Neural Basis of Route-Planning and Goal-Coding During Flexible Navigation

Christoffer J. Gahnstrom

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Experimental Psychology

University College London

April 8, 2022

I, Christoffer J. Gahnstrom, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Animals and humans are remarkable in their ability to flexibly adapt to changes in their surroundings. Navigational flexibility may take many forms and in this thesis we investigate its neural and behavioral underpinnings using a variety of methods and tasks tailored to each specific research aim. These methods include functional resonance magnetic imaging (fMRI), freely moving virtual reality, desktop virtual reality, large-scale online testing, and computational modelling.

First, we reanalysed previously collected rodent data in the lab to better understand behavioural bias that may occur during goal-directed navigation tasks. Based on finding some biases we designed a new approach of simulating results on maze configurations prior to data collection to select the ideal mazes for our task. In a parallel line of methods development, we designed a freely moving navigation task using large-scale wireless virtual reality in a 10x10 space. We compared human behaviour to that of a select number of reinforcement learning agents to investigate the feasibility of computational modelling approaches to freely moving behaviour.

Second, we further developed our new approach of simulating results on maze configuration to design a novel spatial navigation task used in a parallel experiment in both rats and humans. We report the human findings using desktop virtual reality and fMRI. We identified a network of regions including hippocampal, caudate nucleus, and lateral orbitofrontal cortex involvement in learning hidden goal locations. We also identified a positive correlation between Euclidean goal distance and brain activity in the caudate nucleus during ongoing navigation.

Third, we developed a large online testing paradigm to investigate the role of home environment on wayfinding ability. We extended previous reports that street

network complexity is beneficial in improving wayfinding ability as measured using a previously reported virtual navigation game, Sea Hero Quest, as well as in a novel virtual navigation game, City Hero Quest. We also report results of a navigational strategies questionnaire that highlights differences of growing up inside and outside cities in the United States and how this relates to wayfinding ability.

Fourth, we investigate route planning in a group of expert navigators, licensed London taxi drivers. We designed a novel mental route planning task, probing 120 different routes throughout the extensive street network of London. We find hippocampal and retrosplenial involvement in route planning. We also identify the frontopolar cortex as one of several brain regions parametrically modulated by planning demand.

Lastly, I summarize the findings from these studies and how they all come to provide different insights into our remarkable ability to flexibly adapt to navigational challenges in our environment.

Impact Statement

This thesis investigates flexible navigation using functional magnetic resonance imaging (fMRI) and an array of behavioural and computational methods. How our brains are capable of supporting planning and goal-related processes is still unknown. We use state-of-the-art methods including freely moving virtual reality technology to increase the richness of the behavioural data we are able to acquire in a lab setting.

We also develop new navigation tasks that can seamlessly be implemented for large online testing. During the Covid-19 pandemic, we have seen increased use of online platforms to conduct experiments. However, some technological developments were required to adapt our on-site experiment tasks to operate on a large scale. These results of these developments are reported in this thesis and will benefit future large-scale online experiments.

We also report initial findings from our ongoing route planning study on London taxi drivers. These are a unique group of professionals who all share the same detailed information of the London street network. This is an ideal test-bed for many theories of spatial navigation that are usually constrained to be test on a student population where time and monetary constraints means we cannot train them to even a fraction of the proficiency level of a Licensed taxi driver as it takes on average 3-4 years to complete full-time.

Adult brain plasticity as observed in London taxi drivers could also provide with a wealth of insights into brain diseases such as Alzheimer's disease. The same brain areas afflicted, the hippocampus, is the same brain area that changes structurally with years of experience as a London taxi driver. In this thesis, we try to

better understand how taxi drivers are able to mentally plan routes in a short period of time, a process that is thought to be supported by the hippocampus.

Acknowledgements

First, I would like to thank my PhD advisor Hugo Spiers for providing me with supervision and professional opportunities throughout the course of my PhD. I would also like to thank M-GATE, the Marie Sklodowska-Curie training network for generously funding my PhD, enabling me the freedom to pursue a number of training activities.

I thank the IBN for providing a welcoming research environment where everyone was open-minded, collaborative, and did not hesitate to share and discuss their work.

Pursuing a PhD would not have been possible without the support of my family. I thank my grandparents, Inger and Lars Gahnstrom, for always inspiring me to chase my interests wherever they may lead since I was a child. I thank my mother, Kicki, for the kindness and love she always showed me and the world. I also thank my uncle and godfather, Johan, whom I will always look up to.

I thank my London peer support group, a.k.a. the MLCoggers, a.k.a the D&D nerds, Claire, Yihan, Nils, Chris K, Aaron, and Seb.

Most of all thank you to my wife, Sarah, for making these three years the best of my life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Much of our success as a species could be argued to depend on our ability to adapt to changes in the environment. From changes in climate, to ice ages, to nomadic lifestyles in search of greener pastures, the human species spread rapidly over multiple continents soon after our arrival in Africa (Bronowski, 1973). The behavioural flexibility required for such adaptation is remarkable, and can be considered a hallmark of our intelligence. Other mammals are also capable of impressive behaviour flexibility, including laboratory rats (Tolman, 1943). The overall aim of this thesis is to better understand the neural underpinnings of flexible behaviour in the face of navigational complexity.

We will begin with a review summarising the neurobiology supporting flexible spatial navigation. Thereafter, we discuss the application of computational modelling to understanding behavioural flexibility in spatial navigation. Lastly, we present four studies that address flexible navigation in complex environments using a variety of perspectives and methodologies. The first study applies reinforcement learning models to study human freely moving behaviour. The second study uses a revised version of the same task to look at the neural codes of flexible navigation using fMRI. The third study uses large-scale online testing to investigate spatial behavioural strategies and the role of lived environments on wayfinding ability. The last study investigates the neural underpinnings of route planning in London Taxi drivers, a group of professionals whose daily life is based around flexibly updating their spatial behaviour in the ever-changing, complex environment of the London

street network.

1.1 Striatal and hippocampal contributions to flexible navigation

The hippocampus has been firmly established as playing a crucial role in flexible navigation (O'Keefe & Nadel, 1978; Ekstrom et al., 2018). Recent evidence suggests that dorsal striatum may also play an important role in such goal-directed behaviour in both rodents and humans (see Gahnstrom & Spiers, 2020). Across recent studies, activity in the caudate nucleus has been linked to forward planning and adaptation to changes in the environment. In particular, several human neuroimaging studies have found the caudate nucleus tracks information traditionally associated with that by the hippocampus (Javadi et al., 2019a). In this chapter, we examine this evidence and argue the dorsal striatum encodes the transition structure of the environment during flexible, goal-directed behaviour. We highlight that future research should explore the following: 1) Investigate neural responses during spatial navigation via a biophysically plausible framework explained by reinforcement learning models; and 2) Observe the interaction between cortical areas and both the dorsal striatum and hippocampus during flexible navigation.

Flexibility during goal-directed behaviour

Flexible adaptation in response to unexpected changes in the environment is a central challenge of navigation. Tolman adeptly illustrated this in his seminal work exploring the capacity of rodents to accommodate detours and adopt shortcuts in complex mazes (1948). This work led to the proposal of the cognitive map hypothesis for flexible behaviour, by which the brain constructs an internal representation of the environment to support navigation (Tolman, 1948). Subsequent neuroscientific research led O'Keefe & Nadel (1978) to propose that the hippocampus is primarily responsible for supporting this cognitive map. Particularly central to this proposal is the existence of 'place cells' in the hippocampus that show spatially localised activity patterns linked to boundaries and landmarks in an environment

(O'Keefe & Dostrovsky, 1971). This was followed by the discovery of a variety of other spatial coding cells supporting navigation (see Grieves and Jeffery, 2017 for review). Given the ubiquity of spatial representation in the hippocampus and neighbouring parahippocampal structures, several essential questions arise: 1) How is information used during flexible navigation, as suggested by the hypothesis of the cognitive map?; 2) What information does the hippocampus code for downstream regions during navigation?; and 3) What contributions might other regions of the brain's navigation systems, such as the dorsal striatum, have for flexible navigation?

Rodent studies lesioning dorsal striatum and hippocampus provide strong evidence for dissociable behavioural strategies related to intact function of these regions during spatial navigation (Andersen et al., 2006; White & Donald, 2002). 'Place learning' is a flexible process by which an animal learns associations between distal cues and goal locations in the environment; while response learning is an inflexible process whereby an animal learns a series of actions or responses necessary to reach the goal. Place learning can be investigated using the Morris water maze, a task that targets behavioural flexibility and spatial memory (Morris et al., 1982; McDonald & White, 1994; Pearce et al., 1998; Devan & White, 1999; Whishaw et al., 1987). By the original task protocol, a rat is placed at a pseudo-random location within a cylindrical arena filled with opaque water. No local cues other than distal landmarks and boundary distance are provided. Safety is achieved by swimming to a fixed platform located just below the opaque surface, hidden from view. Escape latencies record time to reach the platform during training as well as during probe trials (when the hidden platform is removed). Lesion or inactivation of the hippocampus impacts place learning by increasing escape latencies compared to that of non-lesioned controls (Sutherland et al., 1983; Morris et al., 1982). However, lesions in dorsal striatum impair simple approach behaviour when the platform is visible, and instead, rats will swim to previously learned platform location (McDonald & White, 1994).

A paradigm called Delayed-Matched-to-Place further extended the Morris wa-

ter maze by investigating one-shot learning, a hallmark of behavioural flexibility (Steele & Morris, 1999). In this version of the task, the location of the hidden platform changes each day. This results in a substantial drop in escape latency between the first and second trial. The subsequent trials exhibit latency improvement, but to a much smaller extent. This concept of one-shot learning is an impressive quality of cognitive flexibility difficult to capture by biophysically plausible modelling of place cells (Foster et al., 2000). However, reinforcement learning can capture this behavioural phenomenon by further simulation of cells which estimate real world coordinates (Tessereau et al., 2020; Foster et al., 2000). Together, these simulated cells form an allocentric coordinate system receiving input from the place cells. This coordinate system lacks a biological basis, although this may be analogous to information represented by grid cells in the entorhinal cortex (Hafting et al., 2005). Likewise, simulated deep reinforcement learning agents endowed with grid-like representation can perform flexible spatial navigation tasks such as the Morris water maze (Banino et al., 2018). Additionally, inactivation of the hippocampus by bilateral lesions to the fornix impairs performance in an eight-arm radial maze task, in which rats are trained to revisit certain arms consistently baited with food (Packard et al., 1989). Intact hippocampal function is necessary for place learning in a cross-maze task as well (Packard & McGaugh, 1996). Evidence from neuroimaging studies of humans and patients with hippocampal damage further implicates the hippocampus' role for both place learning and flexible navigation of novel routes and environments (Spiers et al. 2001a,b; Hartley et al., 2003; Bohbot et al., 2007; Iaria et al., 2004; Spiers & Maguire 2006; Xu et al., 2010; Howard et al., 2014; Javadi et al., 2019a; Javadi et al., 2019b; Patai et al., 2019).

In addition to place learning, animals also utilize 'response learning', i.e. learning based on the responses required to reach the goal (Packard & McGaugh, 1996). Such response learning is shown to depend on the functional integrity of the dorsal striatum (Packard & McGaugh, 1996; Packard et al., 1989). Differences in acquisition time of place and response learning, with the latter taking longer to form, also suggest that these behavioural strategies instantiate in the brain as dissociable neural

mechanisms (Packard & McGaugh, 1996). Subsequently, human neuroimaging research has provided convergent evidence for the involvement of the dorsal striatum in such response strategy navigation (Iaria et al. 2004; Voermans et al., 2004; Hartley et al., 2003). Response learning is not traditionally considered flexible because it is tied to the specific features of the environment (e.g. always turn right at the crossroad). By contrast, place learning is thought to be flexible since it is possible to use place-based information from the environment to accommodate detours and identify shortcuts.

Recent studies have begun to explore how different types of spatial information may be tracked by specific brain regions during navigation. Two important metrics for flexible navigation are vector-to-goal and path-to-goal (Bicanski & Burgess, 2020; Spiers & Barry, 2015; Chadwick et al., 2015). Using in situ learning experience and film simulation of Soho in London (UK), Howard et al. (2014) identified neural correlates of path distance to goal in the right posterior hippocampus. Such correlates of distance to goal have also been observed in dorsal hippocampal recordings in rats (Spiers et al., 2018) and bats (Sarel et al., 2017). During detour events, the human posterior right hippocampus was also found to track the increase in path distance when a forced detour occurred (Howard et al., 2014). Based on this finding and other evidence from rats (e.g. Pfeiffer & Foster, 2013; Gupta et al. 2010; Olafsdottir et al. 2015), it has been hypothesised the hippocampus simulates future paths through the environment at key events during navigation, such as at detours (Spiers & Gilbert, 2015). Consequently, detours requiring simulation of a much larger future route will evoke greater demands on the hippocampus than simulation of shorter routes.

In order to test the prediction of Spiers and Gilbert (2015), a recent study by Javadi and colleagues (2019a) examined hippocampal response to respectively small and large changes in distance to goal at forced detours (see Fig. 1a). In this task, participants navigated a virtual desert island riven with lava which blocked certain movements across it. Participants first learned the layout and location of several hidden objects, which later served as cued goal locations. During the test

phase, when participants actively navigated the maze, shifts in the location of lava pools either opened up new paths or blocked old paths, resulting in possible shortcuts and detours, respectively. In contrast to the predictions of Spiers and Gilbert (2015), posterior hippocampus did not index the change in distance to goal at detours, but rather prefrontal regions and bilateral caudate nucleus tracked the change in path distance to goal (Javadi et al., 2019a). Notably, in the work from Howard and colleagues (2014), the hippocampal response to distance changes at detours was also accompanied by a similar response in the dorsal striatum (Fig. 1b). Taken together, these results indicate the dorsal striatum is more consistent in tracking the change in distance at detours than the hippocampus. This suggests it is timely to reconsider the role of dorsal striatum during flexible navigation and understand how the hippocampus interacts with these regions in corticostriatal loops (Goodroe et al., 2018; Brown et al., 2012).

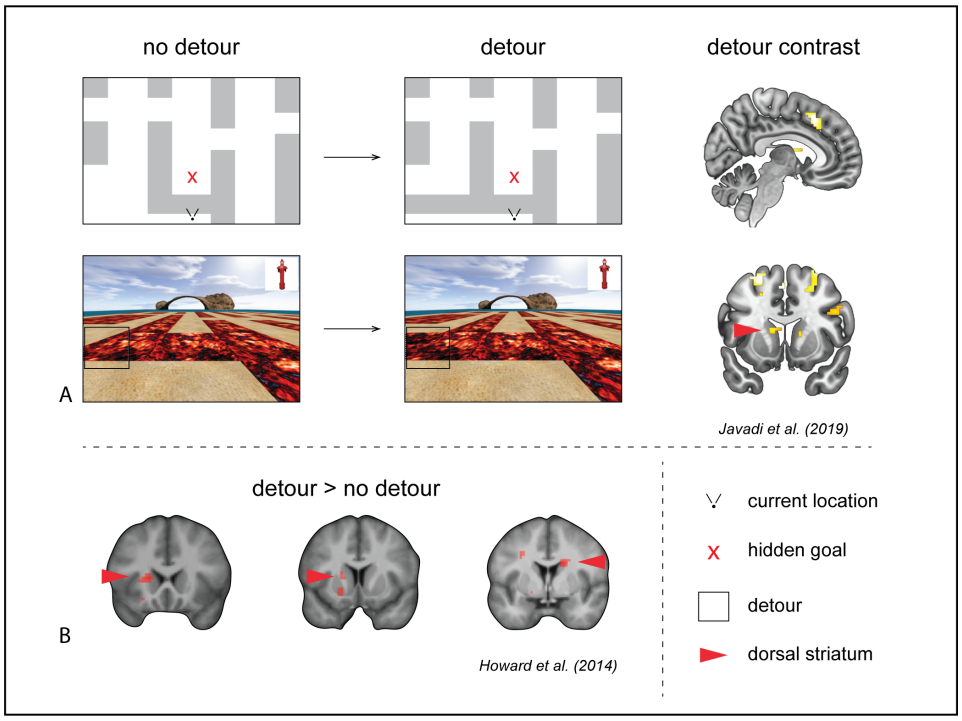

Figure 1.1: Dorsal striatum activity is correlated with the change in distance to goal at detours.

A) Replotted data from Javadi et al., (2019a) in which fMRI and virtual reality desert island riven with lava was used to examine the brain regions responsive to the change in distance to the goal at detours. Top row shows a zoomed in schematic from the larger virtual environment used and the transition that occurs when the path is unexpectedly blocked. Bottom row: the same change but from the first person perspective. Brain image shows bilateral activity in medial caudate nucleus (dorsal striatum) cluster-corrected for activity correlated positively with the change in distance. B) Replotted data from Howard et al., (2014). In this study a film simulation of Soho in London was used to test navigation, including accommodating detours. The amount of change in distance caused by forced detours was correlated with dorsal striatal activity. Red regions show regions activations thresholded at $p < 0.005$ uncorrected, shown on the mean structural image.

How might the striatum contribute to flexible navigation behaviour?

Despite the traditional role of response learning attributed to striatal function, the striatum has been implicated in studies investigating behavioural flexibility in both rodents and humans, suggesting a more nuanced functionality beyond contributing to a less flexible response system (Johnson et al., 2007). Lesions and inactivations in different areas of striatum produce varied behavioural deficits, indicating a dissociation of respective functional roles (Sharpe et al., 2019; Ragozzino et al., 2002).

The striatum is commonly divided up into two anatomically separated regions: the dorsal striatum, composing of the caudate and putamen, and the ventral striatum, composed mainly of the nucleus accumbens, although no clear cytoarchitectonic or histochemical boundary between ventral and dorsal striatum exist (Haber & Knutson, 2010). Furthermore, rodent caudate-putamen is segmented into dorsomedial striatum (homologous to primate caudate) and dorsolateral striatum (homologous to primate putamen) (Cox & Witten, 2019). Early rodent studies did not include strict separation of these regions when using large lesions, which leads to interpretation difficulties (Yin & Knowlton, 2006).

Reinforcement learning models provide a normative framework to investigate neural mechanisms that give rise to flexible and inflexible behaviour (Corrado et al., 2009). Within the reinforcement learning literature, flexible and goal-directed behaviour is often described by a family of algorithms classified as 'model-based.' This is commonly contrasted with habitual behaviour described by a separate family of algorithms classified as 'model-free' (Dolan & Dayan, 2013; Rusu & Pennartz, 2020). These computational models 'learn' states and rewards in the environment by using a component referred to as reward prediction error, i.e. the difference between expected and experienced reward. The goal of a reinforcement learning agent is to take actions which maximise long-term future reward (Sutton & Barto, 2018). Single neurons of the ventral tegmental area in the brainstem of macaques were found to encode reward prediction errors, a region which has direct dopaminergic projection to the nucleus accumbens in ventral striatum (Schultz et al., 1997; Haber & Knutson, 2010). Since then, human fMRI studies on multi-step decision making tasks have identified ventral striatum as a primary region for the processing of reward prediction errors (Gläscher et al., 2010; Daw et al., 2011). Daw and colleagues also found the striatal underpinnings of habitual model-free prediction errors and model-based prediction errors overlap in ventral striatum, suggesting the same neural circuitry is involved in both computations (2011). A recent meta-analysis found correlates for reward prediction errors in both ventral striatum and caudate nucleus (as well as overlapping areas) in model-based and model-free computations (Huang

et al., 2020).

Beyond the classic divisions of model-free and model-based literature in decision-making tasks, there are other families of reinforcement learning (RL) algorithms that provide alternative accounts, including hierarchical RL, linear RL, and successor representation (Tessereau et al., 2020; Botvinick et al., 2009; Piray & Daw, 2019; Dayan, 1993; Stachenheld et al., 2017; Russek et al., 2017; Gershman, 2018). In particular, successor representation can account for flexible behaviour of rats and humans in complex mazes (de Cothi et al., 2020) and humans in reward devaluation (Momennejad et al., 2017). Interestingly, components of the successor representation during simulations show similarities to properties of place cells and grid cells, including the influence of goal locations on place field over-representation and the influence of environmental geometry on grid field integrity (Stachenfeld et al., 2017; Krupic et al., 2015; Ekstrom et al., 2020). It is an interesting future direction for studies to investigate the relationship between neural responses and the internal computations of successor representation shown to account for behaviour flexibility, particularly in some spatial navigation tasks (Russek et al., 2017; for review see Momennejad, 2020). Recent work with rats navigating between four interconnected rooms has revealed that during initial adaptation to obstructed pathways , place cells in CA1 did not adapt their firing fields to accompany the changing behaviour (Duvelle et al., 2020) as might have been predicted by a model in which place cells support successor representation coding (Stachenfeld et al., 2017). It may be that more stereotyped trajectories are required for shifts in place fields as a result of topological manipulations.

The dorsal striatum has commonly been linked to stimulus-association, or habits, in spatial navigation tasks using human fMRI. Doeller and colleagues (2008) employed a virtual object-memory task inspired by the Morris water maze. They found activity in the caudate nucleus to be parametrically modulated by the influence of intramaze landmarks on goal locations, while the right posterior hippocampus correlated with boundary-related influence on goal locations (Doeller et al., 2008). In another study in which participants navigated a virtual town, caudate

activity was preferentially active during route following trials, while anterior hippocampus was preferentially active during wayfinding trials (Hartley et al., 2003). Likewise, Iaria et al. (2003) found place strategy use in an eight-arm radial maze task was associated with increased right hippocampal activity while non-spatial response strategy use was associated with increased activity in caudate nucleus. These studies suggest a dissociation between the roles of dorsal striatum and hippocampus for habitual and flexible behaviour, respectively. However, contextual demands may elucidate a more nuanced role for the striatum in multiple behavioural control circuits (Rusu & Pennartz, 2020; Balleine et al., 2015; Ferbinteanu, 2019).

In rodents, the involvement of dorsal striatum in both flexible and habitual behaviour could be resolved by considering the functional distinction of dorsolateral and dorsomedial regions (Regier et al., 2015; Thorn et al., 2010; van der Meer et al., 2010). Studies investigating the homologous regions in humans are made difficult by the lack of spatially precise recordings of neuronal activity. One account suggests dorsal striatum performs the role of an 'actor' while ventral striatum performs the parallel role of a 'critic' in the 'actor-critic' reinforcement learning framework (Sutton & Barto, 2018). In support of this idea, such a division in computational roles was found during an instrumental learning task using fMRI (O'Doherty et al., 2004). Investigation of functional distinction in dorsal striatum found putamen involvement in habit-based processing from extensive training versus caudate involvement in forward planning (Wunderlich et al., 2012). The role of forward planning at detours could be considered in the task by Javadi and colleagues (2019a) wherein distance changes were tracked by bilateral caudate nucleus (Fig. 1). In a virtual navigation task, Simon and Daw (2011) also found forward planning tracked by striatum using predictions from 'model-based' reinforcement learning.

In a more recent virtual navigation task, Anggriani et al. (2018) identified model-free correlates in dorsal striatum. Model-based correlates were found in the parahippocampus and overlapped with model-free correlates in the retrosplenial cortex. In contrast to Simon and Daw (2011), this study did not utilise visual goal cues and also did not include changes in the maze configuration, more

akin to classical spatial navigation paradigms. The different accounts of striatal involvement in prediction errors can perhaps be reconciled by considering that the behavioural strategies and neural mechanisms are not so easily dissociable as previously thought. One spatial planning task found striatal activity related to the difference in path distance between the shortest path and unchosen longest path to goal as a proxy for exhaustive search or forward planning (Kaplan et al., 2017). This indicates that striatal subregions may be involved in planning, which may be the reason these regions are active in different studies. Perhaps a mixed use of strategies is also an underlying reason for this result. Brown and colleagues (2012) showed that caudate is important for disambiguating context during spatial navigation, together with orbitofrontal cortex and hippocampus. We suggest these findings are in line with a new perspective of these regions. In this view, the caudate encodes learned transition structures, however, the current active transition structure at any point in time is based on the current state of the animal and context within the task, which is proposed to be modulated through cholinergic interneurons between dorsomedial striatum and orbitofrontal cortex (Sharpe et al., 2019). Hippocampus, on the other hand, is involved in learning the structure of the environment (incidental to the task), and also the accompanying association-based learning.

Instrumental learning paradigms in rodents reveal a model-based influence on model-free prediction errors (Langdon et al., 2018). As such, the classical role of dopaminergic prediction errors are more nuanced and can incorporate signals related to behavioural flexibility and the current state of the task in ventral tegmental area (Starkweather et al., 2017; Keiflin et al., 2019) as well as dorsomedial striatum (Stalnaker et al., 2016). Using causal methodology by optogenetically stimulating dopaminergic neurons in ventral tegmental area (the putative cells encoding reward prediction errors), rats could learn associations between cues without endowing them with cached-value, as would be expected based on pure model-free temporal-difference learning models (Sharpe et al., 2020). Another instrumental learning task found an increasing number of neurons encoding task-relevant information in dorsolateral striatum more so than dorsomedial, suggesting the former

may be encoding the development of a habit-based response (Kimchi et al., 2009). Recordings in rats navigating a T-maze found that neurons in dorsomedial striatum were primarily active while choosing between alternative actions after cue-onset, in contrast with neurons in dorsolateral striatum which were primarily active during action execution (Thorn et al., 2010). Stalnaker and colleagues (2016) found that cholinergic interneurons in rodent dorsomedial – and not dorsolateral striatum – represented information about the current state of the choice task. Additionally, this state information was not present in rats with lesions to the orbitofrontal cortex. Taken together, there appears to be shared neural circuitry for model-free and model-based behaviours, and prediction errors may convey more information than the difference between experienced and expected reward (Doll et al., 2012). Perhaps the aforementioned human studies can be reconciled with the notion that caudate can support a mixture of model-free and model-based computations dependent on the task and context at hand. Caudate nucleus activity can be expected in response to changes in transition structure if it also encodes model-based information regarding the task environment.

These recent findings pose a new question: What is the human dorsal striatum coding that drives these observed changes in activity during navigation? Rodent work on dorsomedial striatum suggests this region is necessary for execution of flexible goal-directed behaviour (Rusu & Pennartz, 2020). Similarly, dorsomedial lesions have demonstrated similar behavioural deficits to that of hippocampal lesions in terms of deficiencies in goal-directed flexible behaviour (Sharpe et al., 2019). For effective flexible behaviour, work from Sharpe and colleagues (2019) suggests the hippocampus provides information about the environmental structure, while dorsomedial striatum incorporates information about the transition structure into one's overall world model. In human navigation, novel forced detours are a classic example of a change in the transition structure. If the caudate updates representations of the transition structure, with greater transitional change resulting in greater demand on caudate activity, then this may explain the results of both Javadi et al. (2019a) and Howard et al (2014), see Fig. 1, where a larger change in

distance at detours is associated with greater evoked caudate activity. In contrast, hippocampus may be required to construct simulations of journeys through the environments (Bendor & Spiers 2016). Such simulations may have been much richer in the navigation of London's Soho (Howard et al. 2014), compared with a desert island (Javadi et al., 2019a), explaining the difference in hippocampal engagement. Entorhinal cortex may also be involved in representing low-dimensional features of environments by extracting basis sets (or eigenvectors of the successor representation), some of which look visually similar to the iconic hexagonal nature of grid fields (Behrens et al., 2018; Stachenfeld 2017). Lesions during the Morris water maze have shown the entorhinal cortex to be involved in flexible behaviour, as animals have similar behavioral deficits to those of hippocampal lesions in terms of increased swimming latencies to the hidden platform (Hales et al., 2014). One idea is that the entorhinal cortex supports the ability to form general transition structures of any environment and store information about how distant states or locations are related to each other (Behrens et al., 2018; Constantinescu et al., 2016). However, the unique dorsal striatum contribution may be more closely related to how action-outcome associations are represented and which state is transitioned to as a result of a given motor action (Sharpe et al., 2019).

In conclusion, evidence suggests the dorsomedial striatum / caudate nucleus plays a key role in flexible navigation by representing the transition structure of the environment for guiding future actions (Sharpe et al., 2019). This may explain observed responses at detours where transition structure changes (Howard et al., 2014; Javadi et al., 2019a). Future research will be useful to observe dorsomedial striatal activity in rodents during dynamic changes to the environment's transition structure and variations in update demands (e.g. detours that require larger or smaller shifts in the route to the goal). It would also be important to examine the interplay between the striatum, hippocampal/parahippocampal structures, and prefrontal cortex during such updating and representation for the structure of the environment (see Momennejad, 2020). The entorhinal cortex has also been proposed to play a role in coding the transition structure of the layout of the environment or stimulus set

(Behrens et al., 2018). Understanding how such a code relates to striatal coding of transition structure would be useful for advancing models of the neural systems supporting flexible navigation behaviour.

1.2 Computational view of reinforcement learning modelling in flexible navigation

There is a longstanding lack in our understanding of generalizable mechanisms underlying cognition. At the same time, the subfields of neuroscience investigating rodents invasively, and humans noninvasively, are seriously divided (Badre et al., 2015). This section of the introduction will suggest ways to investigate both of these topics through the implementation of interactionist neuroscience and computational modelling. Specifically, the research paradigm outlined in this chapter combines spatial navigation tasks across species with models of reinforcement learning. However, spatial navigation is just one of many fields this paradigm may be applied to for investigating general mechanisms of information processing during cognition (Behrens et al., 2018).

