscience*, 28(15):3988–3999, 2008.

[89] X. Chen, G. C. DeAngelis, and D. E. Angelaki. Flexible egocentric and allocentric representations of heading signals in parietal cortex. *Proceedings of the National Academy of Sciences*, 115(14):E3305–E3312, 2018.

[90] B. E. Stein and T. R. Stanford. Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, 9(4):255–266.

[91] B. Mimica, B. A. Dunn, T. Tombaz, V. P. T. N. C. Srikanth Bojja, and J. R. Whitlock. Efficient cortical coding of 3d posture in freely behaving rats. *Science*, 362(6414):584–589.

[92] F. Bremmer, J. R. Duhamel, S. Ben Hamed, and W. Graf. Heading encoding in the macaque ventral intraparietal area (VIP). *European Journal of Neuroscience*, 16(8):1554–1568, 2002.

[93] J. R. Whitlock, R. J. Sutherland, M. P. Witter, M. B. Moser, and E. I. Moser. Navigating from hippocampus to parietal cortex. *Proceedings of the National Academy of Sciences*, 105(39):14755–14762, 2008.

[94] G. M. Olsen, K. Hovde, H. Kondo, T. Sakshaug, H. H. Smme, J. R. Whitlock, and M. P. Witter. Organization of posterior parietalfrontal connections in the rat. *Frontiers in Systems Neuroscience*, 13.

[95] K. S. Rockland and G. W. Van Hoesen. Some temporal and parietal cortical connections converge in CA1 of the primate hippocampus. *Cerebral Cortex*, 9(3):232–237, 1999.

[96] S. D. Vann, J. P. Aggleton, and E. A. Maguire. What does the retrosplenial cortex do? *Nature Reviews Neuroscience*, 10(11):792–802, 2009.

[97] P. Byrne, S. Becker, and N. Burgess. Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychological review*, 114(2):340–75, 2007-04.

[98] A. Bicanski and N. Burgess. A neural-level model of spatial memory and imagery. *eLife*, 7, 2018.

[99] R. A. Epstein and C. I. Baker. Scene perception in the human brain. *Annual Review of Vision Science*, 5(1), 2019.

[100] N. Kanwisher, J. McDermott, and M. M. Chun. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, 17(11):4302–4311.

[101] D. Y. Tsao. A cortical region consisting entirely of face-selective cells. *Science*, 311(5761):670–674.

[102] K. Tanaka. Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, 19(1):109–139.

[103] D. D. Dilks, J. B. Julian, A. M. Paunov, and N. Kanwisher. The occipital place area is causally and selectively involved in scene perception. *Journal of Neuroscience*, 33(4):1331–1336.

[104] R. Epstein and N. Kanwisher. A cortical representation of the local visual environment. *Nature*, 392(6676):598–601, 1998.

[105] D. J. Kravitz, C. S. Peng, and C. I. Baker. Real-world scene representations in high-level visual cortex: It's the spaces more than the places. *Journal of Neuroscience*, 31(20):7322–7333, 2011.

[106] K. Ferrara and S. Park. Neural representation of scene boundaries. *Neuropsychologia*, 89:180–190, 2016.

[107] M. F. Bonner and R. A. Epstein. Coding of navigational affordances in the human visual system. *Proceedings of the National Academy of Sciences*, 114(18):4793–4798.

[108] A. Schindler and A. Bartels. Visual high-level regions respond to high-level stimulus content in the absence of low-level confounds. *NeuroImage*, 132:520–525.

[109] A. Bicanski and N. Burgess. Environmental anchoring of head direction in a computational model of retrosplenial cortex. *The Journal of Neuroscience*, 36(46):11601–11618, 2016.

[110] L. L. Chen, L. H. Lin, E. J. Green, C. A. Barnes, and B. L. McNaughton. Head-direction cells in the rat posterior cortex - i. anatomical distribution and behavioral modulation. *Experimental Brain Research*, 101(1):8–23, 1994.

[111] J. Cho and P. E. Sharp. Head direction, place, and movement correlates for cells in the rat retrosplenial cortex. *Behavioral Neuroscience*, 115(1):3–25, 2001.

[112] P. Y. Jacob, G. Casali, L. Spieser, H. Page, D. Overington, and K. Jeffery. An independent, landmark-dominated head-direction signal in dysgranular retrosplenial cortex. *Nature Neuroscience*, 20(2):173–175, 2017.

[113] J. P. Shine, J. P. ValdHerrera, M. Hegarty, and T. Wolbers. The human retrosplenial cortex and thalamus code head direction in a global reference frame. *Journal of Neuroscience*, 36:6371–6381, 2016.

[114] S. A. Marchette, L. K. Vass, J. Ryan, and R. A. Epstein. Anchoring the neural compass: Coding of local spatial reference frames in human medial parietal lobe. *Nature Neuroscience*, 17(11):1598–1606, 2014.

[115] C. Koch, T. A. Li, S. C.and Polk, and N. W. Schuck. Effects of aging on the encoding of spatial direction in the human brain. *bioRxiv*.

[116] G. K. Aguirre and M. D'Esposito. Topographical disorientation: A synthesis and taxonomy. *Brain*, 122(9):1613–1628, 1999.

[117] J. R. Hinman, G. W. Chapman, and M. E. Hasselmo. Neuronal representation of environmental boundaries in egocentric coordinates. *Nature Communications*, 10(1).

[118] A. S. Alexander, L. C. Carstensen, J. R. Hinman, F. Raudies, G. W. Chapman, and M. E. Hasselmo. Egocentric boundary vector tuning of the retrosplenial cortex. *bioRxiv*.

[119] P. A. LaChance, T. P. Todd, and J. S. Taube. A sense of space in postrhinal cortex. *Science*, 365(6449):eaax4192.

[120] X. Gofman, G. Tocker, S. Weiss, C. N. Boccara, L. Lu, M. B. Moser, E. I. Moser, G. Morris, and D. Derdikman. Dissociation between postrhinal cortex and downstream parahippocampal regions in the representation of egocentric boundaries. *Current Biology*, 29(16):2751–2757.e4.

[121] Y. Kobayashi and D. G. Amaral. Macaque monkey retrosplenial cortex: III. cortical efferents. *The Journal of Comparative Neurology*, 502(5):810–833, 2007-06.

[122] Y. Kobayashi and D. G. Amaral. Macaque monkey retrosplenial cortex: II. cortical afferents. *Journal of Comparative Neurology*, 466(1):48–79, 2003.

[123] E. I. Moser and B. L. Moser, M. B.and McNaughton. Spatial representation in the hippocampal formation: A history. *Nature Neuroscience*, 20(11):1448–1464, 2017.

[124] T. Hartley, C. Lever, N. Burgess, and J. O'Keefe. Space in the brain: how the hippocampal formation supports spatial cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1635):20120510.

[125] J.R. Whitlock, G. Pfuhl, N. Dagslott, M. B. Moser, and E.I. Moser. Functional split between parietal and entorhinal cortices in the rat. *Neuron*, 73(4):789–802.

[126] D. A Nitz. Spaces within spaces: rat parietal cortex neurons register position across three reference frames. *Nature Neuroscience*, 15(10):1365–1367.

[127] M. Nau, J. B. Julian, and C. F. Doeller. How the brain's navigation system shapes our visual experience. *Trends in Cognitive Sciences*, 22(9):810–825, 2018.

[128] J. O'Keefe and J. Dostrovsky. The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34(1):171–175, 1971-11.

[129] J. O'Keefe and L. Nadel. *The hippocampus as a cognitive map*. Clarendon Press ; Oxford University Press.

[130] M. P. Witter, H. Kleven, and A. Kobro Flatmoen. Comparative contemplations on the hippocampus. *Brain, Behavior and Evolution*, 90(1):15–24.

[131] A. Rotenberg, M. Mayford, R. D. Hawkins, E. R. Kandel, and R. U. Muller. Mice expressing activated CaMKII lack low frequency LTP and do not form stable place cells in the CA1 region of the hippocampus. *Cell*, 87(7):1351–1361.

[132] N. Ulanovsky and C. F. Moss. Hippocampal cellular and network activity in freely moving echolocating bats. *Nature Neuroscience*, 10(2):224–233.

[133] M. M. Yartsev and N. Ulanovsky. Representation of three-dimensional space in the hippocampus of flying bats. *Science*, 340(6130):367–372.

[134] D. B. Omer, S. R. Maimon, L. Las, and N. Ulanovsky. Social place-cells in the bat hippocampus. *Science*, 359(6372):218–224.

[135] A. D. Ekstrom, M. J. Kahana, J. B. Caplan, T. A. Fields, E. A. Isham, E. L. Newman, and I. Fried. Cellular networks underlying human spatial navigation. *Nature*, 425:184–188, 2003.

[136] J. F. Miller, M. Neufang, A. Solway, A. Brandt, M. Trippel, I. Mader, S. Hefft, M. Merkow, S. M. Polyn, J. Jacobs, M. J. Kahana, and A. Schulze-Bonhage. Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. *Science*, 342(6162):1111–1114.

[137] E. C. Tolman. Cognitive maps in rats and men. *Psychological review*, 55:189, 1948.

[138] K. B. Kjelstrup, T. Solstad, V. H. Brun, T. Hafting, S. Leutgeb, M. P. Witter, E. I. Moser, and M. B. Moser. Finite scale of spatial representation in the hippocampus. *Science*, 321(5885):140–143.

[139] J. O'Keefe and D.H. Conway. Hippocampal place units in the freely moving rat: Why they fire where they fire. *Experimental Brain Research*, 31(4).

[140] G. J. Quirk, R. U. Muller, and J. L. Kubie. The firing of hippocampal place cells in the dark depends on the rat's recent experience. *The Journal of Neuroscience*, 10:2008–2017, 1990.

[141] R. U. Muller and J. L. Kubie. The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *J Neurosci*, 7:1951–1968, 1987.

[142] S. Leutgeb and J. K. Leutgeb. Remapping to discriminate contexts with hippocampal population codes. In *Space, Time and Memory in the Hippocampal Formation*, pages 227–251. Springer, 2014.

[143] E. Bostock, R. U. Muller, and J. L. Kubie. Experience-dependent modifications of hippocampal place cell firing. *Hippocampus*, 1:193–205, 1991.

[144] M. Fyhn, T. Hafting, A. Treves, M. B. Moser, and E. I. Moser. Hippocampal remapping and grid realignment in entorhinal cortex. *Nature*, 446(7132):190–194, 2007-03.

[145] L. L. Colgin, E. I. Moser, and M. B. Moser. Understanding memory through hippocampal remapping. *Trends in Neurosciences*, 31(9):469–477.

[146] D. Hassabis, C. Chu, G. Rees, N. Weiskopf, P. D. Molyneux, and E. A. Maguire. Decoding neuronal ensembles in the human hippocampus. *Current Biology*, 19(7):546–554.

[147] P. F. Rodriguez. Neural decoding of goal locations in spatial navigation in humans with fMRI. *Human Brain Mapping*, pages NA–NA.

[148] V. Sulpizio, G. Committeri, and G. Galati. Distributed cognitive maps reflecting real distances between places and views in the human brain. *Frontiers in Human Neuroscience*, 8.

[149] M. Kim, K. J. Jeffery, and E. A. Maguire. Multivoxel pattern analysis reveals 3d place information in the human hippocampus. *The Journal of Neuroscience*, 37(16):4270–4279.

[150] C. R. Nolan, J. M. G. Vromen, A. Cheung, and O. Baumann. Evidence against the detectability of a hippocampal place code using functional magnetic resonance imaging. *eneuro*, 5(4):ENEURO.0177–18.2018.

[151] T. Hafting, M. Fyhn, S. Molden, M. B. Moser, and E. I. Moser. Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052):801–806, 2005.

[152] M. P. Witter, T. P. Doan, B. Jacobsen, E. S. Nilssen, and S. Ohara. Architecture of the entorhinal cortex a review of entorhinal anatomy in rodents with some comparative notes. *Frontiers in Systems Neuroscience*, 11:1–12, 2017.

[153] H. Stensola, T. Stensola, T. Solstad, K. Frland, M. B. Moser, and E. I. Moser. The entorhinal grid map is discretized. *Nature*, 492(7427):72–78, 2012-12.

[154] D. Bush, C. Barry, D. Manson, and N. Burgess. Using grid cells for navigation. *Neuron*, 87(3):507–520, 2015-05.

[155] Y. Gu, S. Lewallen, A. A. Kinkhabwala, C. Domnisoru, K. Yoon, J. L. Gauthier, I. R. Fiete, and D. W. Tank. A map-like micro-organization of grid cells in the medial entorhinal cortex. *Cell*, 175(3):736–750.e30.

[156] Y. Burak and I. R. Fiete. Accurate path integration in continuous attractor network models of grid cells. *PLoS Computational Biology*, 5(2):e1000291.

[157] M. B. Moser, D. C. Rowland, and E. I. Moser. Place cells, grid cells, and memory. *Cold Spring Harbor perspectives in biology*, 7(2):a021808, 2015.

[158] C. F. Doeller, C. Barry, and N. Burgess. Evidence for grid cells in a human memory network. *Nature*, 463(7281):657–61, 2010.

[159] M. Stangl, J. Achtzehn, K. Huber, C. Dietrich, C. Tempelmann, and T. Wolbers. Compromised grid-cell-like representations in old age as a key mechanism to explain age-related navigational deficits. *Current Biology*, 0(0), 2018-03.

[160] M. Stangl, J. Shine, and T. Wolbers. The GridCAT: A toolbox for automated analysis of human grid cell codes in fMRI. *Frontiers in Neuroinformatics*, 11:1–12, 2017.

[161] A. J. Horner, J. A. Bisby, E. Zotow, D. Bush, and N. Burgess. Grid-like processing of imagined navigation. *Current Biology*, 26(6):842–847, 2016.

[162] J. L. S. Bellmund, L. Deuker, T. Navarro Schroeder, and C. F. Doeller. Grid-cell representations in mental simulation. *eLife*, 5:e17089, 2016-08.

[163] T. Navarro Schroeder, B. W. Towse, N. Burgess, C. Barry, and C. F. Doeller. Optimal decision making using grid cells under spatial uncertainty. *bioRxiv*, pages 1–43, 2017.

[164] L. Kunz, T. Navarro Schroeder, H. Lee, C. Montag, B. Lachmann, R. Sariyska, M. Reuter, R. Stirnberg, T. Stocker, P. C. Messing-Floeter, J. Fell, C. F. Doeller, and N. Axmacher. Reduced grid-cell-like representations in adults at genetic risk for alzheimer's disease. *Science*, 350(6259):430–433, 2015-10.

[165] L. Kunz, S. Maidenbaum, D. Chen, L. Wang, J. Jacobs, and N. Axmacher. Mesoscopic neural representations in spatial navigation. *Trends in Cognitive Sciences*, 23(7):615–630, 2019.

[166] Q. He and T. I. Brown. Environmental barriers disrupt grid-like representations in humans during navigation. *Current Biology*, 29(16):2718–2722.e3.

[167] S. Maidenbaum, J. Miller, J. M. Stein, and J. Jacobs. Grid-like hexadirectional modulation of human entorhinal theta oscillations. *Proceedings of the National Academy of Sciences*, 115(42):10798–10803.

[168] A. O. Constantinescu, J. X. O'Reilly, and T. E. J. Behrens. Organizing conceptual knowledge in humans with a gridlike code. *Science*, 352(6292):1464–1468, 2016-06.

[169] X. Bao, E. Gjorgieva, L. K. Shanahan, J. D. Howard, T. Kahnt, and J. A. Gottfried. Grid-like neural representations support olfactory navigation of a two-dimensional odor space. *Neuron*, 102(5):1066–1075.e5.

[170] M. Nau, T. Navarro Schroeder, J. L. S. Bellmund, and C. F. Doeller. Hexadirectional coding of visual space in human entorhinal cortex. *Nature Neuroscience*, 21(2):188–190, 2018-02.

[171] J. B. Julian, A. T. Keinath, G. Frazzetta, and R. A. Epstein. Human entorhinal cortex represents visual space using a boundary-anchored grid. *Nature Neuroscience*, 21(2):191–194, 2018.

[172] T. Staudigl, M. Leszczynski, J. Jacobs, S. A. Sheth, C. E. Schroeder, O. Jensen, and C. F. Doeller. Hexadirectional modulation of high-frequency electrophysiological

activity in the human anterior medial temporal lobe maps visual space. *Current Biology*, 28(20):3325–3329.e4, 2018.

[173] J. Jacobs, C. T. Weidemann, J. F. Miller, A. Solway, J. F. Burke, X. X. Wei, N. Suthana, M. R. Sperling, A. D. Sharan, I. Fried, and M. J. Kahana. Direct recordings of grid-like neuronal activity in human spatial navigation. *Nature Neuroscience*, 16(9):1188–1190, 2013-08.

[174] Z. Nadasdy, T. P. Nguyen, . Toeroek, J. Y. Shen, D. E. Briggs, P. N. Modur, and R. J. Buchanan. Context-dependent spatially periodic activity in the human entorhinal cortex. *Proceedings of the National Academy of Sciences*, page 201701352, 2017-04.

[175] K. E. Cullen and J. S. Taube. Our sense of direction: Progress, controversies and challenges. *Nature Neuroscience*, 20(11):1465–1473, 2017.

[176] J. S. Taube. Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *The Journal of Neuroscience*, 15(1):70–86, 1995.

[177] J. S. Taube, R. U. Muller, and J. B. Ranck. Head-direction cells recorded from the postsubiculum in freely moving rats. i. description and quantitative analysis. *The Journal of Neuroscience*, 10(2):420–435, 1990.

[178] O. Kornienko, P. Latuske, M. Bassler, L. Kohler, and K. Allen. Non-rhythmic head-direction cells in the parahippocampal region are not constrained by attractor network dynamics. *eLife*, 7, 2018.

[179] L. M. Giocomo, T. Stensola, T. Bonnevie, T. Van Cauter, M. B. Moser, and E. I. Moser. Topography of head direction cells in medial entorhinal cortex. *Current Biology*, 24(3):252–262, 2014.

[180] F. Sargolini, M. Fyhn, T. Hafting, B. L. McNaughton, M. P. Witter, M. B. Moser, and E. I. Moser. Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science*, 312(5774):758–762, 2006.

[181] R. G. Robertson, E. T. Rolls, P. Georges-Frans, and S. Panzeri. Head direction cells in the primate pre-subiculum. *Hippocampus*, 9(3):206–219, 1999.

[182] S. J. Mizumori and J. D. Williams. Directionally selective mnemonic properties of neurons in the lateral dorsal nucleus of the thalamus of rats. *The Journal of neuroscience*, 13(9):4015–4028, 1993.

[183] J. P. Goodridge, P. A. Dudchenko, K. A. Worboys, E. J. Golob, and J. S. Taube. Cue control and head direction cells. *Behavioral Neuroscience*, 112(4):749–761.

[184] O. Baumann and J. B. Mattingley. Medial parietal cortex encodes perceived heading direction in humans. *Journal of Neuroscience*, 30(39):12897–12901, 2010.

[185] M. J. Chadwick, A. E. J. Jolly, D. P. Amos, D. Hassabis, and H. J. Spiers. A goal direction signal in the human entorhinal/subicular region. *Current Biology*, 25(1):87–92, 2015.

[186] I. Indovina, V. Maffei, K. Pauwels, E. Macaluso, G. A. Orban, and F. Lacquaniti. Simulated self-motion in a visual gravity field: Sensitivity to vertical and horizontal heading in the human brain. *NeuroImage*, 71:114–124, 2013.

[187] M. Kim and E. A. Maguire. Encoding of 3d head direction information in the human brain. *Hippocampus*, 29(7):619–629, 2019.

[188] L. K. Vass and R. A. Epstein. Abstract representations of location and facing direction in the human brain. *Journal of Neuroscience*, 33(14):6133–6142, 2013-04.

[189] L. K. Vass and R. A. Epstein. Common neural representations for visually guided

reorientation and spatial imagery. *Cerebral cortex*, 27(2):1457–1471, 2017-01.

[190] J. P. Shine, J. P. ValdHerrera, C. Tempelmann, and T. Wolbers. Evidence for allocentric boundary and goal direction information in the human entorhinal cortex and subiculum. *Nature Communications*, 10(1).

[191] J. Jacobs, J. Miller, S. A. Lee, T. Coffey, A. J. Watrous, M. R. Sperling, A. Sharan, G. Worrell, B. Berry, B. Lega, B. C. Jobst, K. Davis, R. E. Gross, S. A. Sheth, Y. Ezzyat, S. R. Das, J. Stein, R. Gorniak, M. J. Kahana, and D. S. Rizzuto. Direct electrical stimulation of the human entorhinal region and hippocampus impairs memory. *Neuron*, 92(5):983–990, 2016-12.