This part of the introduction will focus on spatial navigation as a test-bed for computational hypotheses underlying cognition for the following reasons: 1) Navigation is ecologically sound. Humans and animals alike navigate daily through their respective environments; 2) Neurobiological representation of spatial information is well established from cellular single-unit recordings in the hippocampus and surrounding brain regions; and 3) Navigation tasks are generalizable across species. The approach described will investigate neural mechanisms which intimately links different levels of understanding, from low-level neural circuits to the overarching computations carried out by the brain.

The first question we consider in this section is: Does the hippocampus support coding of goal information? Studies using real world stimuli have identified correlates of hippocampal activity with distance and direction to goal using fMRI (Howard et al., 2014; Javadi et al., 2017). Similarly, the presence of egocentric goal coding has been found in bats which take long trajectories around the goal

potentially allowing enough angular sampling to be detected (Sarel et al., 2017). Additional evidence also suggests a small subpopulation of reward cells in the hippocampus of mice during virtual navigation (Gauthier et al., 2018). However, it remains to be seen if the same goal coding as identified in the bats is present in rats. The lack of evidence brings up the second question: Do rats and humans use neural mechanisms testable through reinforcement learning to flexible navigate? We hypothesize that a subclass of reinforcement learning models, the successor representation, is best at accounting for the flexible behaviour in the spatial navigation tasks tested in both rodents and humans in this thesis. This is based on recent theoretical work by Stachenfeld et al., (2017) suggesting that the hippocampus instantiates a predictive map which can explain findings such as the influence of reward onto place fields. Other researchers have also identified that successor representation can account for behaviour such as reward revaluation in value-based decision making by utilising mechanisms of offline replay (Momennejad et al., 2018).

Space and the brain

Early behavioural studies with rats navigating through mazes gave rise to the idea that mammals have an internal map-like representation of their environment (Tolman, 1948). Tolman's experiments showed that rodents extrapolate beyond simple stimulus-response mappings during navigation, and he hypothesized that they create a so-called cognitive map. This theoretical map is what allowed the rodents to choose previously unseen and more efficient paths (i.e., shortcuts) in a learned environment. Interestingly, Tolman viewed this cognitive map as a generalized mechanism of cognition, one which is used for rational behaviour (Tolman, 1948; Behrens et al., 2018). He further suggested that behavioural changes caused by disease or brain injury results from damage or narrowing in these cognitive maps. A few decades later, single-unit recordings in the hippocampus of rodents revealed cells which fire in specific spatial locations of the environment – known as place cells (O'Keefe and Dostrovsky, 1971; O'Keefe and Conway, 1978). Each of these so-called place cells has a place field associated with it, a spatial region in the 2D environment where the cell fires most often. The accuracy of these cells are such

that only a few need to be analysed to accurately determine the location of freely moving animals from the recordings alone (Moser et al., 2017). Place-dependent firing of single cells have also been discovered in bats (Ulanovsky and Moss, 2007), mice (Harvey et al., 2009), humans (Ekstrom et al., 2003), and non-human primates (Ludvig et al., 2004; Mao et al., 2021).

Following the discovery of place cells, another set of cells were discovered which further advanced our understanding of encoded spatial information critical for navigation. Head direction cells were found in the dorsal presubiculum and anterodorsal thalamic nucleus and encoded the current heading direction of the animal, irrespective of other behaviour (Ranck Jr, 1984; Taube, 2007; McNaughton et al., 1991). Unlike place cells which encoded environmentally-bound spatial information these head direction cells encoded egocentric information.This information was believed to be important for navigation in conjunction with the allocentric place cells, e.g. for path integration (McNaughton et al., 1991). Path integration is a central aspect of spatial navigation. It is the process by which information about current speed, position, and heading direction allows the organism to precisely calculate their current and future position (McNaughton et al., 2006). One well known issue of path integration is the accumulation of sensory errors due to the poor accuracy of heading direction and/or speed, resulting in, for instance, a case where people can walk in circles despite trying to walk in a straight line (Souman et al., 2009).

In the early 2000s, a new class of spatially selective cells, grid cells, were identified in medial entorhinal cortex, a brain region adjacent to the hippocampus (Fyhn et al., 2007; Hafting et al., 2005), and later in the pre- and parasubiculum (Boccara et al., 2010). In contrast to place cells, these cells encode multiple place fields that span the whole explorable environment in a symmetrical, hexagonal pattern (Moser et al., 2017). Grid cells also adapt to changes in environmental cues, suggesting a role of anchoring in navigation. When cue cards positioned on the inner walls of a box environment were rotated, the firing pattern of the present grid cells followed the same rotation (Hafting et al., 2005). In contrast to place cells, the place fields of grid cells remain constant across changing environments (Bostock et al., 1991;

Fyhn et al., 2007). Grid cells have also been identified in bats, non-human primates, and in mice (Yartsev et al., 2011; Killian et al., 2012; Fyhn et al., 2008). Recently, the presence of rewarding goal locations has been shown to alter the pattern of the grid field around the goal (Boccara et al., 2019; Butler et al., 2019).

Even more recently, cells encoding a boundary vector were discovered in the subiculum of freely moving rats (Lever et al., 2009), and in the entorhinal cortex (Savelli et al., 2008; Solstad et al., 2008), amongst other regions (Grieves and Jeffery, 2017). These cells fire when the animal is positioned at a preferred distance from an environmental boundary. Interestingly, these cells were hypothesized to exist as inputs to place cells based on the influence changes in environmental boundaries have on place field patterns and were also predicted by computational models (Barry et al., 2006). Neurons in the medial entorhinal cortex have also been identified that specifically encode the speed of freely moving rats (Kropff et al., 2015). These cells appear to be solely responsible for encoding speed, unlike many other conjunctive cells that encode speed along with grid fields or heading direction (Wills et al., 2012). An additional subpopulation of cells in medial entorhinal cortex were discovered to encode the vector to objects in the environment (Høydal et al., 2019).

The question remains how this encoded spatial information helps with navigation. For instance, spatial information can be used to execute several different distinct navigational strategies. Two of them are path integration and landmark navigation. Path integration, or dead reckoning, is a process whereby an agent uses self-motion cues like changes in velocity and egocentric heading direction to keep track of their position over time (Mittelstaedt and Mittelstaedt, 1980; Gallistel and King, 2011). Landmark navigation is navigation through the use of external cues like distance between objects in the environment (Yoder et al., 2011). Two other strategy domains are goal-directed and habitual behaviour (Dolan and Dayan, 2013). Goal-directed behaviour is considered a deliberate process requiring future planning based on information beyond immediate sensory cues (Pezzulo et al., 2014). Habitual behaviour constitutes previously learned stimulus-response associations

which may be disassociated from current value outcomes (Dickinson, 1985).

Goal-directed and habitual behaviour in rodents

A longstanding goal in the study of navigation, and of decision-making in general, has been to investigate the extent to which actions are chains of stimulus-response mappings, or whether actions depend on an abstract representation of a goal. The idea being that this goal representation allows the agent to act beyond current or past reinforcement of actions, for example by planning out and executing a shortcut (Tolman, 1948). Another important question is how goal locations are represented in the brain. Studies have shown that place cells will shift their place fields over time to encode reward at goal locations (Komorowski et al., 2009; Dupret et al., 2010; Hok et al., 2007). Using optogenetic stimulation, one study artificially activated dopaminergic neurons projecting to place cells after spatial learning (McNamara et al., 2014). They found an increase in place cell reactivation. Moreover, in the Morris water maze task, place fields were found to be over-represented around the goal location (Hollup et al., 2001). These findings of goal location encoding in place cells demonstrate that they encode more than a collective spatial code of the environment.

Rats exhibit vicarious trial and error behaviour, which means that they halt at a decision point and appear to be considering their options (Tolman, 1938; Redish, 2016). It was recently discovered that during this behaviour, place cells generate brief sequences of spatial trajectories which predict subsequent behaviour even in novel environments (Pfeiffer and Foster, 2013). These generated sequences occur during so-called sharp-wave ripples and consist of sudden high frequency activity in the local field potential during inactive wakefulness and sleep (Buzsáki, 2015). Pronounced theta oscillations in the hippocampus have been established since the earliest recordings of local field potentials (O'Keefe and Conway, 1978). Within each extracellular theta oscillation, populations of place cells have also been found to fire in forward succession of each other, referred to as a theta sequence (Foster and Wilson, 2007). This theta sequence has recently been found to encode goal-related information (Wikenheiser and Redish, 2015). Wikenheiser and Redish (2015) mea-

sured hippocampal place fields of rats during a foraging task. They found that the theta sequences reflected the subsequent rat path trajectories up until goal locations. Moreover, the sequences were longer during longer behavioural trajectories.

During sleep or resting, place cells have been found to fire in a sequence that predicts the following place cell activity while running on a novel linear track (Dragoi and Tonegawa, 2011). This phenomenon of preplay suggests that the current state of the hippocampus pre-configures the future encoding of state representations (Dragoi and Tonegawa, 2014). The activity of place cells in hippocampus also encodes trajectories to regions that have never before been visited (Ólafsdóttir et al., 2015). Rodents were shown rewards positioned at unreachable regions of the environment. The firing of place cells reflected trajectories both towards and away from these reward locations. Another well-established finding in the hippocampus is that of replay. Place cells will fire in a pattern sequence which re-enacts the firing patterns from previous active behaviour during sleep (Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996). Moreover, these replay periods were found to also be present quickly following spatial navigation while the animal was awake but not moving (Foster and Wilson, 2006).

Goal-directed and habitual behaviour in humans

Research into understanding the human brain has drastically increased in the past few decades with the advance of non-invasive brain imaging Making it suddenly possible to map out cognitive functions associated with specific brain areas. One primary imaging modality is functional magnetic resonance imaging (fMRI), a research technique successfully used to detect fluctuations in levels of blood-oxygenation as a function of increased metabolic demands (Poldrack et al., 2011). This setup still allows for navigational processes to be engaged despite being constrained and immobile (Epstein et al., 2017; Huffman & Ekstrom, 2019).

The division of goal-directed (map-based) vs. habitual behaviour (route-based) in humans has often been supported by studies of behaviour in humans and rodent (Iglói et al., 2009; Tolman, 1948). This division is also implicated to recruit separate neural circuits (Wolbers and Hegarty, 2010). One study found that participants

who used spatial landmarks when navigating had increased hippocampal activity, while participants who used a habit-based (non-spatial) strategy had increased caudate nucleus activity (Iaria et al., 2003). The same disassociation was found with participants virtually navigating around a town (Hartley et al., 2003), and during a route-recognition task with early-stage Huntington's disease patients (Voermans et al., 2004). One interesting line of MRI research has focused on expert navigators in the form of London taxi drivers (Woollett and Maguire, 2011). An early study found that the posterior hippocampus of London taxi drivers was larger than that of control groups, while vice versa for anterior hippocampus (Maguire et al., 2000). The level of experience, i.e. years as a taxi driver, was also correlated with grey matter volume in the posterior hippocampus as compared to bus drivers and controls (Maguire et al., 2003; Spiers & Maguire, 2006). That the size of the hippocampus is variable suggests that the anatomy of the hippocampus adapts to the requirements of everyday spatial navigation. Recent work on hippocampal volume and navigation ability in the normal population has found contradictory evidence where in several large sample studies there is no such association between hippocampal volume and navigation ability (Weisberg et al., 2019; Clarke et al., 2020; although see Brunec et al., 2019). It is possible the correlation observed in taxi drivers is due to being extreme outliers in terms of navigational ability (Weisberg & Ekstrom, 2021). Some more recent findings suggest the hippocampus encodes the path distance to goal, while the entorhinal cortex encodes Euclidean distance to goal. Howard and colleagues (2014) designed a task where participants navigated real-world routes and were later shown movies of the same routes. The task was structured so that the Euclidean distance and path distance to goal location varied separately. Posterior hippocampus was found to be associated with the path distance and the change in path distance during detours, while anterior hippocampus and entorhinal cortex was associated with Euclidean distance during passive navigation or when a new goal was presented, respectively. The posterior parietal cortex reflected the egocentric angle to goal direction.

Another set of experiments further investigated brain regions involved in goal

proximity and detour planning. One task had participants first view the position of a figure in a virtual environment and later used mental navigation to calculate the nearest route to the previously shown figure from a random starting position (Viard et al., 2011). They found that the hippocampus, medial prefrontal cortex and parahippocampus tracked goal proximity. Moreover, a subset of detour trials was associated with medial and ventromedial prefrontal cortex activity. Earlier studies have also identified medial prefrontal, subiculum, and entorhinal cortex associated with goal proximity during virtual navigation (Spiers and Maguire, 2007). Recently, the retrosplenial cortex was found to encode goal proximity when navigation familiar virtual environments, while hippocampus encoded goal proximity within novel virtual environments (Patai et al., 2017).

The finding that place cells encode future trajectories to goal locations (Wikenheiser and Redish, 2015) has also been investigated in humans using fMRI and virtual navigation. Brown et al. (2016) designed a task where participants learned goal locations associated with fractal images positioned along a circular maze. They were later prompted with one of the fractal cues before planning and following a route to an encoded location. The hippocampus encoded the location of future goals during the planning stage, and the same hippocampal encoding was present after navigation to that goal. Another recent study using virtual maze navigation also found hippocampal regions to reflect goal-specific planning (Kaplan et al., 2017). One study investigated future planning (Horner et al., 2016) using the previously described process of measuring grid cell-related activity in the entorhinal cortex in humans with fMRI (Doeller et al., 2010). They identified the same signal being present during virtual navigation and during imagined navigation, suggesting its role in route planning. Entorhinal grid cell-like activity also extends beyond navigation and represents locations in visual space (Julian et al., 2018).

The above review of how the rodent and human brain represents space still lacks a computational perspective to explain this wealth of generated data investigating goal-directed navigation across species. By using simulations and through the implementation of reinforcement learning models and freely moving virtual re-

ality technology, one chapter in this thesis attempts to bring these two research domains closer together through consideration of task design.

Chapter 2

Reinforcement learning in freely moving behaviour

The past decades have seen a rapid increase in the number of studies leveraging the power of cognitive computational modelling (Palminteri et al., 2017). These methods could be important tools in understanding and explaining latent variables and building blocks of cognition that may be unattainable through more traditional summary statistics approaches (Corrado et al., 2012). Methods leveraging computational modelling of choice behaviour can also help to bridge the gap between cognitive fields such as decision-making and spatial navigation (Kriegeskorte & Douglas, 2018; Summerfield & Tsetsos, 2012; Gahnstrom & Spiers, 2020).

This chapter describes two projects where the advancement lies in the methods used as data collection was limited due to the Covid-19 pandemic. The first study involves maze simulations for improved decoding of goal-related information. Good experimental design is crucial for advancing the science of any research field. As computational power has risen substantially in the 21th century, new ways to approach good experimental design have also gained popularity (Palminteri et al., 2017). There are many possible parameters that can be tweaked in any given experiment and it is much too costly for any experimental lab to pilot a large subset of these possibilities. This is where simulating experiments can play an important role in discovering a more optimised design to test your experimental hypothesis. This is especially important when utilising computational modelling of choice be-

haviour as you can simulate choice behaviour using your computational models prior to data collection (Wilson & Collins, 2019). If you discover that, for instance, parameter recovery is too noisy, then you may decide to change the design in terms of number of trials, conditions, or any other variable. You can then keep iterating on your experimental design until a desirable paradigm is reached. We utilise part of this approach in designing an experiment enabling us to investigate goal-related information in a complex navigation task for both rats and humans.

2.1 Study 1: Maze simulations for unbiased behaviour in a spatial navigation task.

The main research goal of this project is to find an experimental design capable of investigating goal coding in the rodent and human hippocampus. For this specific purpose we re-analysed rodent data from a land-based modified version of the Morris water maze task (de Cothi et al., 2020). In this task, rats trained in an open-field environment with a single hidden goal that stayed constant throughout all experimental sessions. The rats had one large distal cue placed directly outside the maze to assist with place-learning and orientation. Chocolate milk reward was delivered through a tube connected to a well at the hidden goal location. After reaching a learning criterion during training, the testing sessions began. The testing sessions imposed barriers in the environment which could take a wide variety of geometric shapes (Figure 1B). This was done by removing parts of the flooring, forcing the animal to use new routes to find the optimal trajectory to a learned location of the hidden goal. Each trial consisted of placing the animal in a new starting location, for a total of 10 trials per maze configuration (Figure 1B). The animal only had 45 seconds to reach the goal for a trial to be considered a success after which the animal would get picked up and placed in the next starting location to begin the next trial.

Our first aim was to find a task set-up which resulted in long and unbiased trajectories to the goal without being too difficult for the animals to complete. Our second aim was to ensure uniform sampling of the environment so future neural

recordings would be able to pick up encoding of egocentric goal information (similar to that of bats in Sarel et al., 2017). We found an indication the sampling of the environment in the rodent experiment may be too biased to detect goal-relation information. To address this problem, we made several experimental design changes and further simulated a vast range of possible mazes. This method was later used in subsequent experiments comparing neural recordings and behaviour in rats and humans performing the same complex navigation task (the human fMRI results are described in chapter 3).

2.1.1 Methods

Subjects:

A total of 6 Lister Hooded rats completed 10 trials for each of 25 distinct mazes (for a sample of these mazes, see Figure 1B). The trajectories were analysed to provide a basis for the biases in the sampling of the environment. The rats were trained and housed communally in batches of 3 with slight variations to the training protocol. The rats were food-restricted which was set to still maintain 90% of free-feeding weight. All procedures were according to UCL ethical guidelines and licenses issued by the UK home office according to the Animal Scientific Procedures Act of 1986.

Stimuli and task design:

The physical maze consists of 100 movable wooden modules organised in a 10 by 10 grid (Figure 2c). Every module measures at 20cmx20cm, making the total arena size 2m x 2m, and placed at a height of 52.5cm. The centre of each module was fitted with a sunken well. A tube was connected from the well to an experimenter-controlled syringe filled with the reward (chocolate milk). The goal location was always placed at the same location 4 tiles from the north edge and 4 tiles from the west edge of the maze. Every day a new randomly selected module was used for the reward delivery and placed at the goal location.

The maze was surrounded by a circular white curtain, obscuring all surrounding walls and the experimenter's computer. A single large black distal cue was used

on the north side of the curtain. However, for every day of testing the goal location and distal cue were rotated to reduce navigational strategies that did not rely on the black curtain as a landmark.

Training:

After familiarisation with handling and the maze environment, the rats started the training protocol. The rats were placed in an open field environment and had to fulfil a set of criteria before moving on to the testing stage. First, they were placed on immediately adjacent modules to the goal location. If successfully locating the reward, they were then placed on a module successively further away from the goal. If the rats could complete this from each cardinal direction (south, east, west, and north) and from up to 6 tiles away from the goal, then they were ready for the next stage.

Testing:

All six rats were tested on three different maze configurations per day. Each of the maze configurations consisted of 10 trials with starting locations at increasingly difficult positions as determined by increases in the ratio of path distance over Euclidean distance between the start and goal location.

Simulations:

New mazes were simulated to arrive at a principled way of selecting mazes instead of manually selecting mazes that may or may not give unbiased sampling of the environment. Given that we wanted to investigate goal-coding of single cell firing rates, we also included a second goal location for these simulations to be used in subsequent experiments. The second goal was to enable cells encoding the position of the goal to change their tuning in response to change in goal location during the same session of recording.

The number of possible permutations in a 10x10 binary two-dimensional array is

$$2^{10x10}$$

or

$$1.26x10^{30}$$

. Given this computationally intractable number of possible maze configurations in our task environment, we decided to use a heuristic approach. In order to constrain the space of possibilities, we created nine vectors with the dimensions of 3x1 modules. These vectors were randomly allocated to a position on the 10x10 grid with the possibility of overlap.The orientation of the vectors (vertical or horizontal) was determined using uniform probability (random coin toss). A total of 30,000 maze simulations were performed (Figure 3).

Metrics of unbiased sampling:

In order to investigate egocentric goal coding in the CA1 cells of the hippocampus, a uniform sampling of the egocentric head direction of the animal is crucial. Large biases due to the way maze configuration constrain the possible trajectories to the goal may result in inconclusive results. We investigated the uniformity of egocentric angles using the Kullback-Leibler Divergence (Kullback & Leibler, 1951):

$$D_{KL} = (P\|Q) = -\sum_{x\varepsilon X} P(x) log(\frac{Q(x)}{P(x)})$$

Where P and Q are the probability density functions of two distributions. The KL divergence gives an information-theoretic metric of the amount of information lost (in bits) when trying to approximate one distribution P given a different distribution Q. If these two distributions are identical then DKL = 0. In our case, the distribution of observed egocentric angles in our maze simulations is compared to a perfectly uniform distribution of angles. The closer the observed distribution is to 0, the less biased we argue those observed angles to be.

Calculating the optimal path to goal:

Two different algorithms were used to calculate the optimal paths to goal from all starting locations. For analysing the rodent data, the A* algorithm was used which is a type of best-first search (Russell & Tolvig, 2010) and implemented in Matlab (Mathworks Inc.). However, when comparing all of the maze simulations, Dijk-

stra's algorithm was used and implemented in Python. Both approaches are very similar with the only difference that A* uses an additional heuristic function which estimates the cost of a solution.

Maze similarities:

One concern for the maze simulation was that the best simulations would all result in the same or similar mazes. In order to test for the similarity across the top ranked maze configurations, we computed a range of similarity measurements including the following: Cosine Similarity, Euclidean Distance, Sum Similarity, Pairwise Manhattan, Pairwise Linear, and Structural Similarity Index (SSIM). The SSIM is the most likely candidate for us to group together similar mazes because it is based on the perceptual similarity between two images and was originally developed to assess image quality (Wang et al., 2004).

2.1.2 Results

Rodents approach optimal trajectories during the course of navigation:

Before analysing the task design in terms of sampling for the detection of goal-related information, we wanted to better understand the behaviour. Specifically, we wanted to investigate the relationship between observed rodent trajectories and how an optimal agent would perform. In order to make the trajectory length of the rodents and of the optimal agent comparable, we normalised both by the euclidean distance from the starting point for each trial. In the first batch of rats (n=3) we found a significant correlation between path length and optimal path length (Rat 1: Pearson's r = 0.5, $p < 0.001$; Rat 2: Pearson's r = 0.4, $p < 0.001$; Rat 3: Pearson's r = 0.6, $p < 0.001$).

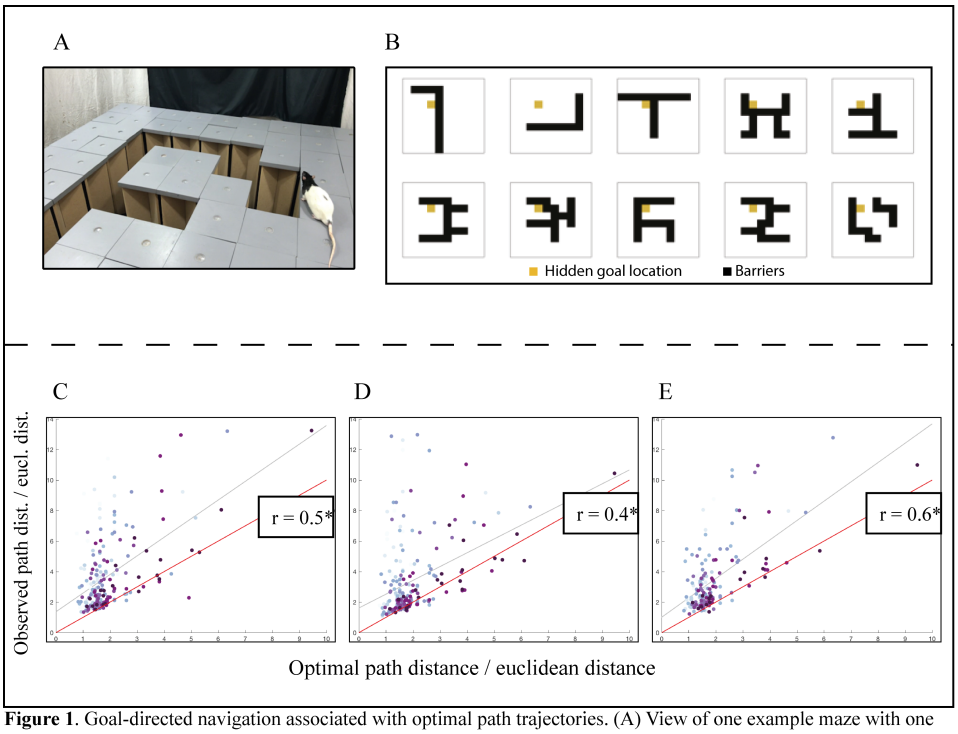

Figure 1. Goal-directed navigation associated with optimal path trajectories. (A) View of one example maze with one of the subjects. (B) Showing 10 out of the 25 maze configurations used along with the goal location. (C) We calculated the optimal path distance for each starting location in each maze and normalised by the euclidean distance in grid space. We then correlated that with the observed path distance and normalised by the euclidean distance in real space within the same trial. The grey line reprents the best fit through all the data points, while red line represents the unit line. The hue of the purple color for each data point indicate the progession of trials where deepest purple is the final trial collected. The correlation observed for rat 1 (Pearson's r = 0.5, p < 0.001). (D) The correlation observed for rat 2 (Pearson's r = 0.4, p < 0.001). (E) The correlation observed for rat 3 (Pearson's r = 0.6, p < 0.001).

Figure 2.1: Goal-directed navigation associated with optimal path trajectories

Egocentric bias in sampling across maze configurations:

The experimental data was collected across two batches with 3 rodents in each batch (total n = 6). We compared the performance of both batches to check if the behaviour was comparable. The two batches were concluded to have similar performance in terms of path distance travelled during successful trials given the overlap in error bars (Figure 2A). However, there were still some important differences such as the time taken on successful trials was non-overlapping, where the first batch of rats were on average slower than the second batch (Figure 2B). Overall, the egocentric angle sampled for each of the 6 rats was fairly consistent when averaged across all maze configurations (Figure 2C).

We investigated the uniformity of egocentric angles by calculating the Kullback-Leibler divergence (see methods for definition). Over the course of the 25 mazes, there was considerable variability in this uniformity and large deviations

from zero (Fig. 3A), as zero would indicate perfect uniformity. Moreover, inspecting the mean egocentric angle of each rat per maze configuration also indicated bias in the sampling of egocentric angles to goal (Fig. 3B). We used this observed variability to motivate that a different set of mazes were necessary in order to ensure a more uniform sampling of the egocentric angles in our upcoming experiments. The most important aspect was to design maze layouts which maximised our ability to detect goal-related information in the firing of hippocampal place cells.

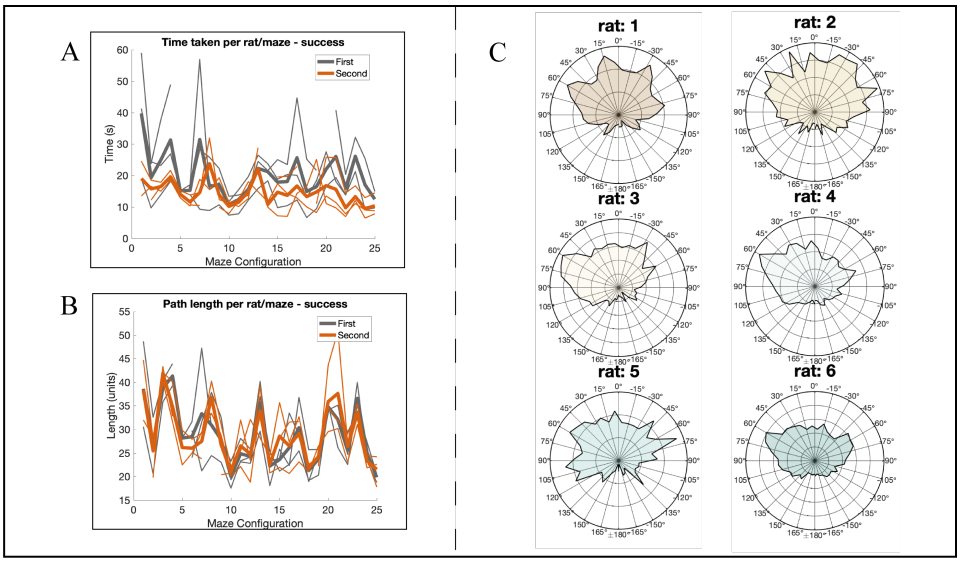

Figure 2. Comparing behaviour across batches and across rats. (A) We compared the time taken to reach the goal on successfull trials between our two batches of collected data. (B) We also compared the path lengths of our two batches across all maze configurations. (C) Polar plots showing the egocentric angle sampling per rat across all maze configurations.

Figure 2.2: Comparing behaviour across batches and across rats

Maze simulations of a new task design:

We made one major design modification with the introduction of a second goal location in the environment. We wanted to be able to show clear switching of goal-related information in spiking activity within a session. In the previous version of the task, there was one constant goal location throughout the entire experiment which could possibly explain part of the bias in egocentric angles to goal. By adding a second rewarded location, we also faced more difficulties in manually constructing maze configurations that would fit all our criteria for a well designed experiment. Below we describe the initial mazes simulated using this approach. However, the final version of our maze simulations and the final task design is described in chapter 3 where we investigate goal coding in humans using fMRI. The final mazes described there are also being used in a parallel rodent experiment which is outside the scope of this PhD thesis.

For the new version of the task we simulated optimal agents on 30,000 mazes (Fig. 4A). In short, the optimal trajectories were computed with Dijkstra's algorithm from each starting location to the position of the two separate goals. These were then up-sampled and smoothed using Savitzky–Golay filtering to get more re-

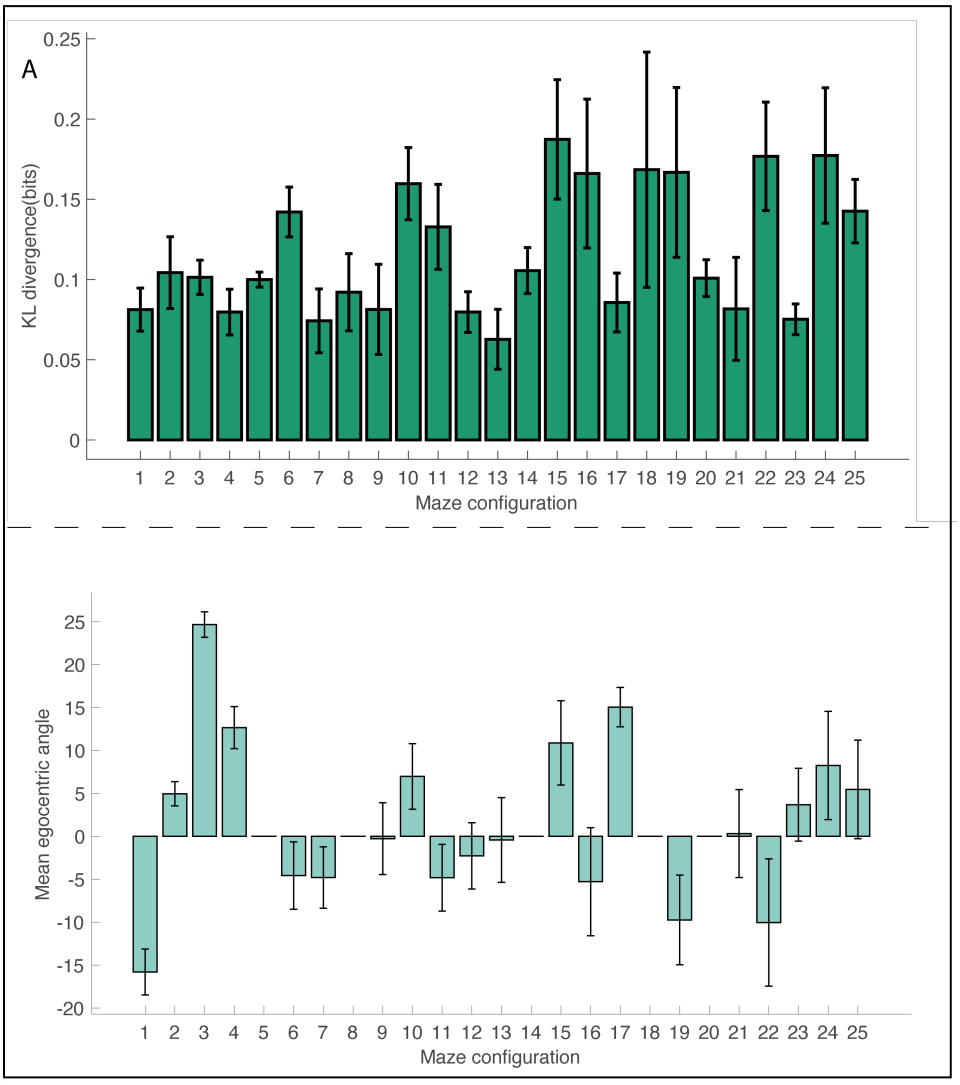

Figure 3. Biased sampling across maze configurations. (A) We computed difference between a uniform distribution and the observed distrubution for each rat on each maze using KL divergence. (B) We identified large consistent biases in certain maze configuraitons across all 6 rats on those maze configurations.

Figure 2.3: Biased sampling across maze configurations

alistic trajectories (Savitzky & Golay, 1964). We aimed to select mazes that best allowed for the separation of Euclidean and path distance coding. To do so, we subsequently ranked all generated mazes according to the average of three metrics using all optimal trajectories:

1. The egocentric heading angles with respect to the goal locations.
2. The path distance divided by the Euclidean distance between the starting position and goal location.
3. The number of tiles removed from the environment.

Each metric was given an equal weight in the ranking. We prioritised mazes that provided uniform sampling of angles, a large path distance over Euclidean ratio to give long trajectories, and a large number of tiles available for the humans and rats to explore. The top 16 mazes selected are shown in Figure 4B.

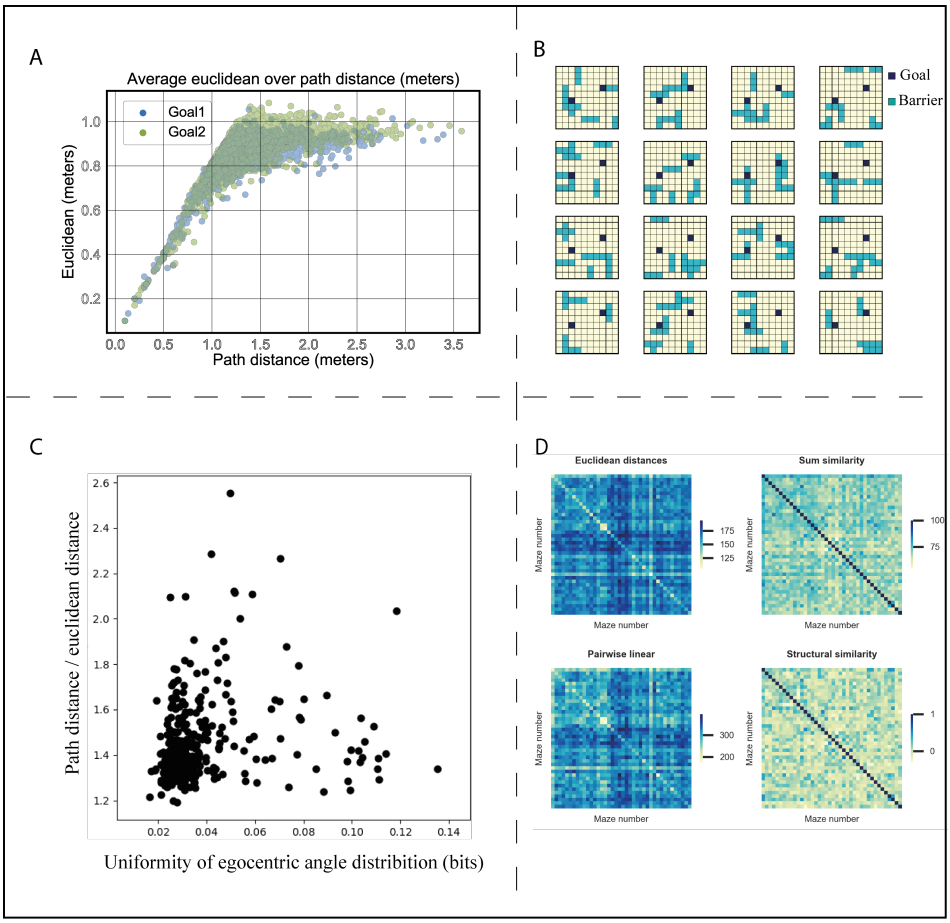

Figure 4. Maze simulations, maze similarity and final ranking. (A) Each circle represents the average euclidean distance and the average path distance from all possible starting locations, plotted separately for each home goal (two circles per maze). Preferred mazes are at the tail-end of the distribution where we maximise average euclidean and path distance. (B) Highest 16 ranked mazes out of 30,000 simulationed maze layouts based on three metrics; uniformity of egocentric angles, ratio of path distance / Euclidean distance, and number of tiles removed. (C) Distribution of uniformity of egocentric angles of the optimal simulated agent, measured in bits using Kullback-Leibler divergence, plotted against the ratio of path distance over the Euclidean distance. The metrics are averaged over both goal locations per maze. (D) Four different metric of maze similarity for the top 40 mazes out of the complete set of simulations. There is low level of similarity across mazes compared to the diagonal – which represents each maze similarity with itself.

Figure 2.4: Maze simulations, maze similarity, and final ranking

2.2 Study 2: Reinforcement learning and freely moving navigation during changing environments.

The main goal of this study was to train reinforcement learning agents on the same spatial navigation task in order to compare navigational strategies used in freely moving behaviour in both rats and humans. As covered in the introduction, many aspects of allocentric and egocentric spatial information are encoded in the hippocampus and neighbouring brain areas. However, the question remains how this encoded information enables the computational processes necessary for success-

ful navigation of the environment. Looking back to the different levels of analysis approach suggested by Marr (1982), several questions arise: 1) What are the computational goals used in navigation?; 2) Moreover, what are the algorithms of information processing of those goals?; and 3)What is the biophysical implementation of the algorithms? This section will focus on reinforcement learning algorithms currently suggested to be implemented by the brain during navigation. These algorithms cover two levels of Marrs of the computational goal (reward optimization) and the algorithmic level (Niv, 2009). This study aims to suggest the ways we will investigate these algorithms (and others) by combining research methods across computational modelling, rodent electrophysiology, and human neuroimaging.

The distinction between goal-directed and habitual behaviour goes back to the early days of behaviour research and was investigated as stimulus-response mappings and Tolman's cognitive map hypothesis, respectively (Tolman, 1948; Dolan and Dayan, 2013). In more recent years, these distinctions have crossed over into computational modelling of behaviour in the form of model-based algorithms for goal-directed behaviour and model-free algorithms for habitual behaviour (Daw et al., 2005; Foster et al., 2000). Moreover, this division is demonstrated in both rodents and humans but may be interacting instead of competing (Balleine and O'Doherty, 2010). The model-based algorithms use experience to construct an internal model of the environment, while model-free algorithms learn the value of the stimulus response mappings of available states and actions or policies (Dayan and Niv, 2008). Importantly, these different classes of algorithms can be used to test specific hypotheses about neural activity and lead us to identify neural mechanisms of behaviour (Daw et al., 2011). Again, the aim of this chapter is to provide the theoretical grounds for investigating generalizable mechanisms of cognition, and the framework of reinforcement learning is central to this approach.

We investigate three reinforcement learning models: (1) model-free algorithm (Q-learner) which is successful at reaching human-level performance in a range of tasks from a suite of ATARI games when combined with deep learning (Mnih et al., 2015); (2) Successor representation (SR) which is more flexible at adapting to en-

vironmental changes than the model-free algorithm and internal components show similarities to findings in place and grid cells (Stachenfeld et al., 2017; Russek et al., 2017); and (3) Monte-Carlo tree search which can achieve superhuman performance compared to professional chess and Go players when combined with deep neural networks (Silver et al., 2018). We hypothesise that successor representation is best at capturing human flexible navigation in our task in line with recent evidence from de Cothi et al. (2020). However, in that study, the human trajectories were also well captured by the A* model-based algorithm. Our MCTS model-based algorithm may have different predictions from the A*implementation. Nevertheless, we expect that model-based trajectories are more similar to human behaviour than model-free trajectories.

2.2.1 Methods

Participants:

A total of eight healthy human participants (two male, M = 23.6, SD=4.72) completed a wireless and freely moving immersive virtual reality task (in either 6mx6m or 10mx10m space). A different set of mazes were used than those from the rodent mazes and maze simulations from Study 1 in this chapter. All participants were recruited on SONA, the UCL undergraduate channel, or through word-of-mouth. Participants were compensated with £10/hour. All participants had normal or corrected to normal vision. Ethical approval was granted by the UCL Department of Psychology Ethics Committee. Each session took approximately one hour.

Virtual Reality Set-Up:

The HTC Vive Pro Eye was used together with four 2.0 lighthouse base stations from Valve Inc. These were each placed in the corners of a 10x10m area cleared from any physical obstacles in a large lecture hall (Figure 1A). When piloting in an area larger than 10x10m, the head-mounted display lost connection near the centre of the environment, enforcing a strict upper limit on the bounds of the space. A wireless transmitter was placed on top of the computer monitor next to one of the corners in the walkable arena. A wireless receiver was attached to the head-

mounted display to enable freedom of movement. A windows PC with a GTX 1060 Ti graphics card was used for rendering the virtual environment in real-time.

The environment was designed and implemented in the Unity3D games engine (Fig. 1B, C). Custom scripts in C Sharp enabled the dynamic changes in maze configuration and goal locations. The data organisation and scripts were based on the Unity Experiment Framework (UXF: Brookes et al., 2019) which reduced frame rate drops by using separate CPU threads for saving data.

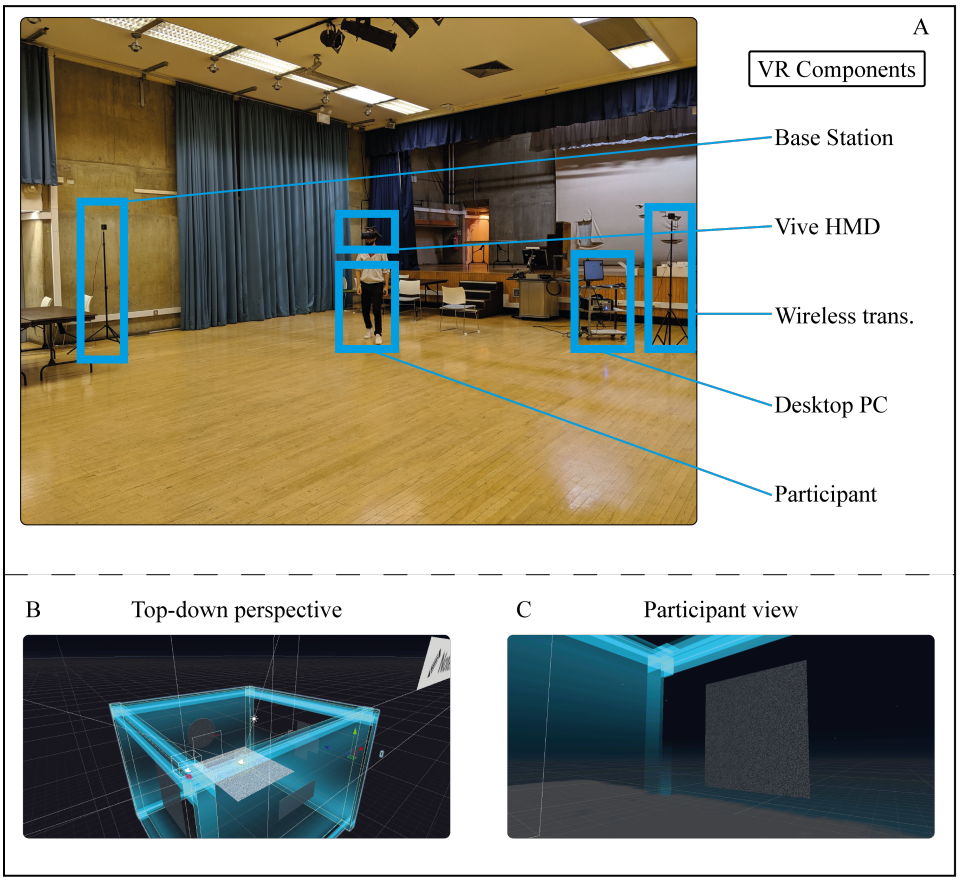

Figure 1. Freely moving virtual reality task set-up. (A) 10x10 meter space was utilised for freely moving virtual reality. Sveral mobile components were set-up and removed after each experimental session, including 4 base stations, wireless transmittor, desktop PC, and HTC Vive Pro Eye Head-Mounted Display with a wireless adaptor. (B) The virtual environment as rendered inside the Unity 3D game engine editor mode. Participants are suspended on a platform in outer space surrounded by four semi-transparent walls. Each wall has a cue on it represented by a different shape. (C) First-person perpective for participants in the environment. The platform they are standing on has transparent fog and so the missing parts in the environment is still partly visible.

Figure 2.5: Freely moving virtual reality task set-up

Stimuli and task design:

The human participants experienced six mazes during one experimental session, taking around 60 minutes to complete. The human task was implemented using fully-immersive virtual reality. The participants freely navigated within a 10x10m room, using a wireless HTC Vive Pro Eye™ head-mounted display to see the task environment (Figure 1). The task was modelled after a rodent experimental paradigm (Pfeiffer & Foster, 2013) with a primary design deviation being that we perturbed the possible routes one can take to a goal location through careful manipulation of the environment (Figure 2A). There were two types of trials: Exploratory trials, where the goal would be hidden at a random location in the environment and

Homing trials, where the goal would be in one of two possible locations. This was to facilitate full exploration of the environment and to have a way of controlling the next starting location without the need to physically move the participant (either human or rat).

To force participants to take new routes to successfully locate the goal, we removed parts of the walkable arena in virtual space. All participants completed 70 trials in total, with the first 20 trials included as training within an open-field environment (no modules removed). This training period was designed for participants to learn the location of two consistent Homing goals – similar to the training protocol used for the rats. The goal locations were hidden from participants until moving near enough to trigger goal consumption when a spinning golden crystal appeared with the text "You got the reward!".

The maze configurations used for this experiment were manually designed based on a few heuristics. We did not use the maze simulation approach described in the previous study because this data collection was part of a pilot project as a proof-of-concept. We designed the mazes on the basis of blocking off previous routes that would result in successful navigation. For instance, if your starting location is in the top right of the first maze, you can reach the goal by going west, south, then east. In the subsequent maze, you will face a barrier if you take that same path. Instead, you need to continue south, turn east, and lastly north to reach the goal. We followed this simple approach throughout the design of the first four mazes. For the final two mazes we also wanted to add the possibility of traversing the space in between the two goal locations.

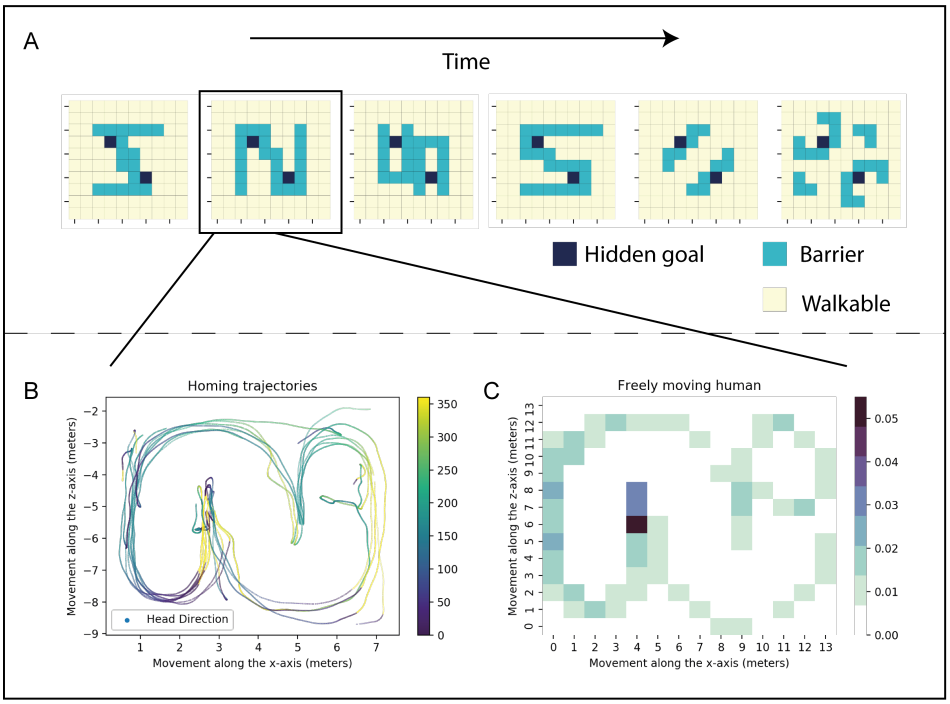

Figure 2. Freely moving human trajectories. (A) All six mazes used for the freely moving task. Each maze consisted of 10x10 grid with two goals on two different grid locations. Participants alternated between navigation to a random location and navigating to a hidden goal location. Each participants experienced all mazes in the same order and the hidden goal location changed halfway through each maze. (B) Example trajectories for one participant in the second maze configuration during homing trials where participant is navigating to the hidden goal locations. Head direction is color coded in allocentric coordinates. (C) Heatmap of the same data plotted in (B) segemented into 13x13 spatial bins indicating the overrepresentaiton in spatial occupancy at the hidden goal location compared to other spatial bins in the environment.

Figure 2.6: Freely moving human trajectories

Metrics for behavioural analysis of observed trajectories:

We used four behavioural metrics to quantify all the trajectories: tortuosity, diffusivity, linear displacement to goal, and angular displacement to goal. These metrics have been used previously to quantify and compare RL agents with rodent and human behaviour (de Cothi et al., 2020). We define these metrics in detail below.

Tortuosity: The amount of curviness of a trajectory. We calculated tortuosity by computing the total path distance travelled on each trial and divided that distance with the Euclidean distance from the starting location to the goal location.

Diffusivity: The amount of dispersion within an event horizon where a low value indicates a random walk. We calculated diffusivity by taking the average Euclidean distance travelled within three time-steps.

Linear displacement to goal: The amount of deviation from a Euclidean direct path to goal. We calculated the linear displacement by taking the average Euclidean

distance to goal for each time-step measured.

Angular displacement to goal: The mean egocentric angle to goal where value of 0 indicates that the agent is always heading toward the goal, on average. We calculated the angular displacement by calculating the egocentric heading angle per time point, and then averaging across all time points.

Mahalanobis distance to compare multiple trajectory features of RL agents against the human trajectories:

We used the Mahalanobis distance metric to get a sense of the deviation of trajectories observed with RL agents to those of the human participants. We chose this metric because it accounts for correlation between the different metrics and outputs the distance of a point from a multivariate distribution in terms of number of standard deviations (de Maesschalck et al., 2000). Mahalanobis distance is defined as follows:

$$D = \sqrt{(x-\mu)^T \times C^{-1} \times (x-\mu)}$$

Where x consists of our metrics from the RL agents, and μ consists of the mean of the behavioural metrics from the human trajectories and C is the covariance matrix of the human trajectory metrics.

Reinforcement Learning models:

There are a range of reinforcement learning models that can be trained and compared against each other in our task. Along the continuum of model-free and model-based algorithms, two models were implemented: a Q-learner and a successor representation (SR) learner using the SARSA algorithm. An analysis pipeline for investigating the behavioural trajectories was developed using the Python Gym framework with custom scripts for environments (Brockman et al., 2016). This framework allows for comparing not only real trajectories but also those of simulated artificial agents. These agents each use a specific implementation of a reinforcement learning (RL) algorithm.

All reinforcement learning agents make use of three quantities to estimate a

value function of all states: the current action from the set of possible actions, the current state from the set of all possible states, and the rewards encountered (Sutton et al., 1998, 1999). Additionally, it is necessary for computational tractability that these agents fall under the definition of a Markov Decision Process (MDP), and can therefore be said to possess the Markov property. This means that the next state and expected reward of the agent can be predicted based on the current state and action. The past states and actions can be ignored. Furthermore, the reinforcement learning agents considered here have a set of hyper-parameters which define the degree of discounting onto future rewards (gamma), the degree of learning which takes place given each experience (learning rate) and the trade-off between exploiting what you have learned and choosing to randomly explore the environment (epsilon).

Model-free agent:

The Q-learner agent infer stimulus-response mappings based on previous history of interactions within the environment. These mappings are updated based on the difference (prediction error) between what was expected, the predicted reward, and what was experienced, the actual reward. The model-free agent trained on our task used a Q learning algorithm which is part of the Temporal Difference approach.

The Q-based agent learns the value function of the environment in terms of all possible state-action pairs:

$$Q(s,a) = Q(s,a) + \alpha\delta$$

Where the prediction error is defined as:

$$\delta = R(s,a) + \gamma Q(s',a') - Q(s,a)$$

The $Q(s,a)$ essentially functions as a look-up table for all possible actions in all possible states. While the agent is learning the environment this look-up table is updated until eventually it can converge on the optimal (Barto and Mahadevan, 2003). Each action is then taken using an epsilon-greedy policy. This means that the agent will take the highest value of the look-up table in each state and use that

to determine its action. However, the epsilon parameter (often tuned to 0.1 or 10%) will choose to take a random action at some threshold decision points.

Successor Representation agent:

The successor representations (SR) algorithm uses a different approach to estimating the value function to both model-free and model-based agents. The SR is originally an extension of the TD and was shown to improve the flexibility of behavior during navigation (Dayan, 1993). The SR decomposes the value function into two components, the successor states matrix M and a vector of the future state reward R.

$$V_s = \sum_{s'} M(s,s') \times R(s')$$

The successor states are defined as the cumulative discounted future state occupancies:

$$M^{\pi}(s,s') = E[\sum_{t=0}^{\infty} \gamma I(S_t = s') | s_0 = s]$$

Where $I(S_t = s')$ is one if a state can be transitioned into; otherwise, it is zero. Parameter y is the discount factor gamma. The successor representation updates its representation of M using a similar process to temporal difference updating. However, instead of a reward prediction error, it uses a successor state prediction error (Stachenfeld et al., 2017).There are several approaches to implementing a successor representation agent, including using off-line policy updating which is a future direction of this work (Russek et al., 2017). In our current case, we used the on-policy TD method SARSA which stands for state-action-reward-state-action to update M. The agent is moving from state-action pair to state-action pairs instead of from state to state. Moreover, the SARSA method is policy-based which is the fundamental difference between Q-learning. SARSA may also be applied to model-free RL agents.

The SR-SARSA is defined in the following equation, where the additional parameter α_M controls the learning rate:

$$M(s_t, s_{t+1}, a_t) = M(s_t, s_{t+1}, a_t) + \alpha_M[I_{s_t=s_{t+1}} + \gamma_M(s_{t+1}, s_{t+2}, a_{t+1}) - M(s_t, s_{t+1}, a_t)]$$

The key advantage with SR is the separation between the long run future occupancy map (the SR) and the reward function. A change in goal location can be easily accounted for without the need to completely relearn the SR.

Model-based agent:

We implemented a Monte-Carlo Tree Search (MCTS) algorithm which uses an explicit model of the environment to make action selections (Sutton & Barto, 2018). The MCTS algorithm generally consists of four separate steps (Browne et al., 2012): (1) Selection of child nodes through already explored states, according to some selection policy, in our case using Upper Confidence Bound (UCB; Auer et al., 2002), until you reach an unexplored state; (2) Expansion of nodes according to available actions; (3) Run simulation from the newly expanded state(s) and choose actions according to a random policy until the goal is reached; and (4) Backpropagation, using the outcome of the simulation in step (3) to update the statistics of each node, for instance the average reward of each state.

In summary, we ran simulations per trial, where the agent chose actions leading to highest rewarding end state based on it's previous simulations of the environment. We limited the number of time steps to 100 and the number of simulations to 200.

Random and optimal agents:

We implemented two non-RL agents that did not take into account any learning across trial episodes or across mazes. These were used for sanity checks and to create an upper and lower bound on performance. For the lower bound, we implemented a random walk agent which for each state it visited would take a random selection from the available actions possible. There was no time limit and so it could take a very long time for this agent to find the hidden reward location. For the upper bound, we implemented an optimal agent that had perfect knowledge of each maze configuration. We used Dijkstra's algorithm to compute the optimal actions to

take in each state of the environment resulting in optimal path trajectories (Dijkstra, 1959).

2.2.2 Results

We only completed a marginal part of the planned data collection for this project due to the Covid-19 pandemic. As such, instead of analysing each participant in isolation we appended all the human trajectories together even though this ignores individual variability which is in large part why computational modelling can be a useful tool to begin with (Wilson & Collins, 2019). We treated this project as a proof-of-concept where most of the advancement lay in the methods developed, more so than the results and final analyses. We also perform no statistical tests to identify significant differences in trajectories of models and humans but instead rely on inspecting qualitative differences to draw tentative conclusions and future suggestions from the data.

Comparing trajectory characteristics between humans and RL agents:

We pooled together all the experimental data from the freely moving participants resulting in 390 trajectories. In order to compare with the simulated agents, we similarly extracted 390 trajectories from each of them (Q-learner, MCTS, SR, Optimal, Random Walk) based on the order of starting locations and maze configuration experienced by the participants. For each RL model where learning was involved, we reset the learned parameters (e.g. cached state-values) after each matched human session.

We found clear differences between the trajectory metrics of the models and human trajectories. For instance, the tortuosity for humans was much lower than that of the SR agent, while the same metric was much higher than that of the MCTS agent, indicating that the tortuosity of human trajectories lay somewhere in between these two RL agents (Figure 3A). Interestingly, for the angular displacement about the goal (or mean egocentric angle per trial), we saw a large variability in values across trials for both human and MCTS trajectories, while the distribution of other agents' trajectories were more stereotyped (Figure 3B).

Lasly, we calculated the mahalanobis distance as a measure of the multivariate deviation for all the different agents from the observed human trajectories (Figure 3C). We observed that there was overlap in the distribution of several agents, especially the Q-learner and MCTS, while the SR appeared further away from human trajectories. Overall, the SR agent seems to be close in multivariate space to the random walk and optimal agents (see Figure 3).