[192] P. Gourtzelidis, C. Tzagarakis, S. M. Lewis, D. A. Crowe, E. Auerbach, T. A. Jerde, K. Ugurbil, and A. P. Georgopoulos. Mental maze solving: Directional fMRI tuning and population coding in the superior parietal lobule. *Experimental Brain Research*, 165(3):273–282, 2005.

[193] A. Schindler and A. Bartels. Parietal cortex codes for egocentric space beyond the field of view. *Current Biology*, 23(2):177–182, 2013.

[194] T. Solstad, C. N. Boccara, E. Kropff, M. B. Moser, and E. I. Moser. Representation of geometric borders in the entorhinal cortex. *Science*, 322:1865–1868, 2008.

[195] F. Savelli, D. Yoganarasimha, and J. J. Knierim. Influence of boundary removal on the spatial representations of the medial entorhinal cortex. *Hippocampus*, 18(12):1270–1282.

[196] C. Lever, S. Burton, A. Jeewajee, J. O'Keefe, and N. Burgess. Boundary vector cells in the subiculum of the hippocampal formation. *The Journal of Neuroscience*, 29:9771–9777, 2009.

[197] A. Hydal, E. R. Skyten, S. O. Andersson, M. B. Moser, and E. I. Moser. Object-vector coding in the medial entorhinal cortex. *Nature*, 568(7752):400–404.

[198] E. Kropff, J. E. Carmichael, M. B. Moser, and E. I. Moser. Speed cells in the medial entorhinal cortex. *Nature*, 523(7561):419–424.

[199] W. E. Skaggs, J. J. Knierim, H. S. Kudrimoti, and B. L. McNaughton. A model of the neural basis of the rat's sense of direction. *Advances in neural information processing systems*, 7:173–80, 1995.

[200] K. Zhang. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *The Journal of Neuroscience*, 16(6):2112–2126.

[201] S. I. Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27(2):77–87.

[202] L.M. Giocomo, M. B. Moser, and E.I. Moser. Computational models of grid cells. *Neuron*, 71(4):589–603.

[203] S. M. Stringer, E. T. Rolls, T. P. Trappenberg, and I. E. T. de Araujo. Self-organizing continuous attractor networks and path integration: two-dimensional models of place cells. *Network: Computation in Neural Systems*, 13(4):429–446.

[204] A. Samsonovich and B. L. McNaughton. Path integration and cognitive mapping in a continuous attractor neural network model. *The Journal of Neuroscience*, 17(15):5900–5920.

[205] A. D. Redish, A. N. Elga, and D. S. Touretzky. A coupled attractor model of the rodent head direction system. *Network: Computation in Neural Systems*, 7(4):671–685.

[206] A. Peyrache, M. M. Lacroix, P. C. Petersen, and G. Buzs. Internally organized mechanisms of the head direction sense. *Nature Neuroscience*, 18(4):569–575.

[207] R. Chaudhuri, B. Ger, B. Pandey, A. Peyrache, and I. Fiete. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. *Nature Neuroscience*, 22(9):1512–1520.

[208] N. Almog, G. Tocker, T. Bonnevie, E. I. Moser, M. B. Moser, and D. Derdikman. During hippocampal inactivation, grid cells maintain their synchrony, even when the grid pattern is lost. *bioRxiv*.

[209] R. J. Gardner, L. Lu, T. Wernle, M. B. Moser, and E. I. Moser. Correlation structure of grid cells is preserved during sleep. *Nature Neuroscience*, 22(4):598–608, 2019-10.

[210] S. G. Trettel, J. B. Trimper, E. Hwaun, I. R. Fiete, and L. L. Colgin. Grid cell co-activity patterns during sleep reflect spatial overlap of grid fields during active behaviors. *Nature Neuroscience*, 22(4):609–617, 2019-10.

[211] J. D. Seelig and V. Jayaraman. Neural dynamics for landmark orientation and angular path integration. *Nature*, 521(7551):186–191.

[212] S. S. Kim, H. Rouault, S. Druckmann, and V. Jayaraman. Ring attractor dynamics in the drosophila central brain. *Science*, 356(6340):849–853.

[213] A. D. Ekstrom. Why vision is important to how we navigate. *Hippocampus*, 25(6):731–735, 2015-06.

[214] M. Carandini and A. K. Churchland. Probing perceptual decisions in rodents. *Nature Neuroscience*, 16(7):824–831.

[215] L. Krubitzer, K. L. Campi, and D. F. Cooke. All rodents are not the same: A modern synthesis of cortical organization. *Brain, Behavior and Evolution*, 78(1):51–93.

[216] E. T. Rolls and S. Wirth. Spatial representations in the primate hippocampus, and their functions in memory and navigation. *Progress in Neurobiology*, 171:90–113, 2018.

[217] M. L. R. Meister and E. A. Buffalo. Getting directions from the hippocampus: The neural connection between looking and memory. *Neurobiology of Learning and Memory*, 134:135–144, 2016.

[218] N. J. Killian and E. A. Buffalo. Grid cells map the visual world. *Nature Neuroscience*, 21(2):161–162.

[219] E. T. Rolls, R. G. Robertson, and P. Georges-Frans. Spatial view cells in the primate hippocampus. *European Journal of Neuroscience*, 9(8):1789–1794, 1997.

[220] P. Georges-Frans, E. T. Rolls, and R. G. Robertson. Spatial view cells in the primate hippocampus: Allocentric view not head direction or eye position or place. *Cerebral Cortex*, 9(3):197–212, 1999-04.

[221] E. T. Rolls. Spatial view cells and the representation of place in the primate hippocampus. *Hippocampus*, 9(4):467–480, 1999.

[222] R. G. Robertson, E. T. Rolls, and P. Georges-Frans. Spatial view cells in the primate hippocampus: effects of removal of view details. *Journal of Neurophysiology*, 79(3):1145–1156, 1998.

[223] S. Wirth, P. Baraduc, A. Plant. Pin, and J. R. Duhamel. Gaze-informed, task-situated representation of space in primate hippocampus during virtual navigation. *PLoS biology*, 15(2):e2001045, 2017.

[224] T. Danjo, T. Toyoizumi, and S. Fujisawa. Spatial representations of self and other in the hippocampus. *Science*, 359(6372):213–218.

[225] N. J. Killian, M. J. Jutras, and E. A. Buffalo. A map of visual space in the primate entorhinal cortex. *Nature*, 491(7426):761–764, 2012.

[226] M. L. R. Meister and E. A. Buffalo. Neurons in primate entorhinal cortex represent gaze position in multiple spatial reference frames. *The Journal of Neuroscience*, 38(10):2430–2441, 2018.

[227] N. Wilming, P. Koenig, S. Koenig, and E. A. Buffalo. Entorhinal cortex receptive fields are modulated by spatial attention, even without movement. *eLife*, 7:e31745, 2018.

[228] A. Bicanski and N. Burgess. A computational model of visual recognition memory via grid cells. *Current Biology*, 29(6):979–990.e4, 2019.

[229] N. J. Killian, S. M. Potter, and E. A. Buffalo. Saccade direction encoding in the primate entorhinal cortex during visual exploration. *Proceedings of the National Academy of Sciences of the United States of America*, 112(51):15743–8, 2015-12.

[230] P. Baraduc, J. R. Duhamel, and S. Wirth. Schema cells in the macaque hippocampus. *Science*, 363(6427):635.

[231] J. Laurens and D. E. Angelaki. The brain compass: A perspective on how self-motion updates the head direction cell attractor. *Neuron*, 97(2):275–289.

[232] T. Evans, A. Bicanski, D. Bush, and N. Burgess. How environment and self-motion combine in neural representations of space: Environment and self-motion in neural representations of space. *The Journal of Physiology*, 594(22):6535–6546.

[233] J. B. Julian, A. T. Keinath, S. A. Marchette, and R. A. Epstein. The neurocognitive basis of spatial reorientation. *Current Biology*, 28(17):R1059–R1073.

[234] R. M. Grieves, S. Jedidi-Ayoub, K. Mishchanchuk, A. Liu, S. Renaudineau, and K. J. Jeffery. The place-cell representation of volumetric space in rats. *bioRxiv*.

[235] M. Kim and E. A. Maguire. Hippocampus, retrosplenial and parahippocampal cortices encode multicompartment 3d space in a hierarchical manner. *Cerebral Cortex*, 28(5):1898–1909, 2018.

[236] A. Finkelstein, D. Derdikman, A. Rubin, J. N. Foerster, L. Las, and N. Ulanovsky. Three-dimensional head-direction coding in the bat brain. *Nature*, 517(7533):159–164.

[237] D. Aronov, R. Nevers, and D. W. Tank. Mapping of a non-spatial dimension by the hippocampal-entorhinal circuit. *Nature*, 543:719–722, 2017.

[238] J. L. S. Bellmund, P. Gaerdenfors, E. I. Moser, and C. F. Doeller. Navigating cognition: Spatial codes for human thinking. *Science*, 362(6415):eaat6766, 2018.

[239] T. E. J. Behrens, T. H. Muller, J. C. R. Whittington, S. Mark, A. B. Baram, K. L. Stachenfeld, and Z. Kurth-Nelson. What is a cognitive map? organizing knowledge for flexible behavior. *Neuron*, 100(2):490–509, 2018.

[240] C. Rennsta, J. E. Lisman, and P. F. M. J. Verschure. A signature of attractor dynamics in the CA3 region of the hippocampus. *PLoS Computational Biology*, 10(5):e1003641.

[241] M. Tsodyks. Attractor neural network models of spatial maps in hippocampus. *Hippocampus*, 9(4):481–489.

[242] T. J. Wills. Attractor dynamics in the hippocampal representation of the local environment. *Science*, 308(5723):873–876.

[243] B. Steemers, A. Vicente-Grabovetsky, C. Barry, P. Smulders, T. Navarro Schroeder, N. Burgess, and C.F. Doeller. Hippocampal attractor dynamics predict memory-based decision making. *Current Biology*, 26(13):1750–1757.

[244] E. T. Rolls. An attractor network in the hippocampus: Theory and neurophysiology. *Learning & Memory*, 14(11):714–731.

[245] E. T. Rolls. The mechanisms for pattern completion and pattern separation in the hippocampus. *Frontiers in Systems Neuroscience*, 7.

[246] K. J. Jeffery. Place cells, grid cells, attractors, and remapping. *Neural Plasticity*, 2011:1–11.