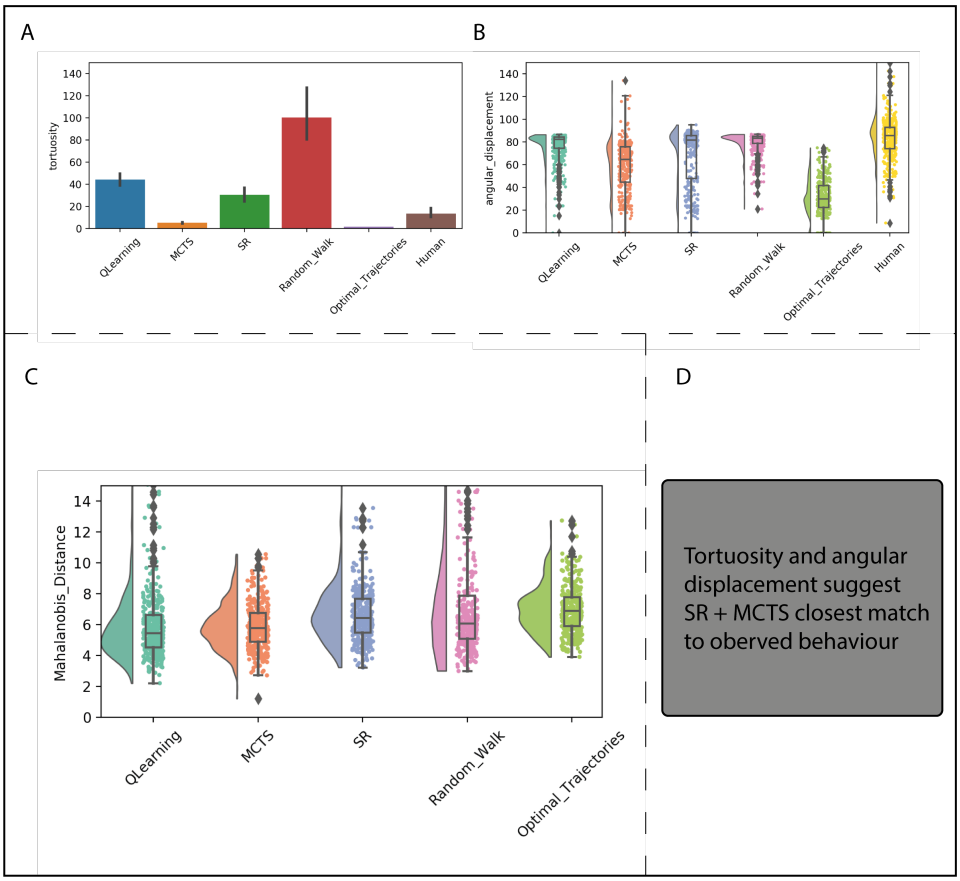

Figure 3. Computational modelling comparisons across RL agents and human trajectories. (A) Tortuosity was calculated for each trial for each agent and human trajectory. We calculated tortuosity by taking the total path length divided by the Euclidean distance to goal from the starting position. (B) Angular displacement was calculated by taking the average egocentric angle to goal for each time point and then plotted per trial for each agent and human trajectory. (C) We computed the mahalanobis distance for each agent which compares the multivariate trajectory metrics (tortuosity, diffusivity, linear displacement, and angular displacement) of the models with those observed during freely moving human behaviour. (D) Summary text: Two spatial metrics indicate that human trajectories are closest match to SR and MCTS models. However, taken together in multivariate mahalanobis distance suggest difficulty to disentangle different models.

Figure 2.7: computational modelling comparisons across RL agents and human trajectories

2.3 Discussion

In this chapter, we have presented two studies that advance methods in spatial navigation in separate but complementary ways. The first study suggests a principled way for improved design of complex task environments. In the future, we can add additional metrics of interest to create a new set of maze configurations. For the second study, we created a test-bed for different reinforcement learning agents on the simulated mazes which in future studies can be used to identify the mazes best at dissociating different types of learning algorithms. In fact, this might be an important way to demonstrate our experimental set-up is capable of identifying which

reinforcement learning agents most closely align with rodent and human behaviour.

In the second study, we found that in our task we were unable to confidently dissociate the different candidate reinforcement learning models when comparing human trajectories to the combined trajectory similarities across multiple metrics. However, two metrics, the tortuosity and angular displacement of a trajectory, indicate that the model-based algorithm may result in more similar behaviour, on average, to the human participants compared to either a model-free Q-learner or to the successor representation. We did not find support for the hypothesis that participant behaviour was best fit to the successor representation algorithm.

Our analysis depends on the assumption that a given participant retains the same behavioural strategy throughout the whole experiment. However, it is likely participants may switch between strategies on a trial-by-trial basis depending on contextual demands. For instance, early in learning, participants may rely on random walk exploration of the environment. Later after initial discovery of the maze layout and hidden goal location, they may rely on a model-based algorithm to efficiently plan and navigate (more equivalent of a place strategy) and later in learning switch to a computationally less demanding algorithm such as the Q-learner (more equivalent of a response strategy). Given the difficulty of assigning clear roles to brain regions (as independent modules) such as the caudate and hippocampus to response vs. place strategy, respectively, reframing these strategies as behaviorally-fitted computational models can better explain the findings. This method may exist as complementary to treating brain regions as part of circuits or networks involved in flexible navigation (Ekstrom et al., 2017).

Another likely explanation for the difficulty in dissociating models is that the metrics extracted from the model trajectories were too similar indicating the task may not sufficiently discriminate between different learning strategies. Participants may also employ other strategies to solve the task based on optic flow or path intergration. To remedy this, one could combine the approaches utilised in study 1 and study 2 and simulate task paradigms that should, in theory, separate different learning strategies as implemented through reinforcement learning models before

collecting behavioural data from human participants. Originally we planned to collect a full dataset of human participants, but the start of the Covid-19 pandemic stopped any in-person testing which our freely moving paradigm required. It is also possible that with more behavioural data we could have sufficient power to differentiate predictions from different reinforcement learning models.

The implementation of freely moving virtual reality opens up a world of possibilities in terms of the kinds of tasks we can use to investigate human cognition and has already been successfully used in studying boundary-related information in human medial temporal cortex (Stangl et al., 2021). It remains to be seen how precisely these behavioural data in freely moving participants can be applied to compare navigational strategies across species, but we now have the tools to do so at our disposal. Beyond testing humans and rats on the same task protocol, we can test reinforcement learning models, such as successor representation, against human and rodent behavior. Beyond behaviour, we can expand these models to fit signatures in the neural activity of single-units in the rat hippocampus as well as in the BOLD time series of human fMRI.

Limitations of current approach

Several of our analyses compared observed trajectories to that of an optimal agent. However, it is unclear the extent to which we should expect humans or rodents to approximate optimal performance in spatial navigation tasks especially where regularities in the environment are unstable, such as changing the hidden goal location or altering the traversable space. Known biases in relative distance estimates such as forward facing bias may also be important to account for as seen in judgment of relative direction tasks (Huffman & Ekstrom, 2019). One recent study found that rats form subgoals at the edges of environmental barriers during escape behaviour and during food-seeking to a learned goal location (Shamash et al., 2021). These subgoals also remained in use after the removal of the barrier, if only for a limited time. This suggests that rodents are neither optimal nor habitual and instead rely on a hybrid process more akin to hierarchical representations of the state space (Botvinick et al., 2009). The use of optimal performance can still be informative

as an upper bound on what is physically feasible in a given task, and the specific instances where performance deviates from the optimal upper bound can be used to inform important behavioural criteria for future models.

One major challenge in spatial navigation is the interpretation of behaviour as showing evidence to the formation of a cognitive map. One of the central tenets of cognitive map hypothesis is the ability of rats and humans to undertake novel shortcuts (Tolman, 1948). However, this behavioural phenomenon has proven difficult to replicate in the absence of maze cues (Benhamou, 1996; Olthof, et al., 1999; Prados, et al., 1999; Grieves et al., 2013) and it has been suggested that use of the term be abandoned completely (Bennett, 1996). In one of the original studies Tolman et al. (1946) trained 56 rats on a circuitous path through a maze. After blocking the learned path and opening 18 alternative paths, the most commonly chosen alternative was on a direct line towards the learned goal location. Importantly, the goal location was located near an overhead light that allowed the use of beaconing, confounding the experiment and calling into question its conclusions. One experiment replicating the Tolman maze in humans also failed to find evidence for novel shortcutting in the absence of a beacon light cue near the goal (Wilson & Wilson, 2018). More recently, real-world tracking of bats suggests evidence for novel shortcuts in mammals (Harten et al., 2020).

Reinforcement learning models have generated a wealth of theoretical accounts of experimental findings in neuroscience (Stachenfeld et al., 2017). However, it is important to consider the limitations of such models and the assumptions which they make (Eckstein et al., 2021). The inferences regarding underlying relationship between parameters of a particular model may only be informative for the current task at hand and not translate to other tasks with differing regularities. It is therefore relevant to interpret the parameter of a cognitive model within the context in which it is fit to participant behaviour.

Moreover, reinforcement learning models are built upon the mathematical notion of states satisfying the Markov property. States of the environment function as a signal which may be constructed, changed or learned in some way (Sutton &

Barto, 2018). For instance, a state signal is Markovian if it encompasses all of the relevant past sensory history of the agent. A concrete example would be a 2-D coordinate point in a maze, as there is no need to know the previous position to know which actions are currently possible. Another example of a state signal which is Markovian are the current positions of all chess pieces on a chess board. If the rules of chess necessitates knowing previous positions of chess pieces, then the state of the environment would be non-markovian. Importantly, the Markovian state signal is all that is necessary to be able to predict subsequent states and reward. Even non-markovian processes such as weather which may depend on additional factors beyond the current state, can still be considered as an approximation to a Markov state within reinforcement learning (Sutton & Barto, 2018).

Much of this chapter centres around the promises of combining rodent and human approaches in spatial navigation despite known physiological differences, for instance in how we process visual information (Zhao, 2018; Ekstrom, 2015). A significant challenge remains in relating the neural activity of single-cells to that of fMRI BOLD response commonly measured in humans (Ekstrom, 2020). The underlying neurophysiological processes that give rise to changes in blood oxygenation are still debated and not fully known. Prior work in monkeys comparing the similarities of BOLD responses in visual cortex to multi-unit activity, local field potentials, and spiking activity, suggest a complex relationship but one where BOLD activity is generally more similar to neuronal input to a regions, as measured via LFP rather than spike activity reflecting output (Logothetis et al., 2001; Logothetis & Wandell, 2004; Ekstrom, 2020). One solution to relating cross-species and cross-modality brain signals to one another could be to investigate the information content present. Techniques such as representation similarity analysis (RSA) provide an integrative approach to analyse neural information content (Kriegeskorte & Bandettini, 2008). Moreover, RSA of neural responses can be extended to compare these to computational predictions, for instance to representational similarities extracted from reinforcement learning models (Barron et al., 2021).

Future directions

In the search for generalizable mechanisms of cognition, there is an impressive range of computational models one could consider. These range from attractor networks, to neural networks, to sequential sampling, to drift-diffusion, to Hebbian models, to integrate-and-fire models, and many more (Dayan & Abbott, 2001; Gerstner et al., 2012). This chapter argues the case for the use of reinforcement learning when investigating underlying strategies of navigation. Additionally, it is suggested that the successor representation might be a candidate for a generalizable mechanism of cognition (Russek et al., 2017; Momennejad et al., 2017; Behrens et al., 2018). In order to test the generalizability of successor representation, or any other algorithms that could apply to several cognitive domains, there is a promising experimental framework that explicitly investigates hypotheses across different levels of analysis. We are at an interesting point in time where advances in human neuroimaging, rodent invasive recording technology, and computational techniques, can simultaneously be utilized to understand the information processing capabilities of the human brain.

There are many examples of interdisciplinary research between neural data and computational modelling, especially in decision making and navigation. However, these studies rarely explicitly and simultaneously spanned all three research modalities of rodent neural recording, human neuroimaging, and computational modelling. For instance, Barron and colleagues (2020) investigated inference using single-unit recording and optogenetics in mice as well as 7T fMRI neuroimaging in humans using comparable tasks. Another study, by Akrami and colleagues (2018), combined single-unit recordings and optogenetics in rats with computational modelling of behaviour in rats and humans to investigate the role of posterior parietal cortex on sensory history in choice behaviour. These studies demonstrate the strength of combining methodological approaches across species to gain deeper insight into cognitive phenomena.

Recent experiments have tested the model-free vs. model-based spatial navigation behaviour and the neural correlates thereof using virtual navigation and fMRI

(Anggraini et al., 2018; Simon & Daw, 2011). In Anggriani and colleagues' work (2018), trial-specific model-based behaviour was associated with parahippocampal and medial temporal lobe activity while model-free behaviour was associated with striatal and ventromedial prefrontal cortex activity. Importantly, a hybrid model incorporating both model versions best explained overall behaviour. The question remains if successor representation would have outperformed these models in accounting for behaviour and neural activity.

There have been studies in perceptual decision-making that use computational modelling and behavioural data from both rats and humans to study optimal evidence accumulation (Brunton et al., 2013) and working memory (Fassihi et al., 2014). However, there is still an underdeveloped frontier of combining these approaches with translational human neuroimaging techniques. Deep reinforcement learning algorithms have shown to be able to capture spiking neural activity during spatial navigation in the form of grid cells, with many similar properties in different layers of the convolutional neural network (Banino et al., 2018).

Information necessary for spatial navigation is encoded and processed in a wide variety of brain regions in humans and rodents (Epstein et al., 2017; Grieves and Jeffery, 2017). The specific (or algorithmic) ways in which this spatial information is processed, manipulated, and transmitted throughout cortical and subcortical brain regions and finally used in behaviour, remains undiscovered. By formulating our hypothesis about the information processing of spatial navigation in the form of testable computational models, we might be able to shed light on our lack of understanding about the different hierarchical spatial and temporal timescales the brain operates.

One danger of the division of human and non-human research is that the chasm separating the two will only keep widening, and develop separate vocabularies and practises, ultimately hindering progress for neuroscience (Badre et al., 2015). We are currently witnessing a massive growth of technology for invasive rodent recordings, similar to the explosion of human neuroimaging developments in the 1990s and early 2000s. It is important to combine these advances alongside advances in

human neuroscience in order to accelerate our understanding of the basic mechanisms underlying cognition.

2.4 Conclusion

In this chapter we consider and apply new methods for investigating neural underpinnings of flexible navigation. In the first part, we investigate and identify behavioural sampling biases in a rodent goal-directed navigation task and suggest methodological advances to remedy these through simulation and evaluation of randomly generated maze layouts. In the second part, we approach the problem of understanding navigation behaviour through the use of computational modelling. We implement three candidate reinforcement learning models, each representing different putative navigation strategies of place learning, response learning, and future planning. The models are compared against freely moving human behaviour and we show the difficulty in separating model predictions in the task with our limited data, although suggestive that model-based and successor representation are most similar to observed human trajectories consistent with recent evidence (de Cothi et al., 2020).

Chapter 3

Goal coding during flexible navigation

Mammals are capable of solving challenging navigation tasks using remarkable behavioural flexibility. The neural mechanisms that support this flexibility is thought to form a common system that underlie multiple cognitive domains including social networks (Tavares et al., 2015), hierarchical segmentation (Peer & Epstein, 2021; Balaguer et al., 2016), perceptual- and economic decision-making, and conceptual knowledge (Barron et al., 2020; Summerfield & Tsetsos, 2012; Shadlen & Shohamy, 2016; Mok & Love, 2019; Bellmund et al., 2018). Brain regions supporting this flexibility including the hippocampus have been studied extensively in humans and rodents. However, and perhaps surprisingly, we lack a unified understanding about how goals are represented in the brain, especially during complex navigation tasks where accurate goal representation is an essential component of successful navigation (Juechems & Summerfield, 2019). This poses two questions: 1) What is the nature of goal-related information?; and 2) What would a goal-like representation in the brain look like?

Compelling evidence of goal representation comes from work in bats which demonstrated goal-vector cells in the hippocampus that preferentially fire at a certain distance and egocentric angle to goal (Sarel et al., 2017). Other work in rats found that hippocampal place cells simulated future trajectories to goals (Pfeiffer & Foster, 2013). Similarly, fMRI studies have found hippocampal BOLD activity cor-

related with distance in tasks based on real-world navigation (Howard et al., 2014) and entorhinal activity correlated with angle to goal in a virtual reality task (Chadwick et al., 2015). Hippocampal BOLD activity was also found to index the path distance to both familiar and recently learned real-world environments (Patai et al., 2019). However, to date, the task environments used across species to investigate goal coding in spatial navigation vary considerably. It is the purpose of the current study to bring the fields closer together in a translational experiment.

Trajectories in most human spatial navigation experiments are highly stereotyped. Participants are constrained to binary or categorical alternatives to choose their path through the environment. In some cases, their choices have no impact on the subsequent navigation (Howard et al., 2014). We decided to model our task after a rodent experiment by Pfeiffer & Foster (2013) for the following reasons: 1) To allow for continuous exploration of the environment; 2) To observe a clear contrast in navigational goals and strategies across trials; and 3) Use non-overlapping starting locations, forcing the participant to take novel routes to the goal each time. We added one significant change to the task structure which was the introduction of barriers, creating a variety of complex mazes. The purpose of the barriers was to increase the variation in Euclidean distance to goal, increase the length of trajectories, and increase the sampling of egocentric goal directions necessary to successfully navigate. We designed the goal locations to change over the course of a session instead of once per day as was the case with the original task design. We wanted to maintain the hippocampal-dependent navigation components. Participants had to continuously re-learn not only the goal locations but also the placement of barriers as they would also change throughout the session. A closely-matched experiment was also run in rats with bilateral hippocampal implants but the findings from this experiment will not be reported in this thesis.

In the current study, we are investigating correlates of goal-related information in a spatial navigation task requiring constant updating of behavioural strategy and long trajectories to learned goal locations. We are also interested in brain regions involved in learning and updating learned goal locations. Based on recent litera-

ture, we hypothesise that the caudate nucleus is especially important for tracking distance-to-goal metrics (Javadi, Patai et al., 2019; Gahnstrom & Spiers, 2020), while the hippocampus is important for tracking goal-vector information including egocentric goal direction (Sarel et al., 2017). We also hypothesise that the hippocampus is involved in updating associations between learned rewards and their associated spatial locations.

3.1 Methods

Participants:

The study was approved by the ethics committee for UCL Division of Psychology and Language Sciences (fMRI/2021/001). For online testing, ethics approval was received from UCL Department of Experimental Psychology (EP/2018/008).

Twenty-five participants (age ranging from 19-36, mean age: 23.9, 15 female and 11 male) completed the full experiment including two training sessions and one fMRI session. Two participants were rejected due to excessive head motion and failure to reach a performance criterion of 90% (1 male, 1 female).

Maze Generation:

Each maze consisted of a 10x10 grid selected from a total set of 180,000 generated mazes. Mazes were generated in batches of 10,000 and used to select one maze per experimental block. Each subsequent batch of 10,000 mazes was non-overlapping with the previously selected maze to ensure that the same mazes were not being selected across blocks. In other words, any non-traversable tile in the current maze had to be traversable in the subsequent maze. We also wanted to make sure that each block pair was non-overlapping so that participants had to take novel routes and explore the whole environment.

Barrier creation:

Nine barriers were created on the grid-world of each maze by the removal of three adjacent tiles. The orientation of each barrier was determined by a random coin toss and the position was determined by 2-dimensional uniform random sampling. The

barriers were allowed to overlap meaning the removal of less than 27 (3 x 9) tiles was possible. The barriers could also create "islands" where direct paths to either goal location was impossible. The tiles of these "islands" were removed meaning more than 27 tiles could also be removed.

An edge detection algorithm identified any mazes with "edges". These were instances of the removal of two tiles along a diagonal. These were simply discarded and new mazes were created to replace them until the full set of 10,000 mazes contained 0 edges. The same mazes are used in a parallel series of rodent experiments and edges could be traversed by the rats if they make a diagonal movement resulting in an undesirable shortcut.

Maze Selection:

In short, optimal trajectories were computed using Dijkstra's algorithm from each possible starting location to the position of the one or two separate goals (Figure 1). The trajectories were up-sampled and smoothed using Savitzky–Golay filtering to get more realistic trajectories (Savitzky & Golay, 1964). The mazes were subsequently ranked according to the average of three metrics using all optimal trajectories: (1) The egocentric heading angles with respect to the goal locations; (2) The path distance divided by the Euclidean between starting position and goal location; and (3) The number of tiles removed from the environment. Each metric was given an equal weight in the ranking. We prioritised mazes that provided uniform sampling of angles, a large path distance over Euclidean ratio to give long trajectories, and a large number of tiles available for exploration. The final 16 mazes were plotted and manually inspected.

Starting location selection:

Each maze had either one or two Home goal locations. Twenty starting locations were randomly selected by dividen the 10x10 grid into four 5x5 quadrants. Each quadrant sampled 5 locations that were not already occupied by a removed tile or a Home goal location. All starting locations were manually inspected per maze and if the starting locations formed large clusters, the process was repeated until

the coverage was even across all quadrants. To reduce the total time taken per experimental block, we further selected only 10 out of the 20 starting locations by disregarding starting locations near the goal location. For the rat experiments all 20 starting locations were used as they require longer sessions with higher degree of spatial sampling of the environment.

Follow trial creation:

We added one additional element to the human version of the task. Every fifth trial for the first 15 trials in each block was designated a Follow trial (i.e. trial number 5, 10 and 15). Participants were instructed to follow a purple line on the floor for these trials, however, the general order of the task was kept the same. The different trial types of Home goal and Explore goal alternated from one trial to the next, starting with an explore trial.

The Follow trials were designed to match the distance of Home goal trials, with varying lengths and tortuosity, i.e. the amount of curvyness. The Follow trials kept the same order of the general task structure, meaning they used the goal location of the trial type they replaced. The first Follow trial occurred during an Explore trial (trial number 5), the second Follow trial occurred during a Home trial (trial number 10) and the final Follow trial occurred during an Explore trial (trial number 15). Participants were therefore aware of the place where the goal was located on trial number 10 even though it was a Follow trial, the goal was still located in the Home position.

Online task:

Participants completed two behavioural sessions at home during the two prior days to coming into the fMRI scan. First behavioural session also included filling out questionnaires including Santa-Barbara Sense of Direction Scale (n = 21, avg = 4.35, min = 2.27, max = 7, std = 1.03) and Navigational Strategies Questionnaire (n = 25, avg =-1.58, min = -10, max = 9, std = 4.86). The behavioural task was developed in the games engine Unity3D with custom scripts modified from Unity-Experimental-Framework (UXF: https://github.com/immersivecognition/unity-

experiment-framework). Additional software include Bitwig Studio for soundtrack creation and Blender for 3D modelling. We modelled the task after Pfeiffer & Foster (2013) with a hidden goal location similar to a land-based Morris water maze (Morris et al., 1983). One major difference is that we removed parts of the environment in order to perturb the possible routes one can take to any given goal location (Figure 1). The environment consisted of a 10x10 (100m by 100m in virtual space) grid of one hundred independently movable modules, each module being 10x10 meters in size. Removal of any given module constrains the navigable area and the removal of several modules creates a maze where changes in the maze layout forces participants to update their navigational strategies over time. Each side of the environment was fixed with a distinct distal landmark so that participants were able to correctly orient themselves and remember the location of the hidden goals. The view of the environment floor was obfuscated by fog centered on the participant. The fog was removed in a diameter of 15m around the participant to allow visual information about the maze layout in their immediate local area. We tried to emulate the poor visual acuity of rodents compared to that of humans by the addition of fog (Meier & Renagel, 2013; Keller et al., 2000).

There were three trial types aimed at investigating different navigational strategies and providing important controls. 1) Exploratory trials where the goal is hidden at a random location in the environment; 2) Homing trials where the goal is in the same location throughout one block of the task; and 3) Follow trials where you follow a purple line on the ground leading you to the goal. Another advantage of our trial structure is that it facilitates complete exploration of the environment and a priori determines the next starting location without the need to physically move or teleport the participant. The trial structure alternated between explore and homing trials. Trial numbers 5, 10, and 15 from the start of each block were Follow trials. Importantly, the navigation was always continuous throughout each block of the task. For example, an Explore trial goal location became the starting location for the Home trial. The current trial type was displayed on the bottom of the screen in the full name for session 1 (Same for Home, Random for Explore and Follow for

Follow) and in only the first letter in session 2 (indicated by R for Random (i.e. Explore), S for Same (i.e. Home) / and F for Follow). Participants could also see the current trial number in the bottom left of the screen during the first two behavioural training sessions.

Participants completed 140 trials for the first online session and 160 trials for the second online session with an increase in difficulty. A failed trial is triggered after 60s in the first session and after 45s in the second session. A failed trial causes the hidden goal in either Explore or Home trials to turn visible and change colour from purple to black. The first block in the first session was considered a training block. There were no removed modules and instead an open-field foraging environment where the hidden goal was visible instead for the first 14 trials. Moreover, for the first block of the first session, the distance necessary to catch the goal was 15 meters and after that it was 11 meters, including for all subsequent sessions.

Each block consisted of 20 trials. Each pair of blocks was considered to belong to either phase 1 or phase 2. Phase 1 meant that the pair of blocks will have the same maze but a different homing goal location. Phase 2 means that the pair of blocks will have different maze layouts but the same Home goal location. We had two phases in order to separate the influence of changing a maze layout vs. changing the goal location. However, after each pair of mazes, both the goal location and the maze layout change irrespective of the phase.

fMRI task:

The third session of the task was the fMRI session where participants went in the scanner (see Figure 2 for first-person view). The task was identical to the second session except that the trial number was no longer displayed in the corner. All three sessions had a different set of mazes for the participants (see maze generation for methods). The order of the phase 1 and 2 was alternated across each behavioural session. Each block of the task corresponded to one fMRI run for a total of 8 runs and blocks. The fMRI acquisition time for each session took an average of 55 minutes (or 405.7s per run). Participants were asked questions about the purpose of the task and reminded about the three different trial types prior to going in the scanner. They were also shown the button box and instructed on its use, including the option to move backwards.

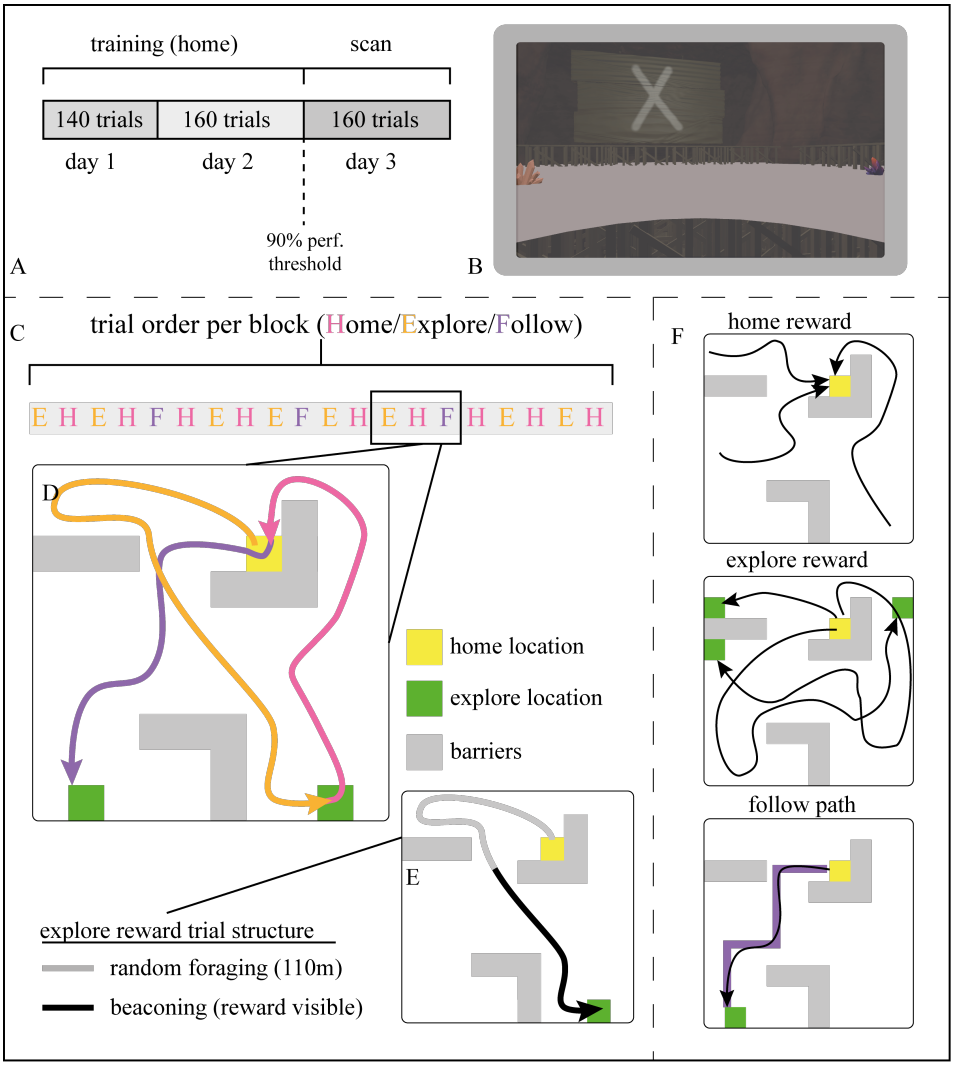

Figure 1. Schematic of the complex fMRI navigation task. (A) Participants completed extensive training on the task before coming in for the fMRI session. Day 1 consisted of an easy version of the task with fewer trials and more time allowed to complete each trial. Day 2 reduced the time allowed per trial from 60 to 45 seconds, and increased the trial number to 160 to match the fMRI session. Only difference between session 2 and 3 was the removal of a trial number indicator at the bottom of the screen. (B) A screenshot depicting first-person view of the navigation task. One out of the four distal cues (a large X) is visible as well as the two crystals present on each trial. The orange crystal was task-irrelevant and the purple crystal was the objective. An opaque fog obscured the ground but left a transparent disc around the participant (diameter 14m), forcing them to remember the layout of the maze in order to successfully and efficiantly navigate to the goal. (C) The task consisted of three distinct trial types, and the same order was preserved across all blocks. Each participant also experienced the same order of trials and maze layouts. (D) Schematic showing the trajectories from three trials in succession. The participant was never teleported or moved within the environment. In the example, a participant begins with an Explore trial with the starting location of the Home goal location in the previous trial. They navigate to the Explore reward and ends the trial. The Home trial begins after an ITI of 4-8 seconds, and concludes with finding the hidden Home goal. After another ITI, the Follow trial begins and the participant will see a purple line on the ground which they were instructed to follow. They follow the path until it ends at the follow goal location. (E) As indicated by a change in path color from grey to black, after 110 meters travelled, the explore goal becomes visible, changing the trial strategy from random exploration to beaconing. (F) The top two squares shows three example non-consecutive trajectories for different starting locations or goal locations for Home trials and Explore trials, respectively. The bottom final square shows an example trajectory for a Follow trial above the purple path they are instructed to follow.

Figure 3.1: Schematic of the complex fMRI navigation task

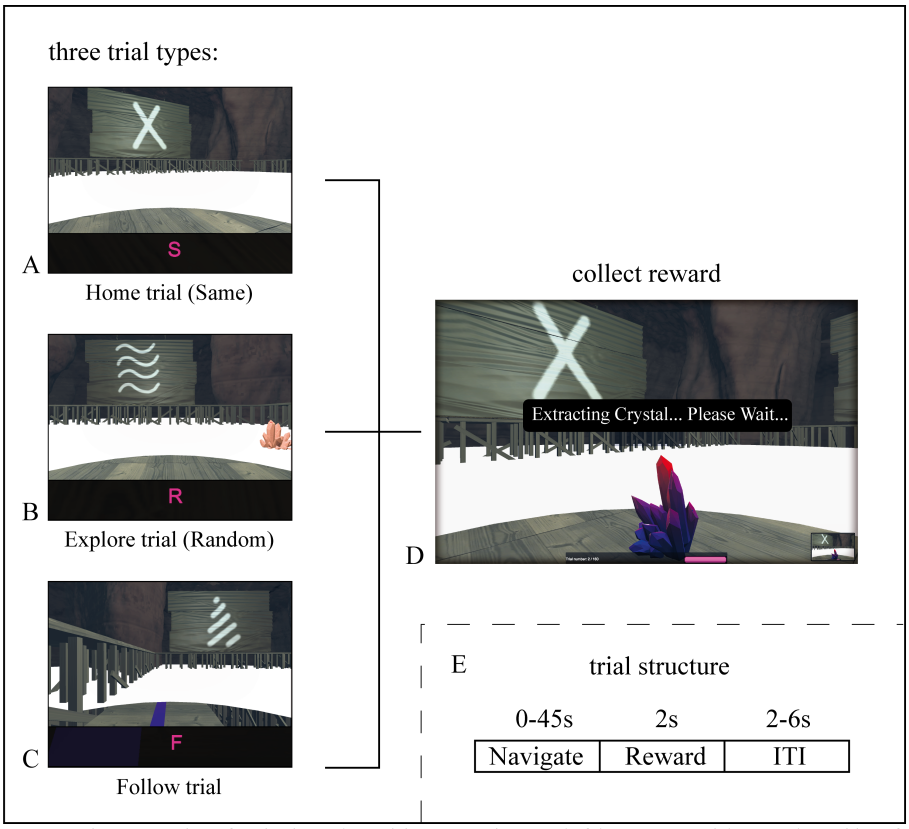

Figure 2. First-person view of navigation task. Participants were instructed of three separate trial types. The position of the hidden goal target, i.e. the purple crystal, changed as a function of the trial type. (A) The goal remained in the same location during Home trials (instructed to participants as Same trials, and always indicated with the letter S at the bottom of the screen) for the duration of each experimental block. (B) The goal was in a pseudo-random location during each Explore trial (instructed to participants as Random trials, and always indicated with the letter R at the bottom of the screen). After exploring the environment for 110 virutal meters, the purple crystal would turn visible and paricipants could utilise a beacon strategy to approach the goal. (C) The goal was either in a Home location or an Explore location, depending on which trial preceeded the Follow trial. The goal location always followed a strict alternation of Home location and Explore location regardless of the occurance of a Follow trial. During the Follow trial, a purple line appeared on the floor, and participants were instructed prior to the scanning session to follow the line on the floor and not veer of its path. The purple line would always lead the participant to the hidden goal location. (D) Goal arrival was initiated if within 11 virtual meters of the hidden goal location. The hidden goal, i.e. purple crystal,turned visible and reward collection began for 2s. After an additional 2-6 seconds uniform randomly jittered, the next trial would begin. (E) An overview schematic of a typical trial structure regardless of trial type. Participants navigate to the hidden goal for up to 45 seconds for Home and Explore trials, for Follow trials there was no time limit. If the participant ran out of time, the hidden goal turns visible but black instead of purple to indicate a failed trial. After arrival at the goal location, there was 2 seconds of reward collection, indicated by a loading bar in the bottom of the screen and a text pop-up stating “Extracting Crystal... Please Wait...”. The text pop-up remained on screen for an additional ITI of 2-6 seconds.

Figure 3.2: First person view of navigation task

MRI acquisition parameters:

Scanning was conducted at the Birkbeck-UCL Centre for Neuroimaging (BUCNI) using a 3-T Siemens Prisma MRI scanner with a 32-channel head coil. Each scanning session lasted around 1 hour and was separated into a maximum of eight runs (each approximately 5 - 7 minutes with 20 trials per run). Approximately 280 functional scans were acquired during each run, depending on participant performance (time taken per trial). A multi-band sequence was used (repetition time [TR] = 1450

msec, echo time [TE] = 35 msec, acquisition time (average) = 3276 seconds, flip angle = 70°). The scans were whole-brain (72 slices) with a multi-band acceleration of 4, slice thickness of 2 mm, spacing between slices of 2 mm, resolution / voxel size of 2x2x2 mm, field of view of 212 mm, and phase encoding A >> P. A top-up sequence was also performed at the end of the scanning session with reverse-phase encoding. A fieldmap was acquired after the first run. A standard T1-weighted high-resolution structural scan (MPRAGE) was acquired about half-way through each scanning session (TR = 2300 msec, TE = 30 msec, 1x1x1 mm resolution). Ear plugs and foam padding on each side of the head were used for noise reduction and to reduce head movements. Task stimuli were projected on a rear projector screen and mirrored for the participant's view. The mirror was adjusted for each participant to see full screen.

Experimental Design:

An event-related design with 160 trials, and 20 trials per block for a total of 8 blocks. There were three trial conditions, Explore, Home, and Follow. We had 8 Explore trials, 9 Home trials, and 3 Follow trials, per block for a total of 64 Explore trials, 72 Home trials, and 24 Follow trials per fMRI scan session. For each trial, the starting location was set to the previous trial's end location, with the only exception of the first trial of the session.

Each block had a specific maze layout which was the same for all participants, and all blocks were presented in the same order (Figure 3B). Each maze configuration was designed to allow for long trajectories with uniform sampling of egocentric goal directions (see maze generation methods). Within each block, starting locations were randomised (5 per quadrant of the environment) and presented to the participants in the same order. Each pair of mazes were non-overlapping in terms of tiles removed to allow for the entire environment to be explorable. The same mazes were used for a parallel rodent experiment where coverage of the environment is important for certain place cell analyses.

Behavioural analysis:

Our primary behavioural metrics were total path distance, euclidean distance to goal, egocentric angle to the goal, and number of crystals collected. We calculated these metrics either per frame or per trial. To test for session-effects and within-block effects, we ran repeated-measures ANOVAs on path distance as an index of performance. We also ran post-hoc t tests (Bonferroni-corrected) to make individual block and session comparisons.

For within-session path distance comparison, we normalised the path distance per trial by the optimal path distance on each respective trial using Dijkstra's algorithm. This allowed us to compare trials from different starting locations across and within task blocks.

MRI preprocessing and analysis:

Preprocessing was performed using fMRIprep (Esteban et al., 2018) with default settings (see Appendix for additional preprocessing details). First- and second-level GLM analysis was performed using SPM12. The first three dummy scans per run were automatically discarded before acquiring volumes.

For the full GLM we entered the following regressors of interest: Euclidean distance to goal, egocentric angle to goal, trial onset (split by trial type: Explore, Homing, and Follow), inter-trial-interval between adjacent trials, button presses (right/left/forward), and reward (goal) consumption (split by trial type: explore, homing, and follow). Path distance was not included as a regressor due to high collinearity with euclidean distance to goal once you are near the goal location.

The parametric regressors for euclidean distance to goal and egocentric angle to goal were entered separately into parametric GLMs. Further work beyond the scope of this thesis will also explore parametric modulation of optimal path distance. We only analysed the Home trials for our parametric GLMs as we were interested in neural responses related to learned goal locations in the environment.

3.2 Results

Behavior

The fMRI session consisted of 8 separate blocks where each new block contained either a change in maze layout or change in Home goal location or both. To test for the effect of block change, we ran a rm-ANOVA on the average path distance, normalised by the optimal distance, within each block separately ($F(1,22)=7.08$, $p < 0.001$***). We found a significant changes in path distance between several block pairs: block 1 (mean = 2.24m, std = 2.14) and block 3 (mean = 3.16m, std = 3.08) (t-stat = -4.57, $p < 0.05$); block 1 and block 7 (mean = 3.43m, std = 4.49) (t-stat = -5.03, $p < 0.001$***); block 2 (mean = 2.30m , std = 2.31) and block 8 (mean = 1.81, std = 1.22) (t-stat = 2.66, $p < 0.01$**); block 3 (mean = 2.79, std = 2.49) and block 5 (mean = 2.24, std = 2.02) (t-stat = 2.68, $p < 0.01$**); block 3 and block 8 (t-stat = 5.16, $p < 0.001$***); block 4 (mean = 2.52, std = 1.80) and block 8 (t-stat = 2.52, $p < 0.05$*); block 5 and block 7 (t-stat = -4.08, $p < 0.05$*); block 6 (mean = 2.81, std = 3.42) and block 8 (t-stat = 4.08, $p < 0.01$**); and block 7 and block 8 (t-stat = 4.87, $p < 0.001$***). All post-hoc results were Bonferroni-corrected, see Table 1 for exhaustive results.

To test one-shot learning, we ran a rm-ANOVA on the average normalised path distance per trial within a single block (Block 1: $F(1,22)=30.5$, $p < 0.001$***; Figure 3A). All even trials were Homing trials during the experiment, and the first Homing trial began with the second trial. We found the following significant changes in path distance between trial pairs during the first block of the fMRI session: trial 2 and all other trials (T4: t-stat = 4.94; T6: t-stat = 5.70; T8: t-stat = 5.56; T12: t-stat = 6.47; T14: t-stat = 5.61; T16: t-stat = 5.90; T18: t-stat = 5.23; T20: t-stat = 5.91; all $p < 0.001$***). (See appendix for full post-hoc results. See Table 2 for post-hoc results, block 1 only). The only significant difference between all trial pairs was found for the first Homing trial compared to the following eight Homing trials (Figure 3C).

To investigate any improvements over the sessions, we ran a rm-ANOVA on the average path distance per session for all the homing trials ($F(1,22)=22.1$, $p <$

Table 1. Post-hoc *t* Tests Path Distance Per Experimental Task Block (Normalised, Bonferroni-corrected)

	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8
Block 1	1.00	0.01*	1.00	1.00	1.00	0.02*	0.41
Block 2		0.31	1.00	1.00	1.00	0.22	0.02*
Block 3			0.17	0.01*	1.00	1.00	<0.001***
Block 4				1.00	1.00	0.14	0.04*
Block 5					1.00	0.02*	0.34
Block 6						1.00	<0.01**
Block 7							<0.001***

Shown are *p* values, with significance indicated by ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

Table 3.1: Post-hoc t Tests Path Distance per Experimental Task Block (Normalised, Bonferroni-corrected)

0.001***). We found significant change in path distance for all session pairs: session 1 (mean = 230.8, std = 64.4) and session 2 (mean = 174.0, std = 13.2) (t-stat = 4.80, p $<$ 0.001***); session 2 and session 3 (mean = 156.2, std = 4.12) (t-stat = 8.83, p $<$ 0.001***); and session 1 and session 3 (t-stat = 5.59, p $<$ 0.001***). See Table 3.

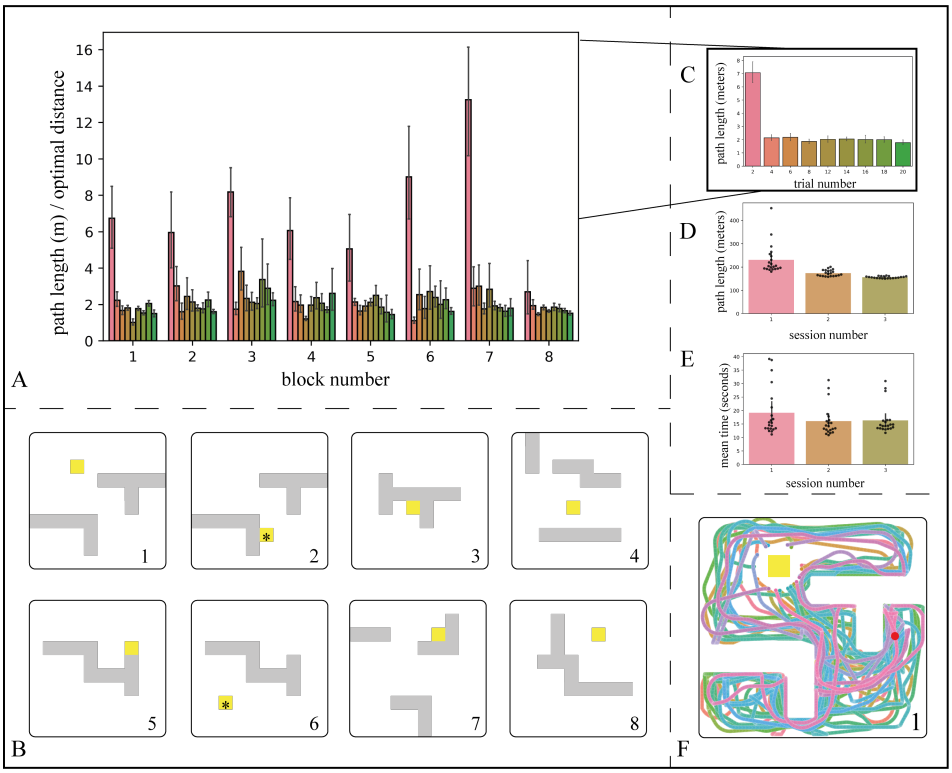

Figure 3. One-shot learning during flexible spatial navigation. (A) The path length for each Homing goal trial was divided by the optimal path from the starting location making different starting locations comparable.The path lengths were averaged across all participants (n=23) and plotted separately for all 8 blocks of the task. (B) Each square depicts the maze layout for the numbered block. The yellow square represents the goal location, and the grey squares represents barriers. * indicates mazes where the layout stayed the same and only the Home goal location changed. (C) Same data as in (A) but averaged over all the blocks. The normalised path length for first Homing goal trial was on average 3.5 times longer than following trials. (D) The path lengths for Home trials averaged across all blocks within a session to show the improvement across task sessions. (E) Mean time taken for Home trials in seconds averaged across all blocks within a session to show the improvement across task sessions. (F) Plotting trajectories for all participants (n=23) on the first Home trial of the first block during the fMRI session. Starting location is a red dot, and the Home Goal location is a yellow square.

Figure 3.3: One-shot learning during flexible spatial navigation

fMRI Results

Arriving at a hidden goal location drives hippocampal, caudate and prefrontal cortex activity

To investigate the processing of goal information we constructed a categorical contrast GLM between the onsets of goal arrival during Home trials vs. during Explore trials. The onsets began once the goal was reached and with duration of 2s for both trial types. After correction (whole-brain family-wise error rate $p < 0.05$, and for display purposes minimum voxel extent of fifty), we found significant bilateral

Table 2. Post-hoc *t* Tests Path Distance Per Trial in Task Block 1 (Normalised, Bonferroni-corrected)

	Trial 4	Trial 6	Trial 8	Trial 12	Trial 14	Trial 16	Trial 18	Trial 20
Trial 2	***	***	***	***	***	***	***	***
Trial 4		0.48	1.00	***	0.82	0.03*	1.00	0.04*
Trial 6			1.00	<0.01**	1.00	1.00	0.16	1.00
Trial 8				***	1.00	0.07	0.47	0.30
Trial 12					***	***	***	0.02*
Trial 14						0.17	0.10	0.66
Trial 16							***	1.00
Trial 18								<0.01**

Shown are *p* values, with significance indicated by ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

Table 3.2: Post-hoc t Tests Path Distance per Trial in Task Block 1 (Normalised, Bonferroni-corrected)

Table 3. Post-hoc *t* Tests Path Distance Per Session (Bonferroni-corrected)

	Session 1	Session 2	Session 3
Session 1	1.00	< 0.001***	< 0.001***

Shown are *p* values, with significance indicated by ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

Table 3.3: Post-hoc t Tests Path Distance per Session (Bonferroni-corrected)

activation in hippocampus (left: t-stat = 8.96, $p < 0.05$; right: t-stat = 8.32, $p < 0.001$***) during Home trial reward consumption. We also found significant activation in prefrontal brain areas including lateral OFC (left: t-stat = 10.94, $p < 0.05$; right: t-stat = 8.65, $p < 0.05$), bilateral ACC (left: t-test = 9.03, $p < 0.05$; right: t-test = 9.44, $p < 0.05$), right dmPFC (t-test = 7.93, $p < 0.05$) during Home trial goal arrival (Figure 4).

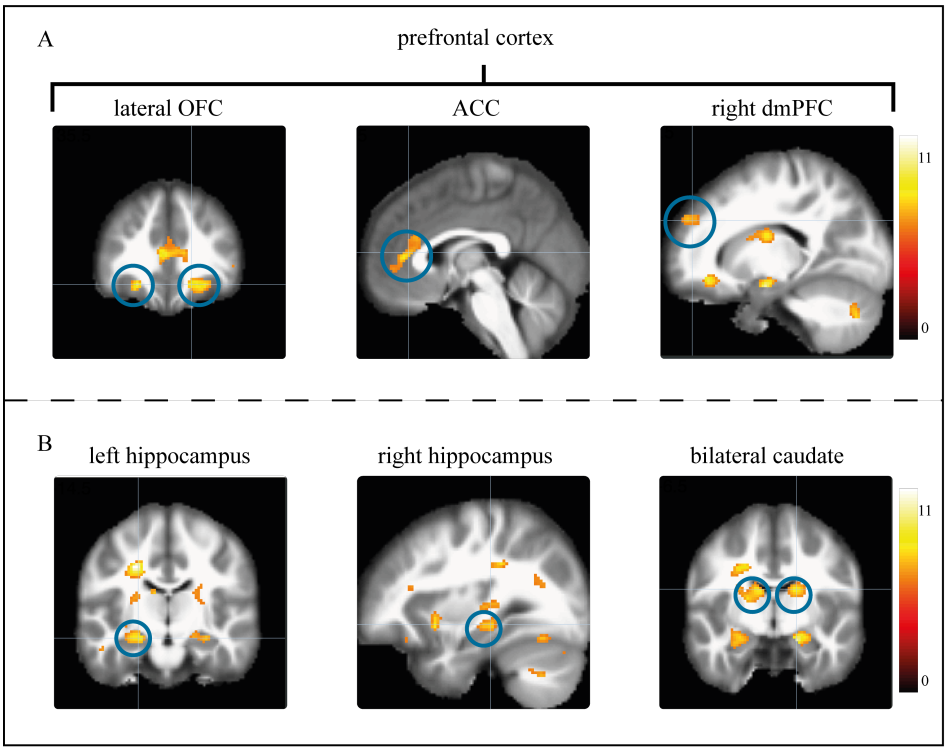

Figure 4. Network of brain areas involved in Home Goal arrival. (A) We identify a network of prefrontal areas that are significantly activated by the Home Goal onset during Home trials including bilateral orbitofrontal cortex (OFC), anterior cingulate cortex (ACC) and right dorsomedial prefrontal cortex (dmPFC) (whole-brain family-wise error corrected, $p < 0.05$). (B) We also find significant bilateral hippocampal activation and bilateral caudate activation (whole-brain family-wise error corrected, $p < 0.05$).

Figure 3.4: Network of brain areas involved in Home Goal consumption

Bilateral Precuneus, Parietal, and lateral Prefrontal Brain Areas Preferentially Active During Navigation vs. Follow Episodes

We contrasted episodes navigating to a learned hidden goal location (Home trials) with the Follow trials to investigate brain areas selectively activated by goal-directed navigation. To do so, we constructed a GLM contrasting the onset and duration of Home trials with that of Follow trials. After correction (family-wise error rate, minimum voxel extent of five), we found significant bilateral activation in posterior intraparietal sulcus (left: t-stat = 8.51, $p < 0.05$; right: t-stat = 9.09, $p < 0.05$), bilateral activation in lateral PFC (left IFG: t-stat = 9.28, $p < 0.05$; right IFG: 8.18, $p < 0.05$), bilateral precuneus (left: t-test = 11.44, $p < 0.05$; right: t-test = 11.73, $p < 0.05$), and bilateral activation in frontal eye fields (left: t-test = 7.31, $p < 0.05$*; right: t-test = 7.43, $p < 0.05$).

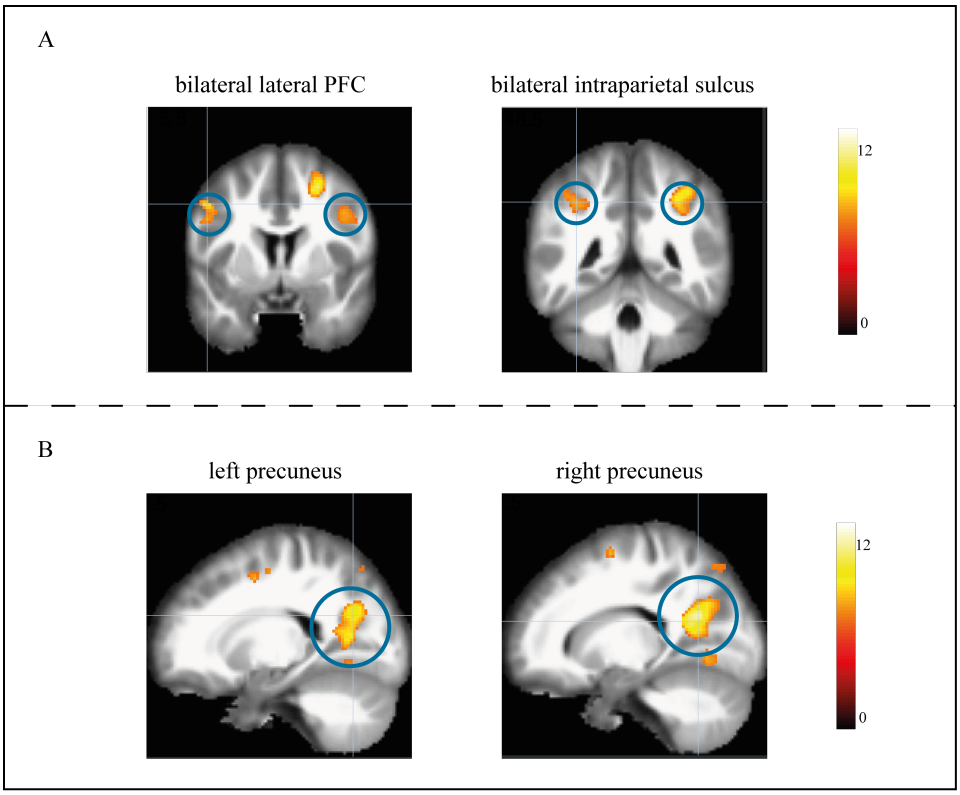

Figure 5. Brain regions activated in contrast Homing vs. Follow trials. (A) We find significant activation in bilateral inferior frontal gyrus, i.e. lateral PFC, as well as bilateral parietal cortex. (whole-brain family-wise error corrected, $p < 0.05$). The two main brain regions involved in state-prediction error detection. (B) We also identify significant activation in bilateral precuneus (whole-brain family-wise error corrected, $p < 0.05$).

Figure 3.5: Brain regions activated in contrast Homing vs. Follow trials

Caudate Tracks of Euclidean Distance To Home Goal

The tracking of goal-related information was tested. We constructed a GLM with parametric regressors for distance to goal calculated at each frame during all Home trials. After correction (family-wise error rate, minimum voxel extent of five), we found significant bilateral caudate (right: t-test = 8.45, $p < 0.05$; left: 8.12, $p < 0.05$). We also found significant activation in inferior parietal lobule (left: t-test: 8.22, $p < 0.05$; right: 10.61, $p < 0.05$), bilateral lateral PFC (IFG), (left: t-test 8.77, $p < 0.05$; right: t-test = 8.29, $p < 0.05$).

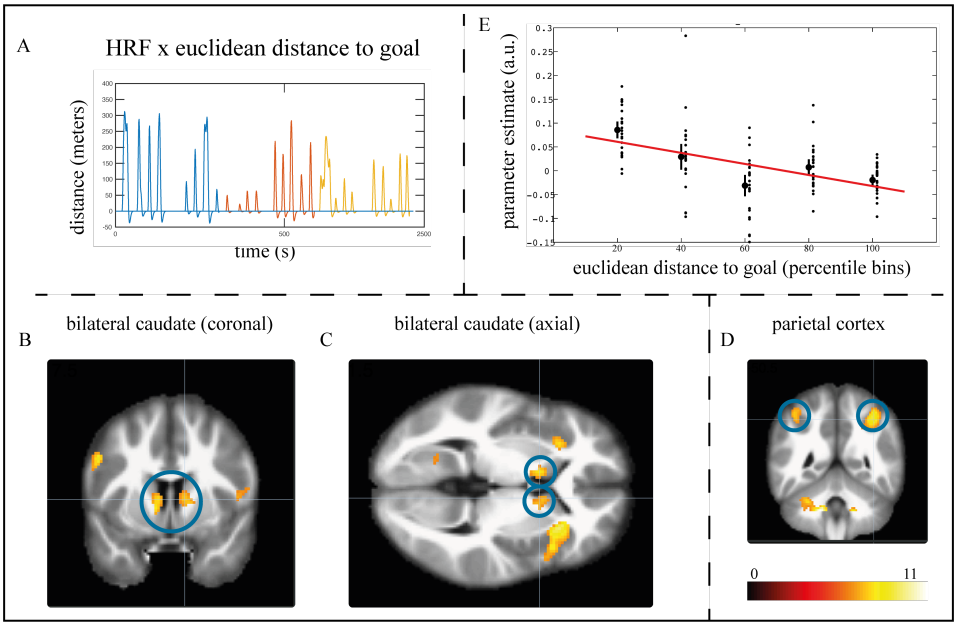

Figure 6. Caudate nucleus, insula, and parietal cortex positively correlated with euclidance distance to goal. (A) The heamodynamic response function (hrf) convolved with the time-course of euclidean distance to goal during Home trials for first three runs of one example participant. (B) We find significant parametric modulation w.r.t changed in euclidean distance to goal in bilateral caudate nucleus, (C) same bilateral caudate response plotted in axial view, and (D) bilateral parietal cortex (whole-brain family-wise error corrected, $p < 0.05$). (E) Parameter estimates for suprathreshold voxels in right caudate nucleus spherical mask centered on peak voxel at (mni-coordinate 16, 2, 8). Errors bars are 95% CI, and each dot represent parameter estimate for a given percential bin for each participant. Red line is best fitting linear regression through all data points.

Figure 3.6: Caudate nucleus, insula, and parietal cortex respond to changes in Euclidean distance to goal

MTL, Hippocampal, and Prefrontal cortex tracking of egocentric goal direction

Previous research has shown a subpopulation of hippocampal cells which fire according to the relative heading direction of bats to hidden goals in their environment (Sarel et al., 2017). We constructed a GLM with parametric regressors for egocentric goal direction to investigate this finding in our data (Figure 7). We restricted the absolute angular values from 0 to 180, instead of ranging from -180 to 180, because we wanted to investigate the modulation in relation to a one-dimensional deviation in direction towards the learned Home goal. We found significant activation in the left hippocampus (t-test = 5.557, $p < 0.0001$, uncorrected) and the right parahippocampus (t-test = 5.42, $p < 0.0001$, uncorrected). We also find significant bilateral lateral temporal lobe (whole-brain family-wise error corrected) (left: t-test = 9.03, $p < 0.05$; right: t-test = 10.52, $p < 0.05$). We also found significant dorsomedial

PFC activation (whole-brain family-wise error corrected, right: t-test = 11.05, $p < 0.05$; left: t-test = 7.34, $p < 0.05$), as well as medial OFC (t-test = 7.07, $p < 0.05$).