[247] M. Nau, A. Schindler, and A. Bartels. Real-motion signals in human early visual cortex. *NeuroImage*, 175:379–387, 2018-07.

[248] M. Nau, T. Navarro Schroeder, M. Frey, and C. F. Doeller. Behavior-dependent directional tuning in the human visual-navigation network. *bioRxiv*.

[249] N. K. Logothetis. What we can do and what we cannot do with fMRI. *Nature*, 453(7197):869–78, 2008-06.

[250] C. Galletti, P.P. Battaglini, and G. Aicardi. 'real-motion' cells in visual area v2 of behaving macaque monkeys. *Experimental Brain Research*, 69(2), 1988-01.

[251] U. J. Ilg, S. Schumann, and P. Thier. Posterior parietal cortex neurons encode target motion in world-centered coordinates. *Neuron*, 43(1):145–51, 2004-07.

[252] C. Galletti, D. F. Kutz, M. Gamberini, R. Breveglieri, and P. Fattori. Role of the medial parieto-occipital cortex in the control of reaching and grasping movements. *Experimental Brain Research*, 153(2):158–170, 2003.

[253] V. Cardin and A. T. Smith. Sensitivity of human visual and vestibular cortical regions to egomotion-compatible visual stimulation. *Cerebral Cortex*, 20(8):1964–1973, 2010-08.

[254] M. B. Wall and A. T. Smith. The representation of egomotion in the human brain. *Current biology : CB*, 18(3):191–4, 2008-03.

[255] C. S. Konen and S. Kastner. Representation of eye movements and stimulus motion in topographically organized areas of human posterior parietal cortex. *The Journal of neuroscience*, 28(33):8361–8375, 2008.

[256] D. Chawla, J. Phillips, C. Buechel, R. Edwards, and K. J. Friston. Speed-dependent motion-sensitive responses in v5: an fMRI study. *NeuroImage*, 7(2):86–96, 1998-02.

[257] D. Chawla, C. Buechel, R. Edwards, A. Howseman, O. Josephs, J. Ashburner, and K. J. Friston. Speed-dependent responses in v5: A replication study. *NeuroImage*, 9(5):508–15, 1999-05.

[258] S. Pitzalis, P. Fattori, and C. Galletti. The functional role of the medial motion area v6. *Frontiers in behavioral neuroscience*, 6:91, 2012.

[259] S. Moeller, E. Yacoub, C. A. Olman, E. Auerbach, J. Strupp, N. Harel, and K. Ugurbil. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 63(5):1144–53, 2010-05.

[260] A. E. Desjardins, K. A. Kiehl, and P. F. Liddle. Removal of confounding effects of global signal in functional MRI analyses. *NeuroImage*, 13(4):751–758, 2001.

[261] M. Brett, J. L. Anton, R. Valabregue, and J. B. Poline. Region of interest analysis using an SPM toolbox. *NeuroImage*, 16:497, 2002.

[262] A. C. Huk, R. F. Dougherty, and D. J. Heeger. Retinotopy and functional subdivision of human areas MT and MST. *The Journal of neuroscience*, 22(16):7195–205, 2002-08.

[263] S. O. Dumoulin, R. G. Bittar, N. J. Kabani, C. L. Baker, G. Le Goualher, G. Bruce Pike, and A. C. Evans. A new anatomical landmark for reliable identification of

human area v5/MT: a quantitative analysis of sulcal patterning. *Cerebral cortex*, 10(5):454–63, 2000-05.

[264] H. Kolster, R. Peeters, and G. A. Orban. The retinotopic organization of the human middle temporal area MT/v5 and its cortical neighbors. *The Journal of neuroscience*, 30(29):9801–20, 2010.

[265] K. Nelissen. Charting the lower superior temporal region, a new motion-sensitive region in monkey superior temporal sulcus. *Journal of Neuroscience*, 26(22):5929–5947, 2006.

[266] S. Pitzalis, P. Fattori, and C. Galletti. The human cortical areas v6 and v6a. *Visual Neuroscience*, 32:E007, 2015.

[267] S. Pitzalis, M. I. Sereno, G. Committeri, P. Fattori, G. Galati, A. Tosoni, and C. Galletti. The human homologue of macaque area v6a. *NeuroImage*, 82:517–530, 2013.

[268] A. M. Dale, B. Fischl, and M. I. Sereno. Cortical surface-based analysis. *NeuroImage*, 9(2):179–194, 1999-02.

[269] C. Galletti and P. P. Battaglini. Gaze-dependent visual neurons in area v3a of monkey prestriate cortex. *The Journal of neuroscience*, 9(4):1112–25, 1989-04.

[270] V. Cardin and A. T. Smith. Sensitivity of human visual cortical area v6 to stereoscopic depth gradients associated with self-motion. *Journal of neurophysiology*, 106(3):1240–1249, 2011.

[271] S. Pitzalis, M. I. Sereno, G. Committeri, P. Fattori, G. Galati, F. Patria, and C. Galletti. Human v6: The medial motion area. *Cerebral Cortex*, 20(2):411–424, 2010.

[272] D. J. Quinlan and J. C. Culham. fMRI reveals a preference for near viewing in the human parieto-occipital cortex. *NeuroImage*, 36(1):167–187, 2007.

[273] F. Bremmer, A. Schlack, N. J. Shah, O. Zafiris, M. Kubischik, KP. Hoffmann, K. Zilles, and G. R. Fink. Polymodal motion processing in posterior parietal and premotor cortex. *Neuron*, 29(1):287–296, 2001.

[274] J. R. Duhamel, C. L. Colby, and M. E. Goldberg. Ventral intraparietal area of the macaque: congruent visual and somatic response properties. *Journal of neurophysiology*, 79(1):126–136, 1998.

[275] T. Zhang, H. W. Heuer, and K. H. Britten. Parietal area VIP neuronal responses to heading stimuli are encoded in head-centered coordinates. *Neuron*, 42(6):993–1001, 2004-06.

[276] C. L. Colby, J. R. Duhamel, and M. E. Goldberg. Ventral intraparietal area of the macaque: anatomic location and visual response properties. *Journal of neurophysiology*, 69(3):902–14, 1993-03.

[277] S. F. Gabel, H. Misslisch, C. C. A. M. Gielen, and J. Duysens. Responses of neurons in area VIP to self-induced and external visual motion. *Experimental brain research*, 147(4):520–8, 2002-12.

[278] T. Zhang and K. H. Britten. Parietal area VIP causally influences heading perception during pursuit eye movements. *The Journal of neuroscience*, 31(7):2569–2575, 2011.

[279] J. H. Maunsell and D. C. van Essen. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *The Journal of neuroscience*, 3(12):2563–2586, 1983.

[280] A. Schlack, S. J. Sterbing-D'Angelo, K. Hartung, K. P. Hoffmann, and F. Bremmer. Multisensory space representations in the macaque ventral intraparietal area. *The Journal of Neuroscience*, 25(18):4616–4625, 2005.

[281] C. Galletti, M. Gamberini, D. F. Kutz, P. Fattori, G. Luppino, and M. Matelli. The

cortical connections of area v6: An occipito-parietal network processing visual information. *European Journal of Neuroscience*, 13(8):1572–1588, 2001.

[282] J. A. Movshon. The velocity tuning of single units in cat striate cortex. *The Journal of physiology*, 249(3):445–468, 1975.

[283] G. A. Orban, H. Kennedy, and J. Bullier. Velocity sensitivity and direction selectivity of neurons in areas v1 and v2 of the monkey: influence of eccentricity. *Journal of neurophysiology*, 56(2):462–480, 1986.

[284] R. Goebel, D. Khorram-Sefat, L. Muckli, H. Hacker, and W. Singer. The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *European Journal of Neuroscience*, 10(5):1563–1573, 1998.

[285] K. Seymour, C. W. G. Clifford, N. K. Logothetis, and A. Bartels. The coding of color, motion, and their conjunction in the human visual cortex. *Current Biology*, 19(3):177–183, 2009.

[286] A. T. Smith, M. W. Greenlee, K. D. Singh, F. M. Kraemer, and J. Hennig. The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). *The Journal of neuroscience*, 18(10):3816–3830, 1998.

[287] S. Zeki, J. D. Watson, C. J. Lueck, K. J. Friston, C. Kennard, and R. S. Frackowiak. A direct demonstration of functional specialization in human visual cortex. *The Journal of neuroscience*, 11(3):641–9, 1991-03.

[288] X. G. Troncoso, M. B. McCamy, A. N. Jazi, J. Cui, J. Otero-Millan, S. L. Macknik, F. M. Costela, and S. Martinez-Conde. V1 neurons respond differently to object motion versus motion from eye movements. *Nature communications*, 6:8114, 2015.

[289] B. E. Russ, T. Kaneko, K. S. Saleem, R. A Berman, and D. A. Leopold. Distinct fMRI responses to self-induced versus stimulus motion during free viewing in the macaque. *The Journal of Neuroscience*, 36(37):9580–9, 2016.

[290] P. Li, S. Zhu, M. Chen, C. Han, H. Xu, J. Hu, Y. Fang, and H. D. Lu. A motion direction preference map in monkey v4. *Neuron*, 78(2):376–388, 2013.

[291] P. Stoerig. Blindsight, conscious vision, and the role of primary visual cortex. *Progress in brain research*, 155:217–234, 2006.

[292] S. Zeki and D. H. Ffytche. The riddoch syndrome: Insights into the neurobiology of conscious vision. *Brain*, 121(1):25–45, 1998.

[293] T. Haarmeier, P. Thier, M. Repnow, and D. Petersen. False perception of motion in a patient who cannot compensate for eye movements. *Nature*, 389(6653):849–852, 1997.

[294] J. A. Movshon, I. D. Thompson, and D. J. Tolhurst. Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex. *The Journal of physiology*, 283(1978):101–20, 1978-10.

[295] N. Qian, R. A. Andersen, and E. H. Adelson. Transparent motion perception as detection of unbalanced motion signals. III. modeling. *The Journal of neuroscience*, 14(12):7381–7392, 1994.

[296] J. M. Hup. C. James, P. Girard, S. G Lomber, B. R. Payne, and J. Bullier. Feedback connections act on the early part of the responses in monkey visual cortex. *Journal of neurophysiology*, 85(1):134–145, 2001.

[297] V. A. F. Lamme and P. R. Roelfsema. The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, 23(11):571–579, 2000.

[298] L. Muckli, A. Kohler, N. Kriegeskorte, and W. Singer. Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS Biology*, 3(8), 2005.