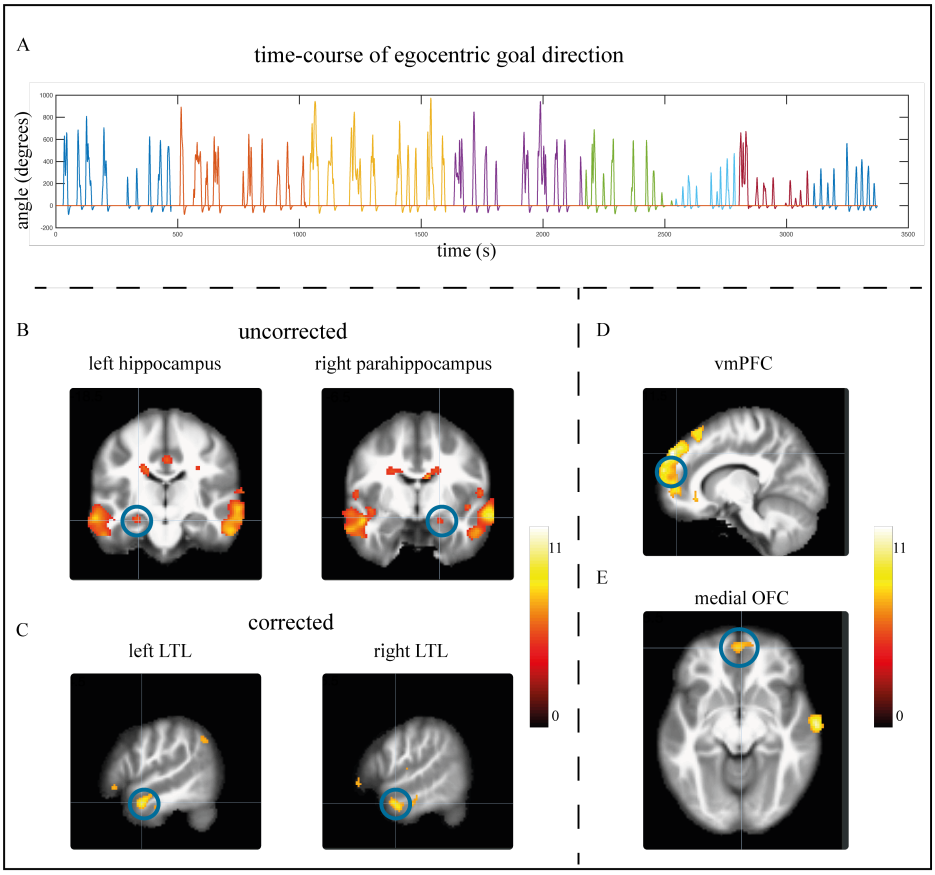

Figure 7. Hippocampal and prefrontal negative correlation with egocentric goal direction. (A) The hemodynamic response function (hrf) convolved with egocentric goal direction during Homing trials for one example participant. Each color represents a different run for a total of 8. Gaps in the signal occurs during Explore and Follow trials. (B) Parametric regression of egocentric goal direction with hippocampal / parahippocampal responses ($p < 0.0001$, whole-brain uncorrected for left hippocampus, and right parahippocampus). (C) Bilateral lateral temoral lobes activated by egocentric goal direction (whole-brain family-wise error corrected, $p < 0.05$) . (D) Right ventromedial prefrontal cortex and (E) medial orbitofrontral parametrically modulated by the egocentric goal angle (whole-brain family-wise error corrected, $p < 0.05$).

Figure 3.7: Hippocampal and prefrontal response to changes in egocentric goal direction

3.3 Discussion

Humans and rodents demonstrate remarkable behavioural flexibility during spatial navigation tasks (de Cothi et al., 2020). We investigated goal representation during a flexible navigation task in humans using fMRI. Our task demonstrated the ability to update learned goal information as well as update behavioural strategies over time. Participants exhibited clear one-shot learning where one exposure to a new hidden Home goal was enough to learn the location for subsequent Home trials. We hypothesised hippocampal involvement with the updating of newly learned goal locations. We found supporting evidence for this hypothesis along with a network of other brain regions including, OFC, ACC, right dmPFC, and caudate nucleus. We

also found evidence for dynamic tracking of euclidean distance to goal in bilateral caudate nucleus, a finding consistent with recent work (Javadi et al., 2019) as the region is suggested to play a role in representing the transition structure of the environment (Gahnstrom & Spiers, 2020; Sharpe et al., 2019). Lastly, we show that facing a hidden goal modulates activity in hippocampus/parahippocampus, lateral temporal lobes, medial OFC, and dmPFC with implications for goal representation during ongoing flexible navigation.

Rapid learning of goals in spatial navigation

We show rapid learning and re-learning of goal locations throughout the experimental session, where participants are able to approximate the shortest path to goal after one exposure to a new goal location (Fig. 3A). This rapid learning is faster than expected based on canonical one-shot learning of a novel spatial environment which normally takes several trials to asymptote, for instance with rodents in the Morris water maze and for humans in a taxi driver task (Morris et al., 1984; Newman et al., 2006). Our participants already learned the relative position of the distal landmarks during two previous behavioural sessions leading up to the scanner sessions which may have resulted in schematization of the spatial environment leading to an ability to incorporate new goal information on a faster timescale (Squire et al., 2015). In future analyses, we can investigate the within-session improvement and observe the difference in learning time-scale as a function of experience in the maze environment.

We calculated the shortest path based on approximations of the real shortest path by reducing the original in-game coordinate space to a 10-by-10 grid. These sometimes lead to incorrect estimations where some path lengths end up shorter than the approximated shortest path. Moreover, the grid calculations of shortest path are based on cardinal directions (up, down, left, righ). Future analysis will include diagonal movement and a larger grid 100-by-100 which will result in more accurate shortest path calculations and eliminate the impossible cases where path distance is shorter than the optimal path.

Path following in navigation

The idea of separate subsystems of spatial memory is supported by findings in human fMRI using virtual navigation paradigms. Several studies have identified hippocampal activation during wayfinding or place strategy and striatal activation during route following or response strategy (Hartley et al., 2003; Iaria et al,. 2003). Hartley et al. (2003) tested participants in three separate conditions, wayfinding, route-following, and trail-following. Although they did not report brain regions preferentially activated during trail-following, they do report the contrast wayfinding condition against the trail-following. Similar to our reported findings, they identify increased BOLD activation in insula, precuneus, calcarine gyrus, right fusiform, and dorsolateral prefrontal cortex. We also find medial prefrontal cortex activation although more posterior than reported by Hartley et al., (2003). These precise coordinates of these effects were not reported and so direct comparisons are limited.

Theories of OFC subregions in value-based decision making: Value learning in lateral OFC and ongoing decision making in medial OFC

Our task identified several key regions involved during flexible spatial navigation including lateral orbitofrontal cortex (lOFC) and medial orbitofrontal cortex (mOFC). The OFC has been studied extensively in reversal learning, devaluation, value-based decision-making, and inference tasks (Gardner & Schoenbaum, 2021; Banerjee et al., 2020 Nature; Rich & Wallis, 2016). Although orbitofrontal responses have been lacking in spatial navigation studies (but see Wikenheiser et al., 2021; Dahmani & Bohbot, 2015; Busu et al., 2021) There are multiple competing account as to its functional role ranging from a role in forming a cognitive map of task space (Wilson et al., 2014) to credit assignment (Noonan et al., 2010), to economic choice (Padoa-Schioppa & Cohen, 2017).

Our task identified two anatomically separate BOLD activations in OFC that are broadly consistent with theories of these regions in value-based decision making tasks. We identified lateral OFC activation when arriving at a hidden goal location

and collecting the reward. Critically, the hidden goal required constant updating and re-learning for each block during the course of the experimental session. The lOFC is suggested to be especially important during value-based learning (Walton et al., 2010; Boorman et al., 2016). In one study, lOFC was shown to be important in switching behaviour based on hidden task state (Vertechi et al., 2020). Mice and humans were tasked with inferring hidden states during a foraging task, and optogenetically silencing lateral OFC reduced the mice ability to perform inference-based decisions and reverted to model-free, or stimulus-based, responses. In our task, participants need to continuously switch between task states. Although these are not hidden, participants need to switch between different strategies depending on current task state, i.e. trial type, which may lead to the observed lateral OFC BOLD activation. Should be noted that it's been proposed that OFC is not involved in driving task context but rather represents current location within and across the trials (Sharpe et al., 2019).

The medial OFC response was negatively correlated with egocentric direction to a hidden goal during the Home goal trials, i.e. more activation while facing towards the hidden goal location. Interestingly, one study contrasted the influence of lesions in lateral vs. medial OFC in macaques during value-guided decision-making (Noonan et al., 2010). Lateral OFC lesions resulted in option values being weighted to the recent average across multiple options, an indication of inability of assigning the correct outcome to the correct option. In other words, an impairment of credit assignment, which is a central problem in classic reinforcement learning (Sutton & Barto, 2018). The unique contribution of lateral OFC to credit assignment has also been observed in macaques who displayed increased choice switching following lesions (Walton et al., 2010). However, they also observed that lesion to medial OFC resulting in impairments of the ongoing decision-making, as opposed to learning value associations. A more recent suggestion is that medial OFC represents terminal states in the task environment while lateral OFC represents initial states (Bradfield & Hart, 2020). Lastly, lateral OFC has also been implicated during ongoing learning of stimulus-outcome associations (Walton et al., 2010; Boorman et al., 2016) and

decision confidence (Lak et al., 2014). Consistent with medial OFC involvement in ongoing value-based decision making, we also find medial OFC involvement in ongoing flexible navigation in tracking the egocentric direction to hidden goal location. Lastly, several recent studies hypothesise that medial OFC tracks your current position within the task space (Wilson et al., 2014; Sharpe et al., 2019; Schuck et al., 2016).

Orbitofrontal cortex during spatial navigation

The unique contribution of OFC during spatial navigation tasks is still unclear but combining ideas from value-based decision making might provide insights (Patai & Spiers, 2021). One navigation study investigated the future goal representation during a prospective planning task in a virtual environment. They found prospective representation of navigational goals in the hippocampus and the strength of this relationship was also reflected in medial OFC (Brown et al., 2016). Another study investigating brain network correlations during spatial navigation suggests that OFC is a brain region which facilitates goal-directed navigation through bilateral projection to both caudate and hippocampus (Brown et al., 2012). Neural responses in the hippocampus along with medial OFC has also been activated by early learning and spatial strategy use (as compared to response strategy) during a virtual radial maze task (Dahmani & Bohbot, 2015). They also found supporting evidence that grey matter volume in OFC is correlated with spatial strategy use and propose as others have that OFC is involved in stimulus-reward associations contrasted with hippocampal involvement in stimulus-stimulus associations (Dahmanu & Bohbot, 2015; Bohbot et al., 2007; Sharpe et al., 2019). Our results support the idea that both OFC and hippocampus are involved in flexible navigation (Brown et al., 2016; Brown et al., 2011; Dahamnu & Bohbot, 2015).

Interestingly, the lack of OFC activation in some studies may be due to methodological issues (Spiers & Patai, 2021). Some prior fMRI studies may have been more susceptible to signal drop-out in prefrontal regions, especially the OFC (Weiskopf et al., 2007). Our current methods use susceptibility correction procedures to correct for the signal distortion in these regions. However, this will not

recover any lost signal but correct for the anatomical distortions that arise. We also use multi-band sequencing which allows us to comfortably collect whole-brain images, where previous single-band sequences may have had to leave out certain brain regions in favour of other ones due to a more narrow field of view.

One recent study explicitly investigated the ability of neurons in the OFC to represent future goal locations (Busu et al., 2021). In their linear maze task, rodents alternated between two different goal locations within each block, and these two goals also changed locations across blocks. They found that the OFC neurons represented the future goal location prior to the start of navigation until reaching the goal location, where they would start to represent the following goal location. Using optogenetics, they also found that silencing OFC neurons prior to the start of navigating resulted in navigation errors. In our data, we found that medial OFC represented hidden goal direction which is consistent with the proposed importance OFC in representing future goal (Busu et al., 2021).

Caudate involvement in distance to goal computation

In spatial navigation tasks, the caudate has long been attributed with stimulus-response or inflexible behaviour (Dahmani & Bohbot, 2015; Hartley et al., 2003; Iaria et al., 2003; Doeller et al., 2008). Recently, caudate is considered to have a role to play to facilitate behavioural flexibility in navigation tasks, mirroring findings in instrumental learning (Sharpe et al., 2019). In the current task, we find that bilateral caudate is dynamically tracking the euclidean distance to a hidden goal in the environment. Interestingly, recent studies have also identified caudate tracking distance to goal, specifically during detour events. One study used real-world learning of street networks later probed using video of the same area of central London (Howard et al., 2014). Another study used a virtual environment (lava world) specifically designed to parametrically change the magnitude of degree in detour (Javadi, Patai et al., 2019). A difference in the current study is that the maze layout of the environment was obscured by fog, a clear difference from the lava world where the maze layout is visible. Participants therefore needed to explore the environment in each block to learn the maze layout and could not simply visually inspect the layout

and plan accordingly.

Taken together, the caudate (i.e. dorsomedial striatum in rodents) involvement during flexible spatial navigation has been proposed to encode the state transitions of the environment (Sharpe et al., 2019; Gahnstrom & Spiers). Our findings could be consistent with this view as bilateral caudate tracked the euclidean distance to goal. Perhaps one interpretation could be that our participants activate their representation of the transition structure of the environment while far away from the hidden goal. If nearby the hidden goal location you no longer need to remember the transition structure (or maze layout) since it is unlikely you have a barrier in the way.

Hippocampal contributions to flexible navigation

We find uncorrected hippocampal involvement during egocentric goal direction rotations where greater involvement occurs while looking directly towards the learned hidden goal during Home trials. Future analysis will investigate if the hippocampal activity survives region of interest small volume correction. However, this finding could support the idea of goal-vector cells in hippocampus that preferentially fire at an angle towards the goal, as has been found in bats (Sarel et al., 2017; although recently suggested to be an artifact based on simulation analysis approach by Carpenter and colleagues, 2021). Our task requires long trajectories around the learned goal location, perhaps more similar to the kind of long trajectories of bats as they fly circling towards the goal location. This could explain how we are able to detect such a signal while other studies have not. Previous reports have found positive correlates as distance to goal increase the posterior hippocampus using real-world stimuli (Howard et al., 2014; Patai, Javadi et al., 2019) while negative correlates as distance to goal increase in virtual environments (Sherrill et al., 2013; Viard et al., 2011). One recent rodent study found that place cells increased the in-field firing rates as an animal approached a goal location which may result in a goal-direction sensitive signal (Aoki et al., 2019). Gauthier & Tank (2018), also identified a separate population of reward-predictive cells in the hippocampus that would fire immediately preceding the goal location. Further analysis of our data may be able to

identify if the goal-direction signal is more strongly activated near the goal as opposed to further away, which would be more consistent with these two recent rodent studies (Aoki et al., 2019; Gauthier & Tank, 2018). Analysing single-cell recordings in human epileptic patients, Kuhn et al. (2021) also identified neurons coding egocentric bearing to locations during a goal-directed navigation task. Although, these bearing cells did not code for egocentric goal-related information.

Interestingly, we also find bilateral hippocampal activation after reaching the hidden Home goal. It is possible that the hippocampal response is related to forming a new spatial association of the location and the hidden goal as these need to be re-learned periodically at the start of each block. Research in rodents has shown evidence of multiple possible mechanisms of spatial learning and planning processes in the hippocampal formation (Foster, 2017), and recent theoretical work has explained these using a normative model based on a reinforcement learning architecture (Mattar & Daw, 2018). For example, upon reaching a goal location, hippocampal place cells have been shown to replay in the reverse sequence (Ambrose et al., 2016; Shin et al., 2019; Bhattarai et al., 2020), which in itself suggests a process of assigning proper credit to actions and states, i.e. temporal credit assignment (Sutton & Barto, 2018). The relationship between goal-directed behaviour and sharp-wave ripples and replay is still under active investigation (Dupret et al., 2010; Pfeiffer & Foster, 2013; Shin et al., 2019; Xu et al., 2019). Forward planning is another mechanism proposed by findings in rodent hippocampal replay (Pfeiffer & Foster, 2013). For instance, one study by Igata and colleagues (2021) found replay of an optimal path to subgoal in an open field environment prior to choosing that path during subsequent trials. Other work has found that replay is more akin to a brownian motion process and reactivating random trajectories instead of the observed behaviour during a random foraging task suggesting replay mechanisms are tied to task demands (Stella et al., 2019). It is possible that the hippocampal involvement we identified is related to a similar replay event occurring after reaching a goal. Recent work in humans has suggested that similar replay mechanisms are at play in the human hippocampus in non-spatial tasks (MEG: Kurth-Nelson et al., 2016; fMRI: Schuck

& Niv, 2019; Vaz et al., 2020), although this has not been studied in spatial tasks. Future work will need to further understand how the different replay mechanisms suggested in rodents may translate to flexible navigation in humans.

Limitations in analysis

Our approach described in this chapter to investigate goal coding in the human hippocampus has important limitations in the underlying analysis. We have used voxelwise inference where statistical significance is evaluated at the level of each voxel. One problem for parametric voxelwise inference is conservative estimates while nonparametric methods can provide nominal results for both voxelwise and clusterwise inference (Woo et al., 2014; Eklund et al,. 2016). We corrected for multiple comparisons by setting a family-wise error rate to 0.05 and for display purposes a voxel extent of 5 voxels. However, choosing a different voxel extent number may mislead the visualisation as the statistical mask appears less spurious. A future approach that may alleviate this issue is the use of nonparametric approaches such as cluster-based bootstrapping where randomly permuting the data creates the null distribution (Eklund et al., 2016).

3.4 Conclusions

Our study shows an engagement of prefrontal regions, with hippocampus and caudate nucleus to support flexible spatial navigation. We find a subdivision of the orbitofrontal cortex between lateral and medial between learning and on-going goal-directed behaviour similar to evidence in macaques (Walton et al., 2010; Gardner & Schoenbaum, 2021). We find evidence of hidden goal processing in the hippocampus supported by previous work in humans and in bats (Sarel et al., 2017; Javadi et al., 2019; Howard et al., 2014). The dichotomy of hippocampal versus caudate involvement in flexible versus inflexible behaviour, respectively, may need a revision as the results of this chapter (in line with other evidence) suggest the caudate to be involved in a task requiring flexible navigation within an ever-changing environment. This task would stereotypically require hippocampal involvement, yet we report evidence of caudate modulation with Euclidean distance to goal in the

absence of hippocampal modulation. As such, this chapter provides some evidence for a rethinking of this dichotomy as well as the cognitive map theory as strictly hippocampal-dependent.

Chapter 4

City Hero Quest

So far in this thesis, we have investigated human behavioural flexibility in spatial navigation tasks following a long tradition starting with Tolman's rodent experiments in the 1940's (Tolman, 1948). He posited that the remarkable ability of rodents to form novel shortcuts implied the use of an internal map, or cognitive map as he called it. Recent work has questioned Tolman's findings, for instance, by suggesting that the shortcut behaviour is due to a light positioned near the goal location facilitating the use of a homing strategy (Grieves et al., 2013). More controlled maze setups have failed to show novel shortcut behaviour in both rats and humans (Grieves et al., 2013; Wilson & Wilson, 2018). Nevertheless, mammals in the wild can display remarkable behavioural flexibility as evidenced by shortcut behaviour in real-world bat flight after translocation to rarely visited locations (Toledo et al,. 2020).

The cognitive map hypothesis posit that rodents and humans could execute behaviours that went beyond simple stimulus-response association. At the time, stimulus-response associations were touted as the only mechanism necessary to explain all of human behaviour (Newcombe, 2018). Many groundbreaking discoveries have identified the hippocampus as a central brain region for memory (Scoville & Milner, 1957) and spatial navigation (O'Keefe & Dostrovsky, 1971; Wilson, 1978). The hippocampus is also suggested to be an instantiation of Tolman's cognitive map (O'Keefe & Nadel, 1978). Despite decades of work, it still remains unclear exactly how the brain instantiates a cognitive map based on Euclidean distances

and to what extent the hippocampus or a network of brain regions implements it, with some suggestions that a cognitive graph better explains experimental findings (Peer et al., 2021; Warren et al., 2017; Warren, 2019).

Perhaps the difficulty in reconciling the idea of a cognitive map in a clear neurobiological framework comes from the individual variability inherent in spatial navigation tasks (Weisberg et al., 2014; Weisberg & Ekstrom, 2021). It has been suggested that perhaps some participants do not form cognitive maps (Weisberg & Newcombe, 2016) or that Euclidean cognitive maps do not exist at all (Warren, 2019). Many research paradigms build on the idea of a classification of navigational strategies that can be either place-based (i.e. allocentric) or response-based (i.e. egocentric) to explain individual variability (Hartley et al., 2003; Igloi et al., 2010; Iaria et al., 2003; Anggriani et al., 2018). One further idea in this line of research is that the brain supports multiple memory systems but that different people have learning biases resulting in a preference for one navigational strategy over another (Marchette et al., 2011).

The origin of the bias in navigational strategy preference and, moreover, navigation ability is still unknown. Recent work has begun to investigate the role of home environment on navigational ability (Coutrot et al., 2020). One study directly compared navigational ability and strategy use between a city in the United States (Salt Lake City, Utah) with a city in Italy (Padua, Veneto) with age- and education-level matched participants. They found that participants in the United States, living in a city with a stereotyped, grid-like street pattern, self-reported a higher use of cardinal direction and a higher self-score of sense of direction. Nevertheless, these participants performed worse in a virtual environment task relative to the participants in Italy, whose city is highly irregular in comparison (Barhorst-Gates et al., 2021). The participants in Padua, Italy, were also more likely to use shortcuts compared to participants in Salt Lake City, Utah – an indication of a higher propensity towards cognitive map formation.

In this chapter, we will extend previous work described in Coutrot et al. (2020) by large-scale online behavioural testing using both mobile and desktop virtual nav-

igation tasks designed to investigate the role of home environment and self-reported navigational strategy use on navigational ability. The original study found evidence from 38 countries that the street network entropy, i.e. the regularity of street layout, was correlated with navigational ability. Specifically, growing up in a less grid-like environment was beneficial to navigational performance (measured as reduced distance travelled in the virtual task Sea Hero Quest). We used six wayfinding levels from Sea Hero Quest with the highest effect size on the relationship between navigational ability and street network entropy to design a new study that directly investigated this relationship. We created a new virtual environment task using a city-themed landscape and directly replicated the Sea Hero Quest layouts and task structure of the selected wayfinding levels. In this chapter, we will present the findings and discuss the generalisability of wayfinding ability across task contexts as well as the consequences of growing up in grid-like environments on wayfinding ability.

4.1 Methods

Participants

Participants were recruited using the online platform Prolific (Palan & Schitter, 2018; www.prolific.co). We divided the data collection into separate batches to be able to manage technical difficulties that could arise. We collected 8 batches total and each batch was prescreened using the same set of restrictions: United States as 'current country of residence', United States as 'Nationality', ethnicity provided by the user (not screened for any specified ethnicities), Prolific approval rate of greater than or equal to 90%, and Mac OS / Ubuntu / or Windows as 'Computer Operating System'. Finally, participants from previous batches were excluded from all subsequent batches to ensure we did not sample the same participants multiple times. Prolific approval was manually inspected by experimenter after confirming the upload of both sets of data – Sea Hero Quest and City Hero Quest – to an anonymous OneDrive server. Partial datasets were approved but excluded from analysis. All approved participants were paid at a standard rate of £7.50 per hour. Average time

to complete the study was around 2 hours.

Sea Hero Quest

Sea Hero Quest is a mobile virtual reality game designed by researchers at University College London and University of East Anglia in collaboration with Deutsche Telekom and Alzheimer's Research UK. The game was built in Unity 3D by Glitchers Games Company and initially released on both Apple and Android app stores (Coutrot et al., 2018). The initial version of the game has since been removed from public download but an unlocked version is available for research purposes. This unlocked version was used for this experiment and provided to participants via a URL download link. We selected a set of levels targeted at testing wayfinding ability. Within each level, players move around within a water environment (e.g. river, lake, sea), sailing a virtual boat. Movement is restricted to left, right, and forward movement with a baseline velocity from the start of each level.

For each Sea Hero Quest level, participants are first shown a map of the maze layout with a series of goals indicated by numbers on the map (Figure 1). This map is viewed for any duration of time chosen by the player. The level starts when 'Start' at the bottom of the screen is pressed by the player. Once the map is removed and the player has initiated the start of the level, the goals must be located in the corresponding numbered order for the level to be successfully completed. Once a goal is travelled through within the collider radius, the collider mesh around the goal turns green, indicating the sub-goal or goal has been completed. Player position and rotation are logged at regular intervals.

The full Sea Hero Quest game has 75 levels. We carefully selected a subset of 5 levels to reduce the time taken to complete the game while still maximising our ability to detect effects on wayfinding ability. The first level was level one in the full game and designed to control for video game experience (Coutrot et al., 2018). The second level was chosen as an intermediary difficulty level and was level 11 in the full game. The third to fifth level was chosen as levels with the largest effect size of the correlation between the street network entropy of the environment where you grew up and performance measured as distance travelled in that level (Coutrout

et al., 2020).

City Hero Quest

City Hero Quest is a desktop virtual reality game designed to directly mimic Sea Hero Quest within a city landscape. The game was built in Unity 3D physics games engine (https://unity.com/) with all virtual structures modelled in Blender 3D modelling software (https://www.blender.org/). To control the back end and log experimental variables, we used Unity Experimental Framework (UXF), an open source set of packages for Unity 3D designed for human behavioral research. The game consists of five levels: one training level and four wayfinding levels. Each level in City Hero Quest was modelled from a corresponding Sea Hero Quest map, with the levels chosen matched along difficulty, entropy, and level effect size. Prior to building each structural model for City Hero Quest, a model city environment was built with urban panels designed to wrap around any 3D polygon. To build the model for each level, we extracted the base map for the corresponding Sea Hero Quest level (1, 16, 52, 56, or 67) and constructed 3D polygons similar in height to the sea environment. After constructing these polygons for each isolated structure within the base map, we tessellated the urban panels on each equally split face of each polygon. These panels were subsequently textured using urban textures (e.g. brick, marble) sourced from Poliigon (https://www.poliigon.com/) and baked for scene lighting to reduce the runtime demand on player platforms. This process was repeated for all levels.

Once built and baked, each model was imported to Unity 3D within separate scenes corresponding to each level and scaled up by 50 units. For the player, we imported a basic Unity car model (pre-fab), as this model contained the features needed for the game baked in. We edited the features of the car to mimic that of the boat in Sea Hero Quest, i.e. turn right, turn left, and accelerate, and de-activated additional features, i.e. lights, engine sounds. For each level, the car [player] was set at an initial position and pointing direction similar to that of the corresponding Sea Hero Quest level such that, at runtime, all players viewed the same perspective of the environment. In addition to the wayfinding component of each level, we added

a pointing task to each City Hero Quest level (not featured in the corresponding Sea Hero Quest versions). On each level, once players reached the final goal, three things happened: (1) Acceleration on the car [player] was reduced to zero, inhibiting forward and backward movement; (2) An arrow appeared extruded from the front of the car; & (3) Players were directed to rotate the car until the arrow was pointing in the direction of their starting location for that level. For both the wayfinding and pointing components of each level, player position and rotation variables were logged at each frame (0.036 seconds, frames per second were locked at 30 Hz) using the UXF position tracking package. This data was subsequently saved to each player's computer, split according to level.

Navigational Strategies Questionnaire (NSQ)

All participants completed a series of questionnaires prior to downloading the two virtual wayfinding tasks. We only analysed one of these, i.e. the navigational strategies questionnaire, for this chapter. The NSQ asks participants to rate themselves on 14 separate questions regarding their propensity to use maps – providing researchers with a 'mapping score' for each participant. Each question has a map-based answer and a non-map-based answer; some have a third alternative which is not coded as either (Brunec et al., 2018). This 'mapping score' has been previously found to correlate with navigational efficiency and posterior-anterior hippocampal ratio in humans (Brunec et al., 2019). For the NSQ in its entirety, please see the appendix.

Experimental design

We used a mixed task design as we looked at hypotheses both within and between subject-level. For the within-subject analyses we looked at the generalisability of wayfinding performance across our two virtual task environments, City Hero Quest and Sea Hero Quest, using Pearson's correlation. We further investigated the relationship between mapping propensity from the NSQ score and wayfinding performance using Pearson's correlation. For the between-subjects group-level analysis, we split the data into participants who grew up in cities and participants who grew up outside cities. These two groups had unequal variances and so we used Welch's

t-test to investigate differences in the means of wayfinding performance.

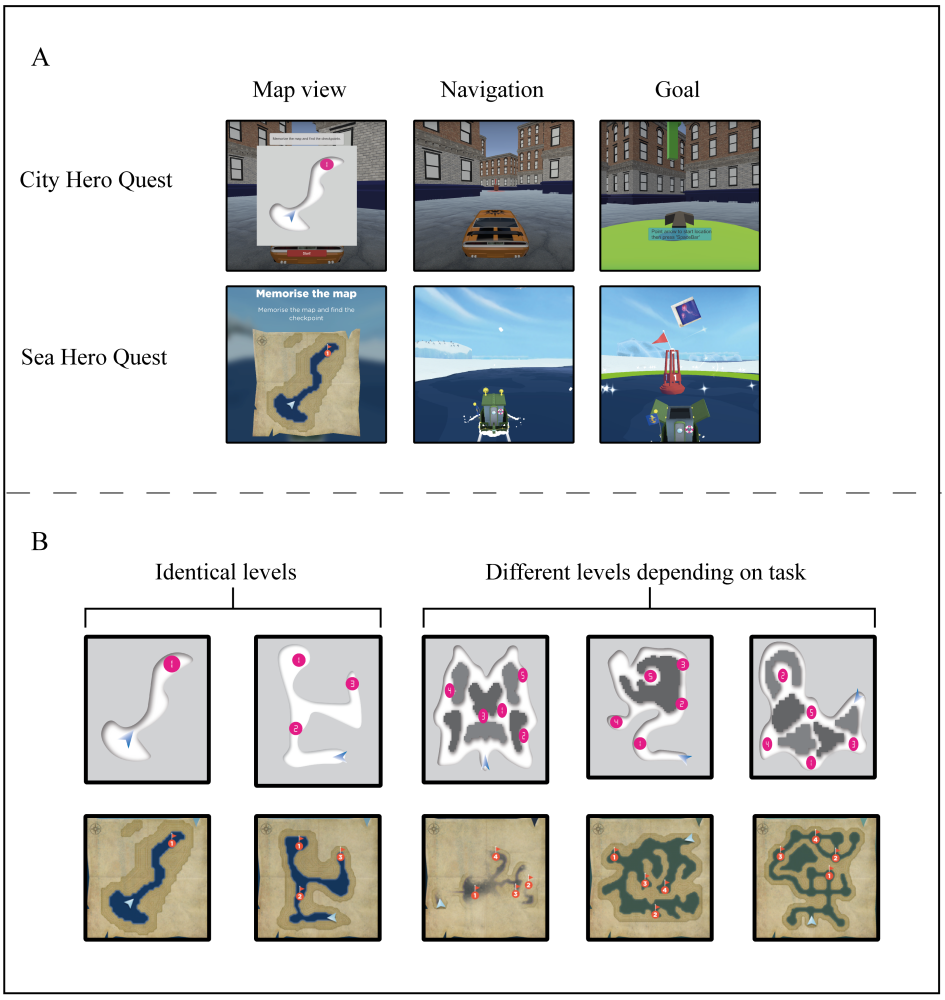

Figure 1. Overview of both virtual navigation tasks (Mobile: SeaHeroQuest. Desktop: CityHeroQuest). Both tasks involved memorising a map of the level layout including the position and order of mutliple goal locations. All the CityHeroQuest (CHQ) was modelled on specific levels of SeaHeroQuest (SHQ). (A) Participants in both tasks could view the map layout for an indeterminant amount of time. The navigation phase began immeadeatly after pressing start (CHQ) or close (SHQ). The navigation phase for CHQ consisted of controlling a car in an urban city environment using arrow keys or AWSD on a keyboard. After reaching the final goal location, participants had to point an arrow towards the initial start location. The arrow could be moved using either arrow keys (left/right) or A/D. The navigation phase for SHQ consisted of controlling a boat in a rural water-based environment using a smartphone or tablet. There was no additional task after reaching the final goal location, instead participants proceeded to the next level. (B) We used identical level layouts for the first two levels of both SHQ and CHQ to normalise the performance of each participant. The following levels were different so we could test the generalisability of results in two different virtual environments.

Figure 4.1: Overview of both virtual navigation tasks (mobile: SeaHeroQuest; desktop: CityHeroQuest)

4.2 Results

Generalisability of wayfinding ability

The wayfinding ability of participants quantified as distance travelled in Sea Hero Quest (SHQ) is predictive of real-world wayfinding ability (Coutrout et al., 2019). We were interested if wayfinding performance in SHQ would generalise to a second

novel virtual environment set in a different context. The SHQ virtual game is set in a water-themed world of rivers and lakes and played on a smartphone or tablet. To contrast this rural setting, we designed an urban city-themed replica titled City Hero Quest (CHQ) using the same maze layouts as in levels from SHQ and which can only be played on a desktop computer. The first two levels of both SHQ and CHQ were identical and used to control for video game ability and baseline visuo-spatial skill (Coutrot et al., 2018). Importantly, the subsequent three levels of each game were used to test participants' wayfinding ability for each virtual world. In total, there were three different level layouts for SHQ and three for SHQ-replicated levels in CHQ. The distance travelled in the first level of each game was used to normalise via division of the distance travelled in all other levels within the same game. Lastly, each normalised distance metric was z-scored for comparison across tasks. We found a significant correlation between the normalised distance travelled in SHQ with the normalised distance travelled in CHQ (Pearson's r = 0.42, $p < 0.001$; figure 2A).

Wayfinding ability associated with mapping propensity across task contexts

People vary greatly in their wayfinding abilities. Performance on wayfinding tasks is related to the ability to form an accurate mental map of the environment (Weisberg et al., 2014; Weisberg et al., 2016). We analysed participants' self-rating on the navigational strategies questionnaire (NSQ) which is designed to score ones' propensity to utilise a mental map of your environment in daily life (Brunec et al., 2018; see appendix A for the questionnaire). We found a significant correlation between mapping score and normalised distance travelled in both SHQ (Pearson's r = -0.13, $p < 0.001$; Figure 2B) and CHQ (Pearson's r = -0.19, $p < 0.001$; Figure 2B).

Advantage of growing up in rural environments on navigational ability

Recent evidence showed that street network entropy has a positive impact on wayfinding ability (Coutrot et al., 2020). Given that US cities tend to follow a

grid layout (and therefore a low street network entropy), we hypothesised that US participants who grew up in cities would have lower wayfinding ability than people who grew up outside of cities. We compared the normalised distance travelled of our city and non-city groups and found significant differences in both of the virtual wayfinding tasks, for City Hero Quest (Welch's t-stat = -3.4, $p < 0.001$) and for Sea Hero Quest (Welch's t-stat = -2.17, p = 0.031). We also tested the difference of mapping propensity from the navigational strategies questionnaire in our city and non-city groups and found a significant difference where the latter scored lower (Welch's t-stat = -1.97, $p < 0.05$; Figure 2D).

Generalisability of wayfinding ability across context for non-city people and relationship with mapping propensity

We found that people who grew up in cities perform worse on both of our virtual wayfinding tasks, even one which was designed in a city-themed landscape. We also found that mapping propensity negatively correlated with normalised distance travelled in both of our virtual wayfinding tasks. In other words, participants who self-report the use map-based strategies perform better in our virtual wayfinding tasks. Our next step was to investigate how mapping propensity could modulate wayfinding ability in people who grew up in cities vs. people who grew up outside of cities. We performed additional analyses after splitting the participants into groups of non-city (n = 643) and city (n = 175).

For the first group, we found a significant correlation for the participants who grew up outside cities between their wayfinding ability in City Hero Quest and mapping propensity (Pearson's r = -0.21, $p < 0.001$; Figure 2E). Similarly, we also found a significant correlation for participants who grew up outside cities between their wayfinding ability in Sea Hero Quest and mapping propensity (Pearson's r = -0.19, $p < 0.001$; Figure 2F). In other words, there is an advantage to self-reporting map-based strategies in both of our virtual environments if you grew up outside cities.

For the second group, we found a significant correlation for the participants who grew up in cities between their wayfinding ability in City Hero Quest and

mapping propensity (Pearson's r = -0.19, p = 0.012; Figure 2E). However, we found no significant correlation between their wayfinding ability in Sea Hero Quest and mapping propensity (Pearson's r = -0.026, p = 0.74; Figure 2D), suggesting that the advantage on wayfinding ability of a having high mapping score does not generalise across contexts for people who grew up in cities as they do for people who grew up outside of cities. Lastly, we found that the correlations between wayfinding ability and mapping score for people who grew up in or outside cities were significantly different (Fisher's z = 1.87, p = 0.03; tested using cocor R package, Diedenhofen & Musch, 2015).

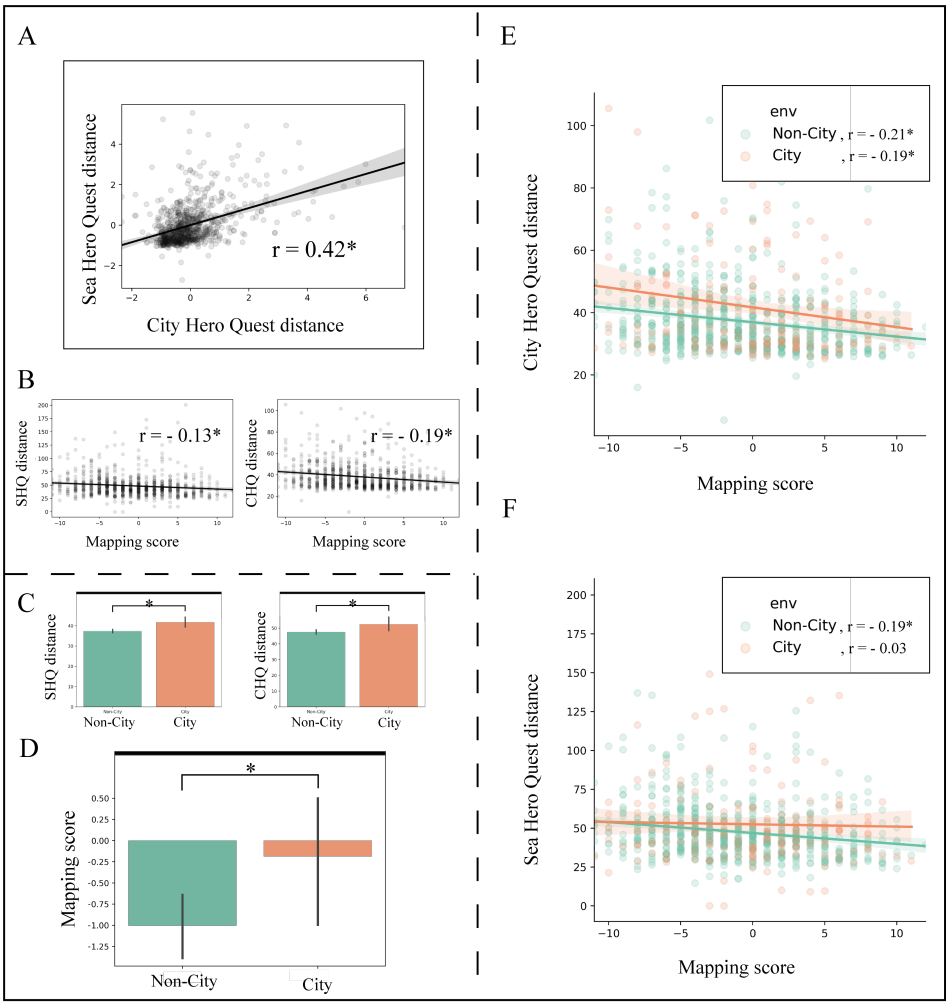

Figure 2. Growing up outside city benefits navigation ability across virtual task environments. The same group of participants completed two virtual wayfinding tasks, one on mobile (Sea Hero Quest or SHQ) and one on desktop (City Hero Quest or CHQ). (A) The performance on both tasks measured as normalised distance travelled was significantly correlated (Pearson's r = 0.42, p < 0.001). (B) Mapping score as measured by a Navigational Strategies Questionnaire significantly correlated with performance in both virtual wayfinding task (Left: Sea Hero Quest, Pearson's r = - 0.13, p < 0.001. Right: City Hero Quest, Pearson's r = - 0.19, p < 0.001). (C) We asked participants to report where in the United States they grew up, including whether they grew up in a city or not. Participants who grew up outside cities outperformed participants who grew up in cities in both wayfinding tasks (Left: Sea Hero Quest, t-stat = 2.78, p = .0056. Right: City Hero Quest, t-stat = 4.23, p < 0.001). (D) Participants who grew up outside cities scored significantly lower on mapping propensity in our Navigational Strategies Questionnaire (t-stat = 1.97, p < 0.05). (E) The distance travelled in City Hero Quest was significantly correlated with mapping score for people who grew up in cities (Pearson's r = - .21, p < 0.001) and for people who grew up outside cities (Pearson's r = - 0.19, p < 0.001). (F) The distance travelled in Sea Hero Quest was significantly correlated with mapping score for people who grew up outside cities (Pearson's r = - 0.19, p < 0.001), but not for participants who grew up in cities (Pearson's r = - 0.03, p = 0.74). The slopes of the two correlations were significantly different (Fisher's z = 1.87, p = 0.03; tested using cocor R package).

Figure 4.2: Growing up outside city benefits navigation ability across virtual task environments

4.3 Discussion

Recent evidence using data collected from 38 countries suggests that street irregularity, quantified using street network entropy, of where you grew up is important for developing one's wayfinding ability (Coutrot et al., 2020). We recruited a new

group of participants (n = 818) from the United States where city streets in certain large and populous cities very regular resulting in low street network entropy and therefore potentially detrimental to wayfinding ability. We tested these participants using the same virtual navigation task called Sea Hero Quest (first reported in Coutrot et al., 2018) as well as a novel virtual navigation task called City Hero Quest modelled after a subset of levels from the original game. We replicated the original finding finding that growing up in a city negatively impacts your wayfinding ability in the virtual environment Sea Hero Quest. We also show new supporting evidence that the impact of growing up in a city on navigation ability extends to a novel wayfinding game, City Hero Quest. Moreover, despite modelling City Quest within a city-theme, participants who grew up in cities still performed worse than people who grew up outside of cities.

Intriguingly, we also find that participants who grew up in cities self-report a higher degree of map-like strategy use. This finding is supported by other studies where participants who report living in cities that follow a grid-like pattern, also self-report a higher reliance on cardinal directions and is considered one aspect of an allocentric or map-like ability (Lawton, 2001). In a more recent study, Barhorst-Cates and colleagues (2021) also report that participants in a regular grid-like city (United States) self-reported more allocentric strategy use compared to participants living in an irregular non-grid-like Italian city. Moreover, the latter study also found that participants who grew up in the city with an irregular layout had improved wayfinding ability in a virtual environment task compared to the participants who grew up in the city with a regular layout – consistent with the results reported in this chapter.

We also identified a new potential advantage for participants who grew up in irregular environments. We found that participants who grew up outside cities and self-report use of map-like strategies have performance boosts in two wayfinding tasks with greatly different contexts (city vs. non-city water environments). However, participants who grew up in a city and self-reported use of map-like strategies, only had performance boosts in the similar city-like environment and not in the

water-based virtual environment task. Perhaps this indicates that people who grew up outside cities are more flexible in their wayfinding ability but also more flexible in their ability to switch between contexts and still maintain a high level of performance.

Limitations

The study reported in this chapter is limited by several aspects. First; the number of participants we were able to recruit differed markedly between the two groups, non-city (n = 643) and city (n = 175). The groups are still high-powered on their own to investigate association with self-reported measures, and this recruitment bias was expected based on the initial sample collected using Sea Hero Quest (Coutrot et al., 2020). We wanted to attempt to prescreen participants based on their geographical location but this was not allowed on our recruitment platform. We used Welch's t-test where appropriate when comparing means of the two groups since we may have violated the assumption of equal variance in the two groups.

Second, we did not counter-balance the order of the virtual environment tasks across participants. Participants received instructions to download the mobile game Sea Hero Quest prior to the instruction for the desktop game City Hero Quest. The order in which most participants completed the two tasks most likely followed the same order. We know from anonymous messaging that at least some participants completed the navigation tasks in the opposite order. Further investigation into timestamps of data collection could provide an accurate estimation of the order in which tasks were completed. One potential problem is that the correlation between participants who grew up in cities and self-reported map-like strategy use and wayfinding performance was influenced by first completing Sea Hero Quest. This would need more work to state whether the order had any influence.

Third, we never explicitly defined for participants what counted as a city vs. not a city. We gave participants the option to select 4 options about where they grew up: city, suburbs, rural, or mixed. We considered all categories besides city to be non-city. We also collected data about the exact street or streets in the US where they grew up which can be used in future analysis for a more fine-grained

classification and to exclude additional confounding factors such as growing up in the west vs. in the east US.

Fourth, the participants we were able to recruit had to overcome some quite challenging technological hurdles. It is difficult to quantify but some participants had to return the study being unable to complete one of a number of steps. The most challenging aspect was correctly installing both virtual environments. At the time of recruitment, Sea Hero Quest was no longer on the App or Google Play store. Instead we provided participants with a download link given to us by the games development company. Some participants struggled with giving Sea Hero Quest the proper certificates meaning they were unable to play the game. Sometimes, these issues could be resolved but they required extensive hands-on technical guidance which is difficult when dealing with upwards of hundreds of participants each day. In short, the participants we ended up recruiting may be more technological savvy and have a higher education level than the general participant pool offered through prolific which can be important for considering the results we presented here. However, this may also mean that our results are more likely to replicate in a psychology student population as usually recruited.

Fifth, the correlations we identified between performance and mapping propensity, are all of low r values and as such should be interpreted with caution as biases. One possible explanation is that measurement noise through online data collection reduced these r values. If true, we could repeat the experiment under more controlled in-person testing and expect to find larger effect sizes. However, it is likely the mapping score simply explained a small variance in the wayfinding performance.

Future Directions

There are several factors we did not control for in this study. These include, age, gender, education level, and the variation in grid-like layout across US cities. As we collected this demographic information including the city participants grew up in, our next steps will investigate if the reported findings hold up. We also collected additional relevant data regarding daily habits GPS use and commute durations and

types. One possibility is that participants living outside of cities commute to them for work and therefore get exposed to a greater variety of environments than participants living in cities.

4.4 Conclusions

Utilizing large-scale online recruitment, we investigated the role of home environment on navigational ability. We identified growing up outside cities as an advantage over growing up inside cities in terms of wayfinding performance across two different wayfinding tasks. Furthermore, we found that use of map-like strategy during navigation improved wayfinding performance overall. However, only participants who grew up outside cities could generalise the advantage of using map-like strategies across both of our tasks. One possible interpretation is that growing up outside cities improves your wayfinding ability and endows you with greater navigational flexibility, allowing you to generalise across tasks and contexts.

Chapter 5

Route planning in London Taxi Drivers

Our everyday life is experienced on a much larger scale than what is typically studied in laboratory experiments. There are a vast number of alternative actions that we can take in any given situation, and it is unknown how the brain is able to form future plans in the face of such large-scale complexity (Wu et al., 2018). We employed a homogenous group of volunteers, who all are faced with learning an identical large-scale state-space and make multi-step decision plans on how to traverse it, namely licensed London Taxi Drivers (Woollett et al., 2009).

Prior work show that London taxi drivers have the remarkable feature of enlarged posterior hippocampi compared to both a normal population sample (Maguire et al., 2000) as well as a matched sample of London bus drivers (Spiers et al., 2006). However, evidence for navigational ability correlated with hippocampal size in a normal population is lacking, suggesting taxi drivers are unique in this regard (Weisberg & Ekstrom, 2021). In a virtual London navigation task, BOLD activity was identified in the hippocampus during route planning in taxi drivers (Spiers & Maguire, 2006), along with bilateral frontopolar activation during spontaneous route planning. However, the question remains of how these drivers are able to leverage their extensive knowledge of the London street network to facilitate the everyday demands of being a London taxi driver.

In this study, we tested whether taxi drivers utilise bottlenecks in the London

street network to facilitate route planning using BOLD fMRI. Specifically, we suggest that taxi drivers initially plan routes across modules, i.e. areas of London that taxi drivers consider to have defined borders, before planning within these areas. A recent study demonstrated that during route call-outs, taxi drivers' response time changed as a function of encountering bottlenecks in the state-space or street network (Griesbauer, 2021). The use of bottlenecks to facilitate route planning in large state-spaces is also supported by theoretical consideration to minimize planning demand from an information-theoretic perspective (Mcnamee, 2016).

We designed a novel route planning task which spanned 360 unique locations around the London street network forming 120 unique routes. Each route contained a start location, a goal location, and a third probe (target) location. We asked taxi drivers, for each route, if their mentally planned route included the probe location or not. By varying the distance of the probe location from bottlenecks in the state-space, we aimed to better understand the neural code of route planning in London taxi drivers.

First, we hypothesise that our route choice task activates the hippocampus during the mental planning phase given the past literature indicating hippocampal processes in planning and spatial memory (Miller & Benditto, 2021; Burgess et al., 2002; Kaplan et al., 2020). Previous work in Licensed Taxi drivers found bilateral frontopolar activation during spontaneous replanning (Spiers & Maguire, 2006). Therefore, we also hypothesise that planning demand, defined using our proxy of distance from predefined boundaries in the street network of London, parametrically modulates prefrontal regions including frontopolar cortex (Griesbauer et al., 2021).

5.1 Methods

Participants:

The study was approved by the ethics committee for UCL Division of Psychology and Language Sciences (fMRI/2021/001). For online testing, ethics approval was received from UCL Department of Experimental Psychology (EP/2018/008).

Twenty-two participants (age ranging from 35 to 67 years, mean age: 52.2, 22 male, 21 right-handed) completed the full experiment including one training session. All participants were London licensed taxi drivers holding green badges authorized by Transport for London. The average years of taxi driver experience was 17.5 years (min = 1, max = 38, std = 10.6).

Task:

We designed a novel route planning task using 450 distinct locations scattered around London. Google street view images were extracted from Google Earth (high-resolution 4800x2643 pixels) at specific longitudes and latitudes for each location. Images were compressed to reduce loading delays during stimulus presentation, average compression from 6MB to 300KB while maintaining the same resolution. The 450 locations consisted of 150 starting locations, 150 goal locations, and 150 target locations. Images were presented using Psychopy (Pierce, 2007).

The task was divided up into two components (Figure 1). (1) Drivers were asked to mentally plan the route between a start and goal location just as they would follow the direct line during a knowledge exam (Griesbauer, 2021). We displayed the street view image and name of the start location at the bottom of the screen for 4-8 seconds, uniformly sampled. After the jittered delay, the street view image of the start location was replaced with the street view image of the goal location, while the text of the start location remained at the bottom of the screen. Along with the appearance of the image for the goal location, a text appeared at the top of the screen stating the location of the image. The street view image of the goal location was present on screen for 4-8 seconds, uniformly sampled. (2) We introduced an accept-reject task where each participant evaluated whether a third street, the target street, formed part of the planned route from start to goal. After the jittered delay for the goal location, the image was replaced by a street view image of a target location along with text in the upper third of the screen, showing the name of the street. The target location image was present until a button press response (yes or no) was made using a button box placed by the right hand of the participant. The yes or no response was mapped to either a response made with the thumb or the index

finger, counterbalanced across participants. Each trial consisted of both components – mental route planning and accept-reject of target street – and was separated with an ITI of 3.5-4.5s. The trial order of each route was maintained for all participants.

Online training and Questionnaire:

Each participant completed 30 training trials at home in the week prior to coming in for the scan. We wanted to familiarise the participants with the mental route planning and the accept-reject aspect of the task. Participants also completed a questionnaire including the Navigational Strategies Questionnaire (Brunec et al., 2018), a demographic questionnaire, and a taxi-driver-specific questionnaire regarding their experience and preferences as a driver.

Routes:

We selected routes which sampled the London street network based on a number of route characteristics. These included the following characteristics: route path distance, number of turns, cardinal direction of the route, target distance to start location, target distance to goal location, detour distance of target, target being on route, target being off route, target being on an A-road (major roads in United Kingdom), and route crossing the river. The different routes and targets also uniformly sampled the geographical area of the 3-mile radius around Charing Cross, the center point for the Knowledge (Figure 1).

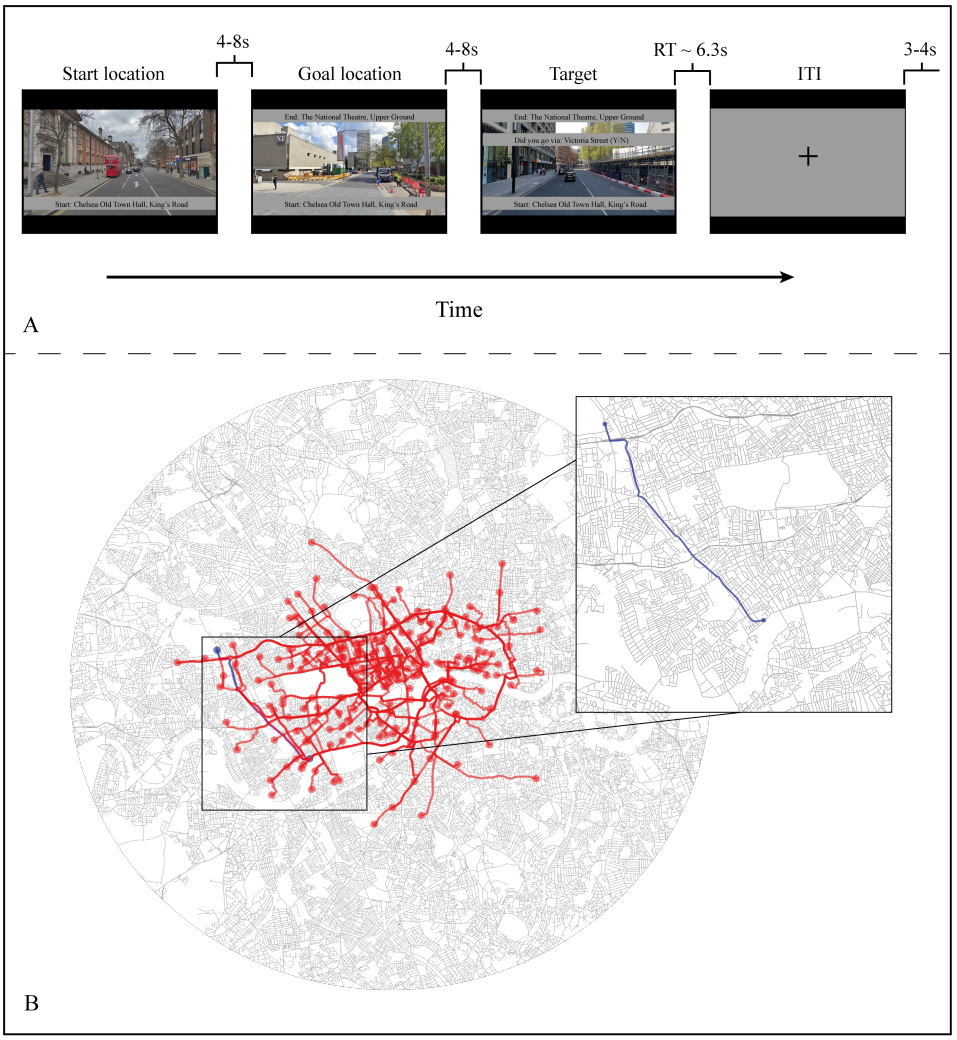

Figure 1. Overview of the route planning task. (A) The stimuli as seen by each participant on the first trial. Each trial began with showing a street view image of the starting location for 4-8 seconds. The landmark along with the street name was shown on the bottom of the screen. The goal location appeared following the display of the starting location. The street name of the goal location also appeared at the same time, and the starting location street name stayed on screen. The target location appeard following the display of the goal location alongside street name. A question was asked "Did you go via: ____ ?" for each trial propmting a yes or no response. (B) Overview of all the routes experienced during the scanning session on a backdrop of the London street network surrounding six-mile radius around Charing Cross. The first trial of the session is shown in a figure pop-out in the blue route.

Figure 5.1: Overview of the route planning task

MRI acquisition parameters:

Scanning was conducted at the Birkbeck-UCL Centre for Neuroimaging (BUCNI) using a 3-T Siemens Prisma MRI scanner with a 32-channel head coil. Each scanning session lasted around 1 hour and was separated into a maximum of eight runs (each approximately 5 - 7 minutes with 15 trials per run). Approximately 280 functional scans were acquired during each run, depending on participant performance (time taken per trial). A multi-band sequence was used (repetition time [TR] = 1450

msec, echo time [TE] = 35 msec, acquisition time (average) = 3276 seconds, flip angle = 70°). The scans were whole-brain (72 slices) with a multi-band acceleration of 4, slice thickness of 2 mm, spacing between slices of 2 mm, resolution / voxel size of 2x2x2 mm, field of view of 212 mm, and phase encoding A >> P. A top-up sequence was also performed at the end of the scanning session with reverse-phase encoding. A fieldmap was acquired after the first run. A standard T1-weighted high-resolution structural scan (MPRAGE) was acquired about half-way through each scanning session (TR = 2300 msec, TE = 30 msec, 1x1x1 mm resolution). Ear plugs and foam padding on each side of the head were used for noise reduction and to reduce head movements. Task stimuli were projected on a rear projector screen and mirrored for the participant's view. The mirror was adjusted for each participant to see full screen.

Experimental Design and Behavioural analysis:

Each trial contained an accept-or-reject response for the question "Did you go via: X street?", where X street refers to a street that may or may not lay on the direct route between a given start and goal location. We collected response times for 120 route planning trials during the fMRI scanning session. Subsequently, we investigated the consistency of responses across participants and factors that may influence the response times.

We calculated optimal paths between each start and goal location pair by first extracting the street network graph of Greater London, UK, from OpenStreetMaps using custom Python code and the osmnx package (Boeing, 2017; Haklay & Weber, 2008). The street network graph contained 568,473 nodes and 1,085,199 edges. Optimal paths were also calculated from start, to target, to goal location. The detour deviation was computed as the fractional increase in distance by going through the additional target street location, and was used in an ordinary least squares, linear regression model to account for response times.