[299] R. P. N. Rao and D. H. Ballard. Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1):79–87, 1999.

[300] J. Silvanto, A. Cowey, N. Lavie, and V. Walsh. Striate cortex (v1) activity gates awareness of motion. *Nature Neuroscience*, 8(2):143–144, 2005.

[301] A. Pascual-Leone and V. Walsh. Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science*, 292(5516):510–512, 2001.

[302] A. Bartels, N. K. Logothetis, and K. Moutoussis. fMRI and its interpretations: an illustration on directional selectivity in area v5/MT. *Trends in neurosciences*, 31(9):444–53, 2008-09.

[303] N. K. Logothetis. The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 357(1424):1003–37, 2002-08.

[304] M. Bannert and A. Bartels. Decoding the yellow of a gray banana. *Current Biology*, 23(22):2268–2272, 2013.

[305] I. Sperandio, P. A. Chouinard, and M. A. Goodale. Retinotopic activity in v1 reflects the perceived and not the retinal size of an afterimage. *Nature Neuroscience*, 15(4):540–542, 2012.

[306] M. A. Williams, C. I. Baker, H. P. Op De Beeck, W. Mok Shim, S. Dang, C. Triantafyllou, and N. Kanwisher. Feedback of visual object information to foveal retinotopic cortex. *Nature Neuroscience*, 11(12):1439–1445, 2008.

[307] D. J. McKeefry, J. D. Watson, R. S. Frackowiak, K. Fong, and S. Zeki. The activity in human areas v1/v2, v3, and v5 during the perception of coherent and incoherent motion. *NeuroImage*, 5(1):1–12, 1997.

[308] A. Schindler and A. Bartels. Connectivity reveals sources of predictive coding signals in early visual cortex during processing of visual optic flow. *Cerebral Cortex*, 27(5):2885–2893, 2017.

[309] J. G. Heys, K. V. Rangarajan, and D. A. Dombeck. The functional micro-organization of grid cells revealed by cellular-resolution imaging. *Neuron*, 84(5):1079–1090, 2014-12.

[310] J. Ross, M. C. Morrone, M. E. Goldberg, and D. C. Burr. Changes in visual perception at the time of saccades. *Trends in neurosciences*, 24(2):113–21, 2001-02.

[311] G. Chen, J. A. King, N. Burgess, and J. O'Keefe. How vision and movement combine in the hippocampal place code. *Proceedings of the National Academy of Sciences*, 110(1):378–383, 2013.

[312] E. M. Klier and D. E. Angelaki. Spatial updating and the maintenance of visual constancy. *Neuroscience*, 156(4):801–818, 2008.

[313] J. J. Summerfield, J. Lepsien, D. R. Gitelman, M. M. Mesulam, and A. C. Nobre. Orienting attention based on long-term memory experience. *Neuron*, 49(6):905–916, 2006.

[314] J. L. Voss, D. J. Bridge, N. J. Cohen, and J. A. Walker. A closer look at the hippocampus and memory. *Trends in Cognitive Sciences*, 21(8):577–588, 2017.

[315] R. A. Epstein, E. Z. Patai, J. B. Julian, and H. J. Spiers. The cognitive map in humans: Spatial navigation and beyond. *Nature Neuroscience*, 20(11):1504–1513, 2017-11.

[316] B. J. Clark, C. M. Simmons, L. E. Berkowitz, and A. A. Wilber. The retrosplenial-parietal network and reference frame coordination

for spatial navigation. *PsyArXiv Preprints*, pages 1–35, 2018.

[317] J. Jacobs, M. J. Kahana, A. D. Ekstrom, M. V. Mollison, and I. Fried. A sense of direction in human entorhinal cortex. *Proceedings of the National Academy of Sciences*, 107(14):6487–6492, 2010.

[318] J. Shine, J. P. Valdes-Herrera, C. Tempelmann, and T. Wolbers. Evidence for allocentric boundary and goal direction information in the human entorhinal cortex and subiculum. *bioRxiv*, page 466789, 2018.

[319] T. E. Nichols and A. P. Holmes. Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, 15(1):1–25, 2002.

[320] E. H. Silson, A. W. Gilmore, S. E. Kalinowski, A. Steel, A. Kidder, A. Martin, and C. I. Baker. A posterioranterior distinction between scene perception and scene construction in human medial parietal cortex. *The Journal of Neuroscience*, 39(4):705–717, 2019.

[321] A. Maass, D. Berron, L. A. Libby, C. Ranganath, and E. Duezel. Functional subregions of the human entorhinal cortex. *eLife*, 4:1–20, 2015.

[322] T. Navarro Schroeder, K. V. Haak, N. I. Z. Jimenez, C. F. Beckmann, and C. F. Doeller. Functional topography of the human entorhinal cortex. *eLife*, 4:1–17, 2015.

[323] A. S. Mitchell, R. Czajkowski, N. Zhang, K. Jeffery, and A. J. D. Nelson. Retrosplenial cortex and its role in spatial cognition. *Brain and Neuroscience Advances*, 2:239821281875709, 2018.

[324] T. M. Mitchell, S. V. Shinkareva, A. Carlson, K. M. Chang, V. L. Malave, R. A. Mason, and M. A. Just. Predicting human brain activity associated with the meanings of nouns. *Science*, 320(5880):1191–1195, 2008.

[325] K. N. Kay, T. Naselaris, R. J. Prenger, and J. L. Gallant. Identifying natural images from human brain activity. *Nature*, 452(7185):352–355, 2008.

[326] B. Thirion, E. Duchesnay, E. Hubbard, J. Dubois, Jean B. Poline, D. Lebihan, and S. Dehaene. Inverse retinotopy: Inferring the visual content of images from brain activation patterns. *NeuroImage*, 33(4):1104–1116, 2006.

[327] R. L. Rademaker, C. Chunharas, and J. T. Serences. Coexisting representations of sensory and mnemonic information in human visual cortex. *Nature Neuroscience*, 2019.

[328] W. N. Butler, K. S. Smith, M. A. A. van der Meer, and J. S. Taube. The head-direction signal plays a functional role as a neural compass during navigation. *Current Biology*, 27(9):1259–1267, 2017.

[329] P. A. Dudchenko, E. R. Wood, and A. Smith. A new perspective on the head direction cell system and spatial behavior. *Neuroscience & Biobehavioral Reviews*, 2019.

[330] M. A. A. van der Meer, Z. Richmond, R. M. Braga, E. R. Wood, and P. A. Dudchenko. Evidence for the use of an internal sense of direction in homing. *Behavioral Neuroscience*, 124(1):164–169, 2010.

[331] S. Valerio and J. S. Taube. Path integration: How the head direction signal maintains and corrects spatial orientation. *Nature Neuroscience*, 15(10):1445–1453, 2012-10.

[332] S. Weiss and D. Derdikman. Role of the head-direction signal in spatial tasks: when and how does it guide behavior? *Journal of Neurophysiology*, 120(1):78–87, 2018-07.

[333] S. Weiss, G. Talhami, X. Gofman-Regev, S. Rapoport, D. Eilam, and D. Derdikman. Consistency of spatial representations in rat entorhinal cortex predicts performance

in a reorientation task. *Current Biology*, 27(23):3658–3665.e4, 2017-12.

[334] A. Ayaz, A. B. Saleem, M. L. Schoelvinck, and M. Carandini. Locomotion controls spatial integration in mouse visual cortex. *Current Biology*, 23(10):890–894, 2013.

[335] A. B. Saleem, A. I. Ayaz, K. J. Jeffery, K. D. Harris, and M. Carandini. Integration of visual motion and locomotion in mouse visual cortex. *Nature Neuroscience*, 16(12):1864–1869, 2013.

[336] J. Fournier, A. B. Saleem, E. M. Diamanti, M. J. Wells, K. D. Harris, and M. Carandini. Modulation of visual cortex by hippocampal signals. *bioRxiv*, page 586917, 2019.

[337] A. B. Saleem, E. M. Diamanti, J. Fournier, K. D. Harris, and M. Carandini. Coherent encoding of subjective spatial position in visual cortex and hippocampus. *Nature*, 562(7725):124–127, 2018.

[338] S. E. Bosch, J. F. M. Jehee, G. Fernandez, and C. F. Doeller. Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *Journal of Neuroscience*, 34(22):7493–7500, 2014-05.

[339] N. C. Hindy, F. Y. Ng, and N. B. Turk-Browne. Linking pattern completion in the hippocampus to predictive coding in visual cortex. *Nature Neuroscience*, 19(5):665–667, 2016.

[340] P. Kok and N. B. Turk-Browne. Associative prediction of visual shape in the hippocampus. *The Journal of Neuroscience*, 38(31):6888–6899, 2018.

[341] J. S. Taube. The head direction signal: Origins and sensory-motor integration. *Annual Review of Neuroscience*, 30(1):181–207, 2007.

[342] L. Acharya, Z. M. Aghajan, C. Vuong, J. J. Moore, and M. R. Mehta. Causal influence of visual cues on hippocampal directional selectivity. *Cell*, 164(1):197–207, 2016.

[343] S. Coletta, M. Frey, K. Nasr, P. Preston-Ferrer, and A. Burgalossi. Testing the efficacy of single-cell stimulation in biasing presubicular head direction activity. *The Journal of Neuroscience*, 38(13):3287–3302, 2018.

[344] K. Cheng. A purely geometric module in the rat's spatial representation. *Cognition*, 23(2):149–178, 1986.

[345] J. B. Julian, A. T. Keinath, I. A. Muzzio, and R. A. Epstein. Place recognition and heading retrieval are mediated by dissociable cognitive systems in mice. *Proceedings of the National Academy of Sciences*, 112(20):6503–6508, 2015.

[346] T. Wolbers. Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *Journal of Neuroscience*, 25(13):3333–3340, 2005.

[347] S. C. Berens, B. H. Joensen, and A. J. Horner. Tracking the emergence of location-based spatial representations. *bioRxiv*, page 547976, 2019.

[348] S. D. Auger, P. Zeidman, and E. A. Maguire. Efficacy of navigation may be influenced by retrosplenial cortex-mediated learning of landmark stability. *Neuropsychologia*, 104:102–112, 2017-09.

[349] K. M. O'Craven and N. Kanwisher. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience*, 12(6):1013–1023, 2000.

[350] J. J. Knierim, H. S. Kudrimoti, and B. L. McNaughton. Place cells, head direction cells, and the learning of landmark stability. *The Journal of Neuroscience*, 15(3):1648–1659, 1995.

[351] A. T. Keinath, J. B. Julian, R. A. Epstein, and I. A. Muzzio. Environmental geometry aligns the hippocampal map during spatial reorientation. *Current Biology*, 27(3):309–317, 2017.

[352] N. K. Logothetis, J. Pauls, M. Augath, T. Trinath, and A. Oeltermann. Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412(6843):150–157, 2001.

[353] R. Kaplan, A. J. Horner, P. A. Bandettini, C. F. Doeller, and N. Burgess. Human hippocampal processing of environmental novelty during spatial navigation. *Hippocampus*, 24(7):740–750, 2014.

[354] C. F. Doeller and N. Burgess. Distinct error-correcting and incidental learning of location relative to landmarks and boundaries. *Proc Natl Acad Sci U S A*, 105:5909–5914, 2008.

[355] K. Gramann, J. Onton, D. Riccobon, H. J. Mueller, S. Bardins, and S. Makeig. Human brain dynamics accompanying use of egocentric and allocentric reference frames during navigation. *Journal of Cognitive Neuroscience*, 22(12):2836–2849, 2010.

[356] J. L. Gardner and T. Liu. Inverted encoding models reconstruct an arbitrary model response, not the stimulus. *Eneuro*, 6(2):ENEURO.0363–18.2019, 2019.

[357] T. Liu, D. Cable, and J. L. Gardner. Inverted encoding models of human population response conflate noise and neural tuning width. *The Journal of Neuroscience*, 38(2):398–408, 2018.

[358] T. C. Sprague, J. T. Serences, and G. M. Boynton. Inverted encoding models estimate sensible channel responses for sensible models. *bioRxiv*, pages 1–12, 2019.

[359] T. C. Sprague, K. C. S. Adam, J. J. Foster, M. Rahmati, D. W. Sutterer, and V. A. Vo. Inverted encoding models assay population-level stimulus representations, not single-unit neural tuning. *Eneuro*, 5(3):ENEURO.0098–18.2018, 2018.

[360] J. D. Power, K. A. Barnes, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3):2142–2154, 2012.

[361] B. A. Poser, P. J. Koopmans, T. Witzel, L. L. Wald, and M. Barth. Three dimensional echo-planar imaging at 7 tesla. *NeuroImage*, 51(1):261–266, 2010.

[362] R. Cusack, A. Vicente-Grabovetsky, D. J. Mitchell, C. J. Wild, T. Auer, A. C. Linke, and J. E. Peelle. Automatic analysis (aa): efficient neuroimaging workflows and parallel processing using matlab and XML. *Frontiers in Neuroinformatics*, 8, 2015.

[363] B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, 54(3):2033–2044, 2011.

[364] T. Yarkoni, R. A Poldrack, T. E. Nichols, D. C. Van Essen, and T. D. Wager. Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, 8(8):665–670, 2011-08.

[365] P. Berens. CircStat: A MATLAB toolbox for circular statistics. *Journal of Statistical Software*, 31(10), 2009.

[366] C. Galletti and P. Fattori. Neuronal mechanisms for detection of motion in the field of view. *Neuropsychologia*, 41(13):1717–1727, 2003.

[367] C. M. Zingale and E. Kowler. Planning sequences of saccades. *Vision research*, 27(8):1327–41, 1987.

[368] M. M. Hayhoe, A. Shrivastava, R. Mruczek, and J. B. Pelz. Visual memory and motor planning in a natural task. *Journal of vision*, 3(1):6–6, 2003.

[369] G. Megardon, C. Ludwig, and P. Sumner. Trajectory curvature in saccade sequences: spatiotopic influences vs. residual motor activity. *Journal of Neurophysiology*, 118(2):1310–1320, 2017.

[370] N. Burgess, J. O'Keefe, and M. Recce. Using hippocampal'place cells' for navigation, exploiting phase coding. In *Advances in neural information processing systems*, pages 929–936, 1993.

[371] W. B. Scoville and B. Milner. Loss of recent memory after bilateral hippocampal lesion. *Journal of Neurology, Neurosurgery & Psychiatry*, 20(1):11–21.

[372] H. Eichenbaum and N. J. Cohen. Can we reconcile the declarative memory and spatial navigation views on hippocampal function? *Neuron*, 83(4):764–770, 2014.

[373] E. Tulving. Episodic memory: From mind to brain. *Annual Review of Psychology*, 53(1):1–25.

[374] J. Sugar and M. B. Moser. Episodic memory: Neuronal codes for what, where, and when. *Hippocampus*.

[375] J. J. Knierim and J. P. Neunuebel. Tracking the flow of hippocampal computation: Pattern separation, pattern completion, and attractor dynamics. *Neurobiology of Learning and Memory*, 129:38–49.

[376] K. L Stachenfeld, M. M Botvinick, and S. J. Gershman. Author correction: The hippocampus as a predictive map. *Nature Neuroscience*, 21(6):895–895, 2018.

[377] S. Theves, G. Fernandez, and C. F. Doeller. The hippocampus encodes distances in multidimensional feature space. *Current Biology*, 29(7):1226–1231.e3.

[378] S. Leutgeb and J. K. Leutgeb. Pattern separation, pattern completion, and new neuronal codes within a continuous CA3 map. *Learning & Memory*, 14(11):745–757.

[379] M. E. Wheeler, S. E. Petersen, and R. L. Buckner. Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences*, 97(20):11125–11129.

[380] P. W. Frankland and B. Bontempi. The organization of recent and remote memories. *Nature Reviews Neuroscience*, 6(2):119–130.

[381] N. Dijkstra, S. E. Bosch, and M. A. J. van Gerven. Vividness of visual imagery depends on the neural overlap with perception in visual areas. *The Journal of Neuroscience*, 37(5):1367–1373.

[382] P. Zeidman and E. A. Maguire. Anterior hippocampus: The anatomy of perception, imagination and episodic memory. *Nature Reviews Neuroscience*, 17(3):173–182, 2016-03.

[383] B. P. Staresina, R. N. A. Henson, N. Kriegeskorte, and A. Alink. Episodic reinstatement in the medial temporal lobe. *Journal of Neuroscience*, 32(50):18150–18156.

[384] K.Z. Tanaka, A. Pevzner, A.B. Hamidi, Y. Nakazawa, J. Graham, and B.J. Wiltgen. Cortical representations are reinstated by the hippocampus during memory retrieval. *Neuron*, 84(2):347–354.

[385] A. Clark. Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3):181–204.

[386] F. P. de Lange, M. Heilbron, and P. Kok. How do expectations shape perception? *Trends in Cognitive Sciences*, 22(9):764–779.

[387] J. O'Keefe and M. L. Recce. Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, 3(3):317–330.

[388] J. Lisman and A. D. Redish. Prediction, sequences and the hippocampus. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521):1193–1201.

[389] T. Hafting, M. Fyhn, T. Bonnevie, M. B. Moser, and E. I. Moser. Hippocampus-independent phase

precession in entorhinal grid cells. *Nature*, 453(7199):1248–1252.

[390] A.C. Schapiro, L.V. Kustner, and N.B. Turk-Browne. Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current Biology*, 22(17):1622–1627.

[391] N. C. Hindy, E. W. Avery, and N. B. Turk-Browne. Hippocampal-neocortical interactions sharpen over time for predictive actions. *Nature Communications*, 10(1).

[392] C. L. Colby and M. E. Goldberg. Space and attention in parietal cortex. *Annual Review of Neuroscience*, 22(1):319–349, 2002.

[393] I. C. Fiebelkorn, M. A. Pinsk, and S. Kastner. A dynamic interplay within the frontoparietal network underlies rhythmic spatial attention. *Neuron*, 99(4):842–853.e8, 2018.

[394] R. F. Helfrich, I. C. Fiebelkorn, S. M. Szczepanski, J. J. Lin, J. Parvizi, R. T. Knight, and S. Kastner. Neural mechanisms of sustained attention are rhythmic. *Neuron*, 99(4):854–865.e5, 2018.

[395] A.N. Landau and P. Fries. Attention samples stimuli rhythmically. *Current Biology*, 22(11):1000–1004.

[396] R. VanRullen, T. Carlson, and P. Cavanagh. The blinking spotlight of attention. *Proceedings of the National Academy of Sciences*, 104(49):19204–19209.

[397] I. C. Fiebelkorn and S. Kastner. A rhythmic theory of attention. *Trends in Cognitive Sciences*, 23(2):87–101.

[398] I. C. Fiebelkorn and S. Kastner. Functional specialization in the attention network. *Annual Review of Psychology*, 71(1).

[399] M. P. Brandon, A. R. Bogaard, C. P. Libby, M. A. Connerney, K. Gupta, and M. E. Hasselmo. Reduction of theta rhythm dissociates grid cell spatial periodicity from directional tuning. *Science*, 332(6029):595–599, 2011.

[400] J. Koenig, A. N. Linder, J. K. Leutgeb, and S. Leutgeb. The spatial periodicity of grid cells is not sustained during reduced theta oscillations. *Science*, 332(6029):592–595, 2011.

[401] U. Rutishauser, I. B. Ross, A. N. Mamelak, and E. M. Schuman. Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature*, 464(7290):903, 2010.

[402] M. J. Jutras, P. Fries, and E. A. Buffalo. Oscillatory activity in the monkey hippocampus during visual exploration and memory formation. *Proceedings of the National Academy of Sciences*, 110(32):13144–13149, 2013.

[403] K. L. Hoffman, M. C. Dragan, T. K. Leonard, C. Micheli, R. Montefusco-Siegmund, and T. A. Valiante. Saccades during visual exploration align hippocampal 38 hz rhythms in human and non-human primates. *Frontiers in systems neuroscience*, 7:43, 2013.

[404] H. McCartney, A. D. Johnson, Z. M. Weil, and B. Givens. Theta reset produces optimal conditions for long-term potentiation. *Hippocampus*, 14(6):684–687.

[405] M. Aly, C. Ranganath, and A. P. Yonelinas. Detecting changes in scenes: The hippocampus is critical for strength-based perception. *Neuron*, 78(6):1127–1137, 2013-06.

[406] M. Aly and N. B. Turk-Browne. How hippocampal memory shapes, and is shaped by, attention. In *The Hippocampus from Cells to Systems*, pages 369–403. Springer, 2017.

[407] N. Rungratsameetaweemana, L. R. Squire, and J. T. Serences. Preserved capacity for learning statistical regularities and

directing selective attention after hippocampal lesions. *Proceedings of the National Academy of Sciences*, page 201904502.

[408] N. A. Ruiz, M. R. Meager, S. Agarwal, and M. Aly. The hippocampus is critical for spatial relational attention. *bioRxiv*.

[409] J. S. Wynn, K. Shen, and J. D. Ryan. Eye movements actively reinstate spatiotemporal mnemonic content. *Vision*, 3(2):21.

[410] J. M. Henderson. Gaze control as prediction. *Trends in cognitive sciences*, 21(1):15–23, 2017.

[411] D. E. Hannula. Worth a glance: using eye movements to investigate the cognitive neuroscience of memory. *Frontiers in Human Neuroscience*, 4, 2010.

[412] B. Laeng, I. M. Bloem, S. D'Ascenzo, and L. Tommasi. Scrutinizing visual images: The role of gaze in mental imagery and memory. *Cognition*, 131(2):263–283, 2014.

[413] M. B. Bone, M. St-Laurent, C. Dang, D. A. McQuiggan, J. D Ryan, and B. R. Buchsbaum. Eye movement reinstatement and neural reactivation during mental imagery. *Cerebral Cortex*, 29(3):1075–1089.

[414] A. J. Ryals, J. X. Wang, K. L. Polnaszek, and J. L. Voss. Hippocampal contribution to implicit configuration memory expressed via eye movements during scene exploration: Implicit configuration memory. *Hippocampus*, 25(9):1028–1041.

[415] D. E. Hannula and C. Ranganath. The eyes have it: Hippocampal activity predicts expression of memory in eye movements. *Neuron*, 63(5):592–599, 2009-09.

[416] A. C.H. Lee, T. J. Bussey, E. A. Murray, L. M. Saksida, R. A. Epstein, N. Kapur, J. R. Hodges, and K. S. Graham. Perceptual deficits in amnesia: challenging the medial temporal lobe mnemonic' view. *Neuropsychologia*, 43(1):1–11, 2005-01.

[417] M. Behrmann, A.C.H. Lee, J.Z. Geskin, K.S. Graham, and M.D. Barense. Temporal lobe contribution to perceptual function: A tale of three patient groups. *Neuropsychologia*, 90:33–45, 2016-09.

[418] T. Hartley, C. M. Bird, D. Chan, L. Cipolotti, M. Husain, F. Vargha-Khadem, and N. Burgess. The hippocampus is required for short-term topographical memory in humans. *Hippocampus*, 17(1):34–48, 2007-01.

[419] J. S. Matthis, J. L. Yates, and M. M. Hayhoe. Gaze and the control of foot placement when walking in natural terrain. *Current Biology*, 0(0), 2018-04.

[420] L. Piccardi, M. De Luca, R. Nori, L. Palermo, F. Iachini, and C. Guariglia. Navigational style influences eye movement pattern during exploration and learning of an environmental map. *Frontiers in behavioral neuroscience*, 10:140, 2016.

[421] M. L. Mittelstaedt and H. Mittelstaedt. Homing by path integration in a mammal. *Naturwissenschaften*, 67:566–567, 1980.

[422] A. S. Etienne and K. J. Jeffery. Path integration in mammals. *Hippocampus*, 14:180–192, 2004.

[423] C. R. Gallistel. *The organization of learning*. The organization of learning. The MIT Press, 1990.

[424] S. Zhang, F. Schoenfeld, L. Wiskott, and D. Manahan-Vaughan. Spatial representations of place cells in darkness are supported by path integration and border information. *Frontiers in behavioral neuroscience*, 8:222, 2014.

[425] D. A. Dombeck, C. D. Harvey, L. Tian, L. L. Looger, and D. W. Tank. Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. *Nature Neuroscience*, 13(11):1433–1440, 2010-11.

[426] S. S. Winter, M. L. Mehlman, B. J. Clark, and J. S. Taube. Passive transport disrupts grid signals in the parahippocampal cortex. *Current Biology*, 25(19):2493–2502, 2015.

[427] T. Stensola, H. Stensola, M. B. Moser, and E. I. Moser. Shearing-induced asymmetry in entorhinal grid cells. *Nature*, 518(7538):207–212, 2015.

[428] J. Krupic, M. Bauza, S. Burton, C. Barry, and J. O'Keefe. Grid cell symmetry is shaped by environmental geometry. *Nature*, 518(7538):232–235, 2015.

[429] K. Hardcastle, S. Ganguli, and L. M. Giocomo. Environmental boundaries as an error correction mechanism for grid cells. *Neuron*, 86(3):827–839, 2015.

[430] S. A. Lee, J. F. Miller, A. J. Watrous, M. R. Sperling, A. Sharan, G. A. Worrell, B. M. Berry, J. P. Aronson, K. A. Davis, and R. E. Gross. Electrophysiological signatures of spatial boundaries in the human subiculum. *Journal of Neuroscience*, 38(13):3265–3272, 2018.

[431] A. Treisman. Preattentive processing in vision. In *Human and machine vision II*, pages 313–334. Elsevier, 1986.

[432] A. C. H. Lee, L. K. Yeung, and M. D. Barense. The hippocampus and visual perception. *Frontiers in Human Neuroscience*, 6:1–17, 2012.

[433] R. Q. Quiroga. Neuronal codes for visual perception and memory. *Neuropsychologia*, 83:227–241, 2016-03.

[434] P. Zeidman, S. L. Mullally, and E. A. Maguire. Constructing, perceiving, and maintaining scenes: Hippocampal activity and connectivity. *Cerebral Cortex*, 25(10):3836–3855, 2015-10.

[435] Z. J. Urgolites, R. O. Hopkins, and L. R. Squire. Medial temporal lobe and topographical memory. *Proceedings of the National Academy of Sciences*, 114(32):8626–8630, 2017-08.

[436] C. J. Hodgetts, J. P. Shine, A. D. Lawrence, P. E. Downing, and K. S. Graham. Evidencing a place for the hippocampus within the core scene processing network: The HC's place in the core scene processing network. *Human Brain Mapping*, 37(11):3779–3794, 2016-11.

[437] A. D. Ekstrom and C. Ranganath. Space, time, and episodic memory: The hippocampus is all over the cognitive map. *Hippocampus*, pages 1–8, 2017.

[438] D. Hassabis, D. Kumaran, and E. A. Maguire. Using imagination to understand the neural basis of episodic memory. *Journal of Neuroscience*, 27(52):14365–14374, 2007-12.

[439] N. Niehof, J. J. Tramper, C. F. Doeller, and P. W. Medendorp. Updating of visual orientation in a gravity-based reference frame. *Journal of vision*, 17(12):4–4, 2017.

[440] C. Marendaz, P. Stivalet, L. Barraclough, and P. Walkowiac. Effect of gravitational cues on visual search for orientation. *Journal of Experimental Psychology: Human Perception and Performance*, 19(6):1266, 1993.

[441] J. Laurens, B. Kim, J. D. Dickman, and D. E. Angelaki. Gravity orientation tuning in macaque anterior thalamus. *Nature Neuroscience*, 19(12):1566–1568, 2016.

[442] M. C. Morrone, M. Cicchini, and D. C. Burr. Spatial maps for time and motion. *Experimental brain research*, 206(2):121–128, 2010.

[443] B. L. McNaughton, F. P. Battaglia, O. Jensen, E. I. Moser, and M. B. Moser. Path integration and the neural basis of the 'cognitive map'. *Nature Reviews Neuroscience*, 7(8):663–678, 2006.

[444] J. Cavanaugh, R. A. Berman, W. M. Joiner, and R. H. Wurtz. Saccadic corollary discharge underlies stable visual perception. *Journal of Neuroscience*, 36(1):31–42, 2016.

[445] K. Mirpour and J. W. Bisley. Remapping, spatial stability, and temporal continuity: From the pre-saccadic to postsaccadic representation of visual space in LIP. *Cerebral Cortex*, 26(7):3183–3195, 2016-07.

[446] J. D. Golomb and N. Kanwisher. Retinotopic memory is more precise than spatiotopic memory. *Proceedings of the National Academy of Sciences*, 109(5):1796–1801, 2012.

[447] A. Schindler and A. Bartels. Integration of visual and non-visual self-motion cues during voluntary head movements in the human brain. *NeuroImage*, 172:597–607, 2018.

[448] A. B. Vog, M. D. Finch, and R. C. Olson. Functional heterogeneity in cingulate cortex: The anterior executive and posterior evaluative regions. *Cerebral Cortex*, 2(6):435–443, 1992.

[449] A. S. Alexander and D. A. Nitz. Spatially periodic activation patterns of retrosplenial cortex encode route sub-spaces and distance traveled. *Current Biology*, 27(11):1551–1560, 2017.

[450] A. S. Alexander and D. A. Nitz. Retrosplenial cortex maps the conjunction of internal and external spaces. *Nature Neuroscience*, 18(8):1143–1151, 2015.

[451] D. M. Smith, J. Barredo, and S. J. Y. Mizumori. Complimentary roles of the hippocampus and retrosplenial cortex in behavioral context discrimination. *Hippocampus*, 22(5):1121–1133, 2012.

[452] H. L. Dean and M. L. Platt. Allocentric spatial referencing of neuronal activity in macaque posterior cingulate cortex. *Journal of Neuroscience*, 26(4):1117–1127, 2006.

[453] M. Vz-Fort, E. F. Bracey, S. Keshavarzi, C. V. Rousseau, L. Cossell, S. C. Lenzi, M. Strom, and T. W. Margrie. A circuit for integration of head-and visual-motion signals in layer 6 of mouse primary visual cortex. *Neuron*, 2018.

[454] S. Van der Stigchel and A. Hollingworth. Visuospatial working memory as a fundamental component of the eye movement system. *Current Directions in Psychological Science*, 27(2):136–143, 2018.

[455] A. C. Marino and J. A. Mazer. Perisaccadic updating of visual representations and attentional states: Linking behavior and neurophysiology. *Frontiers in Systems Neuroscience*, 10, 2016.

[456] M. M. Umeno and M. E. Goldberg. Spatial processing in the monkey frontal eye field. II. memory responses. *Journal of Neurophysiology*, 86(5):2344–2352, 2001.

[457] A. Banino, C. Barry, B. Uria, C. Blundell, T. Lillicrap, P. Mirowski, A. Pritzel, M. J. Chadwick, T. Degris, J. Modayil, G. Wayne, H. Soyer, F. Viola, B. Zhang, R. Goroshin, N. Rabinowitz, R. Pascanu, C. Beattie, S. Petersen, A. Sadik, S. Gaffney, H. King, K. Kavukcuoglu, D. Hassabis, R. Hadsell, and D. Kumaran. Vector-based navigation using grid-like representations in artificial agents. *Nature*, 557(7705):429–433, 2018.

[458] D. E. Warren, M. C. Duff, N. J. Cohen, and D. Tranel. Hippocampus contributes to the maintenance but not the quality of visual information over time. *Learning & Memory*, 22(1):6–10, 2015-01.

[459] J. Erez, A. C. H. Lee, and M. D. Barense. It does not look odd to me: Perceptual impairments and eye movements in amnesic patients with medial temporal lobe damage. *Neuropsychologia*, 51(1):168–180, 2013.

[460] I. R. Olson, K. S. Moore, M. Stark, and A. Chatterjee. Visual working memory is impaired when the medial temporal lobe is damaged. *Journal of Cognitive Neuroscience*, 18(7):1087–1097, 2006.

[461] J. D. Schall, A. Morel, D. J. King, and J. Bullier. Topography of visual cortex

connections with frontal eye field in macaque: convergence and segregation of processing streams. *The Journal of Neuroscience*, 15(6):4464–4487, 1995.

[462] S. L. Ding. Comparative anatomy of the prosubiculum, subiculum, presubiculum, postsubiculum, and parasubiculum in human, monkey, and rodent. *Journal of Comparative Neurology*, 521(18):4145–4162, 2013.

[463] R. Insausti and D. G. Amaral. Entorhinal cortex of the monkey: IV. topographical and laminar organization of cortical afferents. *Journal of Comparative Neurology*, 509(6):608–641, 2008.

[464] L. G. Ungerleider, T. W. Galkin, R. Desimone, and R. Gattass. Cortical connections of area v4 in the macaque. *Cerebral Cortex*, 18(3):477–499, 2008.

[465] C. Distler, D. Boussaoud, R. Desimone, and L. G. Ungerleider. Cortical connections of inferior temporal area TEO in macaque monkeys. *The Journal of Comparative Neurology*, 334(1):125–150, 1993-08.

[466] M. Yoshida, Y. Naya, and Y. Miyashita. Anatomical organization of forward fiber projections from area TE to perirhinal neurons representing visual long-term memory in monkeys. *Proceedings of the National Academy of Sciences*, 100(7):4257–4262, 2003.

[467] A. W. Roe, L. Chelazzi, C. E. Connor, B. R. Conway, I. Fujita, J. L. Gallant, H. Lu, and W. Vanduffel. Toward a unified theory of visual area v4. *Neuron*, 74(1):12–29, 2012.

[468] H. Hong, D. L. K. Yamins, N. J. Majaj, and J. J. DiCarlo. Explicit information for category-orthogonal object properties increases along the ventral stream. *Nature Neuroscience*, 19(4):613–622, 2016-04.

[469] M. Mund R. Insausti. Cortical efferents of the entorhinal cortex and the adjacent parahippocampal region in the monkey (macaca fascicularis). *European Journal of Neuroscience*, 22(6):1368–1388, 2005.

[470] A. Kafkas and D. Montaldi. Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *Quarterly Journal of Experimental Psychology*, 64(10):1971–1989, 2011.

[471] R. J. Molitor, P. C. Ko, E. P. Hussey, and B. A. Ally. Memory-related eye movements challenge behavioral measures of pattern completion and pattern separation. *Hippocampus*, 24(6):666–672, 2014.

[472] Z. X. Liu, K. Shen, R. K. Olsen, and J. D. Ryan. Visual sampling predicts hippocampal activity. *Journal of Neuroscience*, 37(3):599–609, 2017.

[473] T. Staudigl, E. Hartl, S. Noachtar, C. F. Doeller, and O. Jensen. Saccades are phase-locked to alpha oscillations in the occipital and medial temporal lobe during successful memory encoding. *PLoS biology*, 15(12):e2003404, 2017.

[474] L. L. Colgin. Rhythms of the hippocampal network. *Nature Reviews Neuroscience*, 17(4):239–249, 2016-04.

[475] C. N. Smith and L. R. Squire. Experience-dependent eye movements reflect hippocampus-dependent (aware) memory. *Journal of Neuroscience*, 28(48):12825–12833, 2008.

[476] M. J. Jutras, P. Fries, and E. A. Buffalo. Gamma-band synchronization in the macaque hippocampus and memory formation. *Journal of Neuroscience*, 29(40):12521–12531, 2009.

[477] R. K. Olsen, V. Sebanayagam, Y. Lee, M. Moscovitch, C. L. Grady, S. R. Rosenbaum, and J. D. Ryan. The relationship between eye movements and subsequent recognition: Evidence from individual differences and amnesia. *Cortex*, 85:182–193, 2016-12.

[478] T. I. Brown, V. A. Carr, K. F. LaRocque, S. E. Favila, A. M. Gordon, B. Bowles, J. N. Bailenson, and A. D. Wagner. Prospective representation of navigational goals in the human hippocampus. *Science*, 352(6291):1323–1326, 2016.

[479] T. K. Leonard, J. M. Mikkila, E. N. Eskandar, J. L. Gerrard, D. Kaping, S. R. Patel, T. Womelsdorf, and K. L. Hoffman. Sharp wave ripples during visual exploration in the primate hippocampus. *Journal of Neuroscience*, 35(44):14771–14782, 2015.

[480] Z. Wang and R. M. Klein. Searching for inhibition of return in visual search: A review. *Vision research*, 50(2):220–228, 2010.

[481] S. Wang, A. N. Mamelak, R. Adolphs, and U. Rutishauser. Encoding of target detection during visual search by single neurons in the human brain. *Current Biology*, 28(13):2058–2069.e4, 2018.

[482] S. P. Jadhav, C. Kemere, P. W. German, and L. M. Frank. Awake hippocampal sharp-wave ripples support spatial memory. *Science*, page 1217230, 2012.

[483] T. K. Leonard and K. L. Hoffman. Sharp-wave ripples in primates are enhanced near remembered visual objects. *Current Biology*, 27(2):257–262, 2017.

[484] L. T. S. Yee, D. E. Warren, J. L. Voss, M. C. Duff, D. Tranel, and N. J. Cohen. The hippocampus uses information just encountered to guide efficient ongoing behavior. *Hippocampus*, 24(2):154–164, 2014.

[485] C. B. Alme, C. Miao, K. Jezek, A. Treves, E. I. Moser, and M. B. Moser. Place cells in the hippocampus: eleven maps for eleven rooms. *Proceedings of the National Academy of Sciences*, 111(52):18428–18435, 2014.

[486] C. Lever, T. Wills, F. Cacucci, N. Burgess, and J. O'Keefe. Long-term plasticity in hippocampal place-cell representation of environmental geometry. *Nature*, 416:90–94, 2002.

[487] C. Kentros, E. Hargreaves, R. D. Hawkins, E. R. Kandel, M. Shapiro, and R. V. Muller. Abolition of long-term stability of new hippocampal place cell maps by NMDA receptor blockade. *Science*, 280:2121–2126, 1998.

[488] M. M. Chun and Y. Jiang. Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive psychology*, 36(1):28–71, 1998.

[489] B. Colagiuri and E. J. Livesey. Contextual cuing as a form of nonconscious learning: Theoretical and empirical analysis in large and very large samples. *Psychonomic bulletin & review*, 23(6):1996–2009, 2016.

[490] A. J. Greene, W. L. Gross, C. L. Elsinger, and S. M. Rao. Hippocampal differentiation without recognition: an fMRI analysis of the contextual cueing task. *Learning & memory*, 14(8):548–553, 2007.

[491] E. V. Goldfarb, M. M. Chun, and E. A. Phelps. Memory-guided attention: Independent contributions of the hippocampus and striatum. *Neuron*, 89(2):317–324, 2016.

[492] M. M. Chun and E. A. Phelps. Memory deficits for implicit contextual information in human amnesics. *Nature Neuroscience*, 2(9):844–847, 1999.

[493] J. R. Manns and L. R. Squire. Perceptual learning, awareness, and the hippocampus. *Hippocampus*, 11(6):776–782, 2001.

[494] A. R. Preston and J. D. E. Gabrieli. Dissociation between explicit memory and configural memory in the human medial temporal lobe. *Cerebral Cortex*, 18(9):2192–2207, 2008.

[495] S. Negash, D. Kliot, D. V. Howard, J. H. Howard, S. R. Das, P. A. Yushkevich, J. B. Pluta, S. E. Arnold, and D. A. Wolk.

Relationship of contextual cueing and hippocampal volume in amnestic mild cognitive impairment patients and cognitively normal older adults. *Journal of the International Neuropsychological Society*, 21(4):285–296, 2015.

[496] J. R. Brockmole, D. Z. Hambrick, D. J. Windisch, and J. M. Henderson. The role of meaning in contextual cueing: Evidence from chess expertise. *Quarterly journal of experimental psychology*, 61(12):1886–1896, 2008.

[497] T. Beesley, M. A. Vadillo, D. Pearson, and D. R. Shanks. Pre-exposure of repeated search configurations facilitates subsequent contextual cuing of visual search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(2):348, 2015.

[498] T. J. Preston, F. Guo, K. Das, B. Giesbrecht, and M. P. Eckstein. Neural representations of contextual guidance in visual search of real-world scenes. *Journal of Neuroscience*, 33(18):7846–7855, 2013.

[499] M. Geva-Sagiv, S. Romani, L. Las, and N. Ulanovsky. Hippocampal global remapping for different sensory modalities in flying bats. *Nature Neuroscience*, 19(7):952–958, 2016.

[500] S. Wirth and P. Baraduc. Spatial orientation in the primate: I see, there i am. *mcine/sciences*, 34(1):33–36, 2018-01.

[501] K. Jezek, E. J. Henriksen, A. Treves, E. I. Moser, and M. B. Moser. Theta-paced flickering between place-cell maps in the hippocampus. *Nature*, 478(7368):246, 2011.

ISBN 978-82-326-4418-6 (printed version)
ISBN 978-82-326-4419-3 (electronic version)
ISSN 1503-8181

Doctoral theses at NTNU, 2020:32

Matthias Nau

Perception and the cognitive map

Deriving a stable world from visual inputs

NTNU
Norwegian University of
Science and Technology
Faculty of Medicine and Health Sciences
Kavli Institute for Systems Neuroscience

NTNU
Norwegian University of
Science and Technology

Doctoral Thesis
Matthias Nau
Doctoral theses at NTNU, 2020:32
NTNU

NTNU
Norwegian University of
Science and Technology