MRI preprocessing and analysis:

Preprocessing was performed using fMRIprep (Esteban et al., 2018) with default settings (see Appendix for additional preprocessing details). First- and second-level GLM analysis was performed using SPM12. The first three dummy scans per run were automatically discarded before acquiring volumes. We constructed general linear models contrasting different stages during the route planning task. Specifically, we contrasted the onset of the initial starting street with the onset of the target street, to look at brain regions selectively active during planning. We also constructed a parametric general linear model to investigate the influence of target placement with respect to boundaries in the London street network (Griesbauer & Spiers, 2021). We hypothesized that London taxi drivers segment the London's street network into a hierarchy of modules, where each module is separated by boundaries. We classified each target as being on boundary, near boundary, or far from boundary, with an assigned value of 1, 2 or 3 in order of increased planning demands.

5.2 Results

Behaviour

Our route planning task consisted of 120 carefully selected routes that uniformly sampled across a number of factors. We investigated which of these were behaviourally relevant in a linear ordinary least squares model. We found no correlation with number of turns, path distance from start to goal location, or Euclidean distance. We also categorised the planning demand of each target street in terms of its location relative to the nearest boundary. A target on a boundary had a low planning demand (value of 1), a target near a boundary had a higher planning demand (value of 2), a target far away or near no boundaries had the highest planning demand (value of 3). There was no significant correlation identified between planning demand and response times.

We further calculated the increase in path distance by the inclusion of the target street. Path distance is the distance along the shortest path between each starting

and goal location. To calculate the increase in path distance, we summed the path distance between start and target street with the path distance between the target and goal street. We then divided the change in path distance by the original path distance. We found a negative correlation between response times and the fractional increase in path distance ($r = -0.56$, $p < 0.001$; Figure 2C). This implies that taxi drivers were faster at rejecting a target street if it were far away from the optimal path.

Participants varied substantially in the amount of agreement across different routes (Figure 2A). We quantified this as a fractional disagreement value per trial, where 0.5 means maximal disagreement and 0 means complete agreement. Response times were found to be positively correlated with the amount of agreement ($r = 0.67$, $p < 0.001$; Figure 2B).

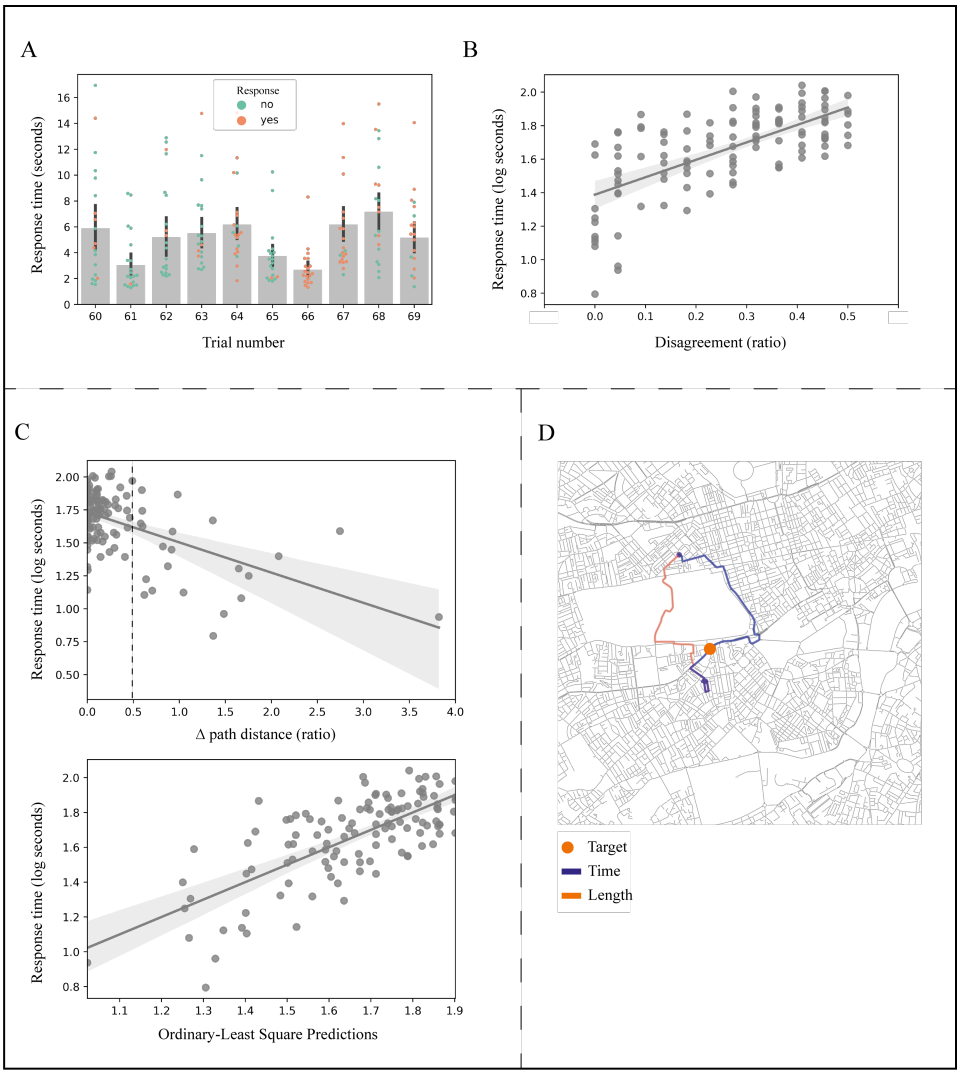

Figure 2. Response times during route planning task. (A) Example response and time variability across trials. Trial number 61 had 95% agreement that target did not lay on route, while trial 66 had 100% agreement thtat target lay on route. (B) Response time is significantly correlated with disagreement, measured as proportion of taxi drivers with the same response in a given trial (Pearson's correlation; $r = 0.68$, $p < 0.001$). (C) Top: Response times influenced by the increase in path distance caused by the addition of a target location to any given start-goal route (Pearson's correlation; $r = -22$, $p = 0.01$). Bottom: Both change in path distance ($ß = -0.131$, t-stat = 7.5, $p < 0.001$) and response disagreement ($ß = 0.84$, t-stat = -4.3, $p < 0.001$) significantly account for variance explained in multiple linear regression ($R^2 = 0.52$, F-stat = 62.8, $p < 0.001$). (D) Two different optimal routes from a trial with maximum disagreement among taxi drivers are computed using Dijkstra's algorithm; orange is based on path distance; blue is based on edges weighted by speed limits (Python package OSMNx through OpenStreeMaps).

Figure 5.2: Response times during route planning task

fMRI results

Bilateral hippocampal and bilateral retrosplenial cortex activation driven by route planning

Our fMRI task was designed to probe taxi drivers' knowledge of a wide variety of streets and areas of London (Figure 1). Each trial of the task presented three im-

ages in succession, start image, goal image, and lastly target image, and planning across all three locations was required to solve the accept/reject task. We investigated which brain regions were involved in route planning in our task with a GLM contrasting the onset of the target locations by the onset and duration of the start locations from all 120 trials. The start locations served as a baseline where planning is limited since neither the goal or target location is known. We also modelled the inter-trial-intervals, and the button responses as separate regressors.

We found that bilateral hippocampus was preferentially activated during the target vs. start location contrast indicating hippocampal involvement in mental route planning of London licensed taxi drivers (Figure 3; left: t-stat = 10.93, $p < 0.05$; right: t-stat = 9.03, $p < 0.05$). We also identified bilateral retrosplenial cortex activation in the same contrast (left: t-stat = 11.85, $p < 0.05$; right: t-stat = 10.1, $p < 0.05$). All results are whole-brain family-wise error-corrected.

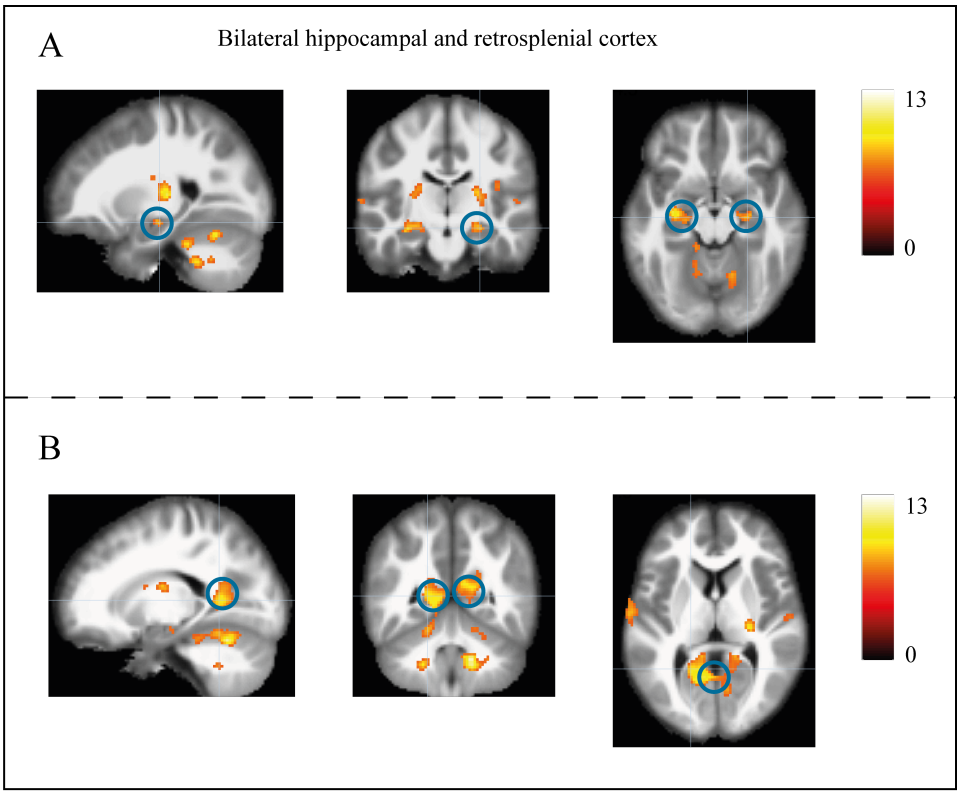

Figure 3. Bilateral hippocampal and bilateral retrosplenial cortex driven by route planning. We constructed a GLM by contrasting the onset of the target street with the onset and duration of the start street image. (A) We find significant activation in bilateral hippocampus (whole-brain family-wise error corrected, $p < 0.05$). (B) We also identify significant activation in bilateral retrosplenial cortex (whole-brain family-wise error corrected, $p < 0.05$).

Figure 5.3: Bilateral hippocampal and bilateral retrosplenial cortex driven by route planning

Boundary-related planning demand elicits frontopolar, preSMA, and ACC activation

The target streets in our route planning task were specifically selected to probe the influence of boundaries in the London street network. A previous study by Griesbauer & Spiers (2021) identified a number of boundaries that London Taxi drivers identified although these had varying degrees of agreement. We hypothesised that the planning demand of each target street would depend on the distance of a target street from the nearest boundary. Theoretical work on planning in spatial navigation suggests that the process of planning first occurs over global modules, where each module is separated by boundaries (McNamee, 2016; Weiner & Mallet, 2003). As such, target streets that form part of a module boundary should be easier to access and use in route planning, compared to targets that are further away from a bound-

ary. These boundaries can also be viewed as bottlenecks in the state-space which may also coincide with main roads and bridges.

We constructed a parametric GLM, using all 120 trials, with planning demand varying from low (target street is part of boundary; Figure 4A), to medium planning demand (target street is adjacent to a boundary street), to high planning demand (target is far from any boundary). We found significant activation in ACC (right: t-stat = 6.54, $p < 0.05$; whole-brain family-wise error corrected). We also identified significant activation in bilateral frontopolar cortex (right: t-stat = 6.54, $p < 0.05$; whole-brain family-wise error corrected).

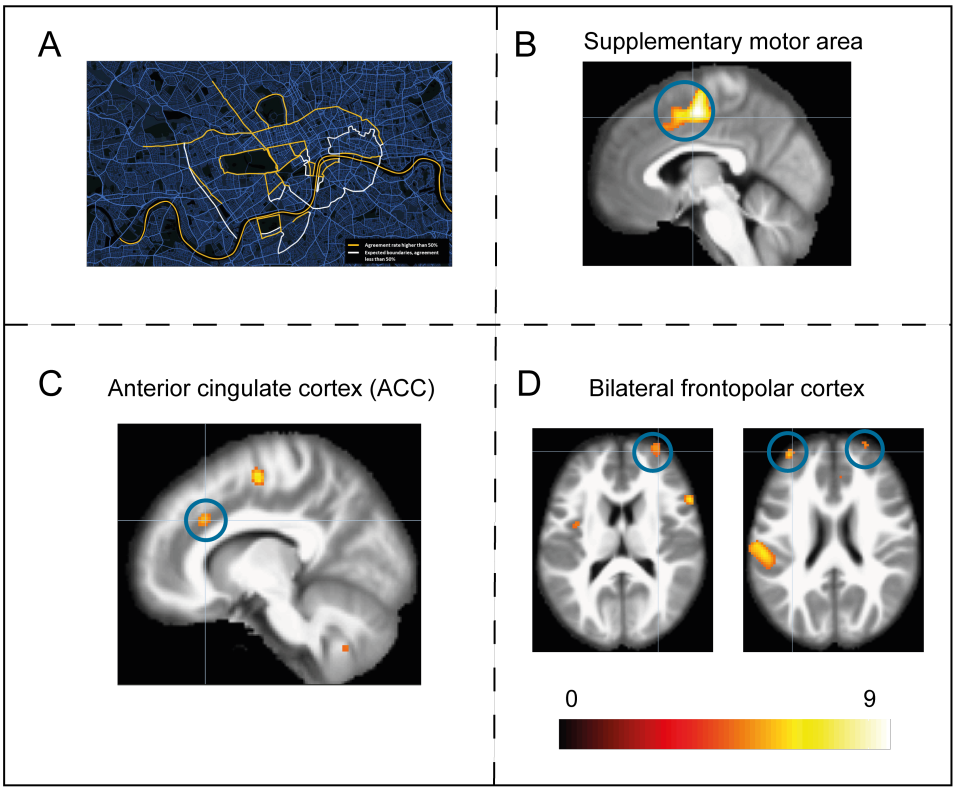

Figure 4. Boundary-related planning demand influence on ACC, suppl. motor area, and frontopolar cortex. (A) Boundaries of London street network as defined by Taxi Drivers (figure from Griesbauer & Spiers, 2021). Yellow boundaries have agreement rate higher than 50%, and white boundaries have agreement of less than 50%. (B) We find significant activation in a parametric regression GLM (see methods for details) in suppl. motor area (family-wise error corrected, $p < 0.05$). (C) The right anterior cingulate cortex (ACC) was significantly modulated by variation in boundary-related planning demand (family-wise error corrected, $p < 0.05$). (C) Bilateral frontopolar cortex was also modulated by boundary-related planning demand (family-wise error corrected, $p < 0.05$).

Figure 5.4: Boundary-related planning demand influence on ACC, suppl. motor area, and frontopolar cortex

5.3 Discussion

In this study we investigated the neural code of mental route planning in large state-spaces, namely the street network of London. We recruited licensed London taxi drivers, who have extensive knowledge and experience planning within this state space and had them perform a novel route choice task using a diverse set of street view locations around London. We identified the joint contribution of retrosplenial cortex and the hippocampus in the retrieval and usage of target streets during the process of mental route planning. This finding is consistent with recent evidence of goal coding in these regions when one is navigating in highly familiar, real-world environments (Patai, Javadi et al., 2019).

The hippocampus is implicated in both memory and spatial navigation processes (Epstein et al., 2017; Eichenbaum, 2017). Given the proposed structural changes that occur in the hippocampus of licensed London taxi drivers, an important element to understanding its role in spatial navigation may come from human patient studies where the hippocampi are damaged (Maguire et al., 2000; Maguire et al., 2006). In one amnesic case study, a former taxi driver sustained bilateral hippocampal damage and subsequently navigated in a virtual version of London alongside a group of matched controls (Maguire et al., 2006). Several aspects of the spatial navigation process were still intact including "general orientation in London, rich and detailed topographical knowledge of landmarks, and the spatial relationships between them, or even navigating along some routes". The main impairment was only apparent when navigating along smaller roads, in the UK known as non-A roads, as opposed to main roads known as A roads, where the patient would easily get lost compared to controls. In the current study, we identify hippocampal involvement while taxi drivers imagined travelling 120 routes through London. One such explanation for this involvement is that the hippocampus is active due to the nature of the street view images priming discrete memory retrieval, ultimately resulting in a memory task instead of a planning task. However, when contrasting the onset of the target location with the onset of a start location – in which one would expect equal engagement of the hippocampus – we find hippocampal and retrosplenial activity, suggesting their involvement in imagined routes. Interestingly, a study of patients with primary damage to their hippocampus bilaterally also reported impairment to their ability to imagine novel experiences (Hassabis et al., 2007). Some of the routes in our task may be novel for any given participant. As such, the act of imagining routes may be a possible explanation for hippocampal involvement.

To further understand how taxi drivers are able to seamlessly plan in a large state-space, we turned to testing the hypothesis that they utilise characteristics of the street network, specifically the bottlenecks. These bottlenecks were operationalised as the boundaries separating different regions of London, for instance the bridges crossing the river Thames which cut through central London, dividing it in two.

Recent behavioural evidence suggests that taxi drivers use these bottlenecks in their route planning (Griesbauer, 2021). We went one step further, and manipulated the distance of target streets from boundaries in our novel route planning task. We split these distances into three categories: low, medium, and high planning demand corresponding to the target street existing either on bottlenecks, adjacent to a bottleneck, or far from any bottleneck, respectively.

We identify brain activity in bilateral frontopolar cortex parametrically modulated by the boundary-related planning demand. This result replicates an earlier finding by Spiers & Maguire (2006), in which they found bilateral frontopolar cortex was similarly activated compared to a control condition during spontaneous replanning of a route while driving in a virtual simulation of London. We also identified additional brain regions modulated by the planning demand, including the right ACC and pre-SMA. We found no significant correlation between response times and planning demand, nor did we find a difference in response between trials with target street on bottleneck vs. trials with target street away from bottlenecks. This suggests that our results are not driven by differences in response times across trials. The boundary-related planning demand is our approach at probing a hierarchical representation of the London street network. Previous work has investigated hierarchical representations in the general population using learned state-spaces such as virtual subway maps (Balaguer et al., 2016),

The planning demand modulation in pre-SMA is similar to another finding in a multi-step planning task (Balaguer et al., 2016). In their task, participants navigated along a fictional subway network, separated hierarchically into unique subway lines. preSMA activity increased with the cost of hierarchical plan and context switches of subway lines. Kaplan et al. (2017) also identified lateral PFC along with pre-SMA as brain regions involved in demanding planning. In their task, participants were shown a map of a maze and asked to choose the shortest path between alternative routes. Trials with small differences in alternative routes, and therefore higher planning demand, resulted in the observed increased activity in preSMA and lateral PFC. Our findings are consistent with this work as our proxy

of planning demand correlated with activity in preSMA and frontopolar PFC.

We also found dorsal ACC activation in our parametric modulation of boundary-related planning demand. This brain regions is involved in several cognitive domains including error detection (Carter et al., 1998), conflict monitoring (Botvinick et al., 2004), model updating (O'Reilly et al., 2013), foraging (Kolling et al., 2012), and tracking multiple agents in a spatial task (Yoo et al., 2021). Recent evidence from a virtual navigation task also demonstrated the involvement of dACC in backtracking behaviour (Javadi et al., 2018). Interestingly, the addition of the target street in our task may sometimes function as mental backtracking where participants need to think back on the route they already mentally planned to make the decision whether or not the target street lay along that route. Another plausible explanation is that ACC is commonly seen in conjunction with task demands and effortful behaviour (Porter et al., 2019). In our task, we hypothesise that target streets placed far from boundaries require more planning. If more planning leads to more effortful behaviour then this could be consistent with our results. However, we did not find any effect of target distance to boundary on response times as might be expected if indeed more effort is required for these routes. Exactly which factors drive response times in our mental route planning task will require further investigation and detailed analyses on the route characteristics. There may also be interesting individual variability where some taxi drivers weigh certain route characteristics as more planning intensive compared to others, for instance number of turns or distance of a route.

Limitations and Future Directions

We describe our task as route planning; however, we are not able to evaluate the planning process in terms of each road a given participant would select to navigate from start to goal location. Instead, our task asks participants to mentally plan the route and selectively accept or reject the target street we show them. The task could be rephrased as a route memory task if it was not for the case that some of the routes from a given start to goal are bound to be completely novel even if the locations are familiar.

Our response time analysis has so far focused on the average response time per trial across participants. It is important that we better understand individual variability in the response times by also modelling on the level of each participant. There is substantial variability in the agreement of accept vs. reject on certain trials which is correlated with response times (Figure 1B). One possible explanation is that taxi drivers form certain preferences, and analysis on individual level may reveal a consistent division where subsets of taxi drivers are consistent amongst each other. We will be increasing the number of licensed taxi drivers who complete the behaviour task to help understand the variability in route preference.

We hypothesise that taxi drivers use boundaries and bottlenecks in the state-space to efficiently plan and reduce the computational demands (McNamee et al., 2016). The boundaries were quantified using a separate study where a different set of taxi drivers drew separating lines on a map of London of where they considered these boundaries to be (Griesbauer et al., 2021). In the current study, we consider any given point on these boundaries as a bottleneck in the state-space. However, these boundaries, and therefore bottlenecks, are subjectively defined and some of the boundaries only had

All of our fMRI route analysis thus far looked at the contribution of all trials. However, our design inherently includes different trial types including, reject vs. accept trials, target on appearance run (direct path), and varying degrees of difficulty based on the deviation of target path from a plausible route. In future analysis we will investigate the network of brain regions involved in these different trial types and compare them to predictions based on the literature.

5.4 Conclusions

In conclusion, we investigated route planning in London taxi drivers to understand the ease at which humans can solve the problem of planning in large state-spaces. We found that the hippocampus and retrosplenial cortex are activated by route planning, solidifying a role for these regions for long-term spatial memory in real-world navigation settings. Furthermore, we found a neural signature of hierarchical plan-

ning demand parametrically modulated by activity in lateral frontopolar cortex, preSMA, and dACC. The lateral frontopolar response is consistent with previous reports of spontaneous route planning during a virtual driving task in London taxi drivers (Spiers & Maguire, 2006). However, the current findings take this signature one step further by operationalizing the planning demand as the distance between streets and bottlenecks in the state-space. We interpret our data to suggest that taxi drivers exploit these bottlenecks during route planning in order to make the challenging problem of planning in a large state-spaces solvable (McNamee, 2016). When considering the theoretical extent of this work, it is important to note that the findings of this chapter posit a role for regions (e.g. retrosplenial cortex and frontopolar PFC) beyond the hippocampus in spatial navigation and, specifically, spatial planning within large state space. These findings must be considered in terms of the stereotypically hippocampal-dependent cognitive map theory, as our results suggest a more nuanced role across additional brain regions in large state-spaces.

Chapter 6

General Conclusions

The study of spatial navigation is often based on two strategies: place-based and response-based. Early rodent lesion work and human fMRI studies suggests that these map onto neurobiological substructures in the hippocampus for place-based strategies and in the caudate for response-based strategies. Recent work suggests this dichotomy may need some revision as the caudate has been implicated in flexible spatial and non-spatial tasks (Gahnstrom & Spiers, 2020). This thesis suggests new ways forward including a stronger focus on investigating the computations underlying different spatial strategies observed in navigation behaviour. We also design a novel spatial navigation task to investigate dynamic flexible behaviour in a normal population and a novel route choice task to investigate mental planning in large state-spaces with a population of expert navigators, licensed London taxi drivers. Both our fMRI studies implicate a network of brain regions supporting flexible navigation and mental planning that go beyond hippocampal as the single locus of a cognitive map (Ekstrom et al., 2017).

The investigation into flexible behaviour is not unique to spatial navigation. In decision-making, a wide range of studies have tried to understand flexible behaviour from many perspectives using different classes of rational, normative, descriptive, aggregate, or mechanistic computational models (Corrado et al., 2009). In recent years, computational modelling approaches have gained popularity in spatial navigation (de Cothi et al., 2020; Mattar & Daw, 2018), but no study we know of has yet to fully take advantage of the possibilities (Wilson & Collins, 2019). We review

and pilot the feasibility of some of these possibilities using reinforcement learning models, task simulations, and freely moving human behaviour in chapter 2 of this thesis. We suggest successor representation and monte-carlo tree search as better matched to observed human behaviour compared to a Q-learner algorithm although caution our interpretation for lack of data collection as a result of the COVID-19 pandemic.

Subsequently, we used some of the lessons learned in chapter 2 to devise a new study to investigate goal coding in humans using desktop virtual reality and functional magnetic resonance imaging (fMRI). We simulated agents on 180,000 pseudo-randomly generated maze configurations to create mazes which afforded ideal circumstances to investigate goal coding in the whole-brain with humans and in dorsal CA1 with rodents. We present our human findings in chapter 3, demonstrating that our participants were able to rapidly learn and relearn new hidden goal locations after a single trial. We identified a network of brain regions involved in updating these hidden goal locations including bilateral hippocampus, lateral orbitofrontal cortex (OFC), bilateral caudate nucleus, dorsal cingulate cortex, and dorsomedial prefrontal cortex. Recent work has suggested that the caudate nucleus is involved in representing the transition structure, or the action-outcome associations, of a task (Gahnstrom & Spiers, 2020; Sharpe et al., 2019). We report evidence consistent with this theory where the caudate nucleus was parametrically modulated by Euclidean distance to goal in our task. This is consistent with other recent findings by Javadi and colleagues (2019). Lastly, medial OFC, left hippocampus, and right hippocampus/parahippocampal regions track egocentric hidden goal direction in our task. Taken together, these results overlap with a suggested dissociation of OFC subregions, where lateral regions are involved in learning state values (switching between behavioural goals), while medial OFC is involved in ongoing decision making (Noonan et al., 2010).

Behavioural flexibility and the formation of cognitive maps in spatial navigation tasks is surprisingly varied (Weisberg et al., 2014). In light of the neural correlates of flexible navigation we found in chapter 3, we wondered what could cause

this variation in the general population, as this could possibly limit the generalisability of our results. The typical sample size recruited for neuroimaging experiments makes it next to impossible to investigate this variability in a lab setting. Leveraging massive online testing, recent evidence suggests that part of this variability in behavioural flexibility, measured as wayfinding ability in the mobile navigation game Sea Hero Quest, was found to be related to the complexity of the street layout of where you grew up (Coutrot et al., 2020). In chapter 4, we extended this finding by comparing the wayfinding ability of participants in the United States who grew up in cities with those who grew up outside of cities. We also found that a navigational strategies questionnaire (NSQ; Brunec et al., 2018) provided a possible explanation for the difference in wayfinding ability. Participants with a high mapping score, which indicates more behavioural flexibility during spatial navigation, outperformed participants with a low mapping score in two virtual tasks, Sea Hero Quest, and a novel city-themed variation, City Hero Quest. Participants who grew up outside cities, in places with higher street network complexity than in US cities, had improved wayfinding abilities and improved generalisability of behavioural flexibility across task contexts compared to those who grew up in cities. All our reported mapping score correlation feature low effect sizes so need to be interpreted with caution as slight biases. Future neuroimaging studies could recruit participants based on understanding the baseline variability in wayfinding ability and investigate how that influences neural correlates of flexible behaviour during complex navigation.

The variability of navigational ability in the general population described in chapter 4 begs the question: What makes expert navigators, those who have mastered behavioural flexibility during complex navigation tasks, different? Luckily, we were able to study expert navigators who have spent years learning and planning within one of the most complex street networks in the world – licensed London Taxi Drivers (Griesbauer & Spiers, 2021). We devised a novel task targeting the aspect of their job that would drive the most planning demand, namely planning a new route (Spiers & Maguire, 2006). Our route planning task was carefully

designed to sample locations and routes that spread all across London. We found bilateral hippocampal and retrosplenial cortex activation during planning across our 120 routes. Rodent studies investigating spatial planning have identified the ability of the hippocampus to simulate future trajectories which may even translate to humans but further work is necessary (Pfeffer & Foster, 2013; Kurth-Nelsen et al., 2016). We also investigated the neural correlates of planning demand. We found that anterior cingulate cortex, pre-SMA, and bilateral frontopolar cortex was parametrically modulated by our proxy of planning demand defined as the distance of target street from nearby boundary, where London boundaries were defined by a separate group of licensed Taxi Drivers (Griesbauer & Spiers, 2021). Our findings in ACC and pre-SMA are consistent with other reports of the involvement of these regions during hierarchical planning of a virtual subway task (Balaguer et al., 2016). The frontopolar cortex activation is also consistent with a report of spontaneous route planning during spatial navigation in a virtual simulation of London (Spiers & Maguire, 2006).

Overall the body of work presented in this thesis provides evidence for a more nuanced relationship between stereotypically spatial brain regions and flexible navigation. In chapter 2, we found human trajectories in freely moving tasks mostly overlap with model-based and successor representation algorithms. In chapter 3, we found a network of brain regions to include caudate, hippocampus, and PFC regions, involved in updating goals information in an ever-changing environment. In chapter 4, we found that where you grew up influenced future navigation ability, tested using two separate video games, and that this effect was moderated by a self-reported measure of mapping propensity. In chapter 5, we identified hippocampal and retrosplenial involvement with mental route planning (and frontopolar activity correlated with variations in planning demand) in a group of expert navigators (London taxi drivers). Taken together, these results provide evidence for a distributed network of brain regions supporting flexible navigation and a role for regions beyond the hippocampus in supporting a cognitive map - if such exists. In addition, the work considered in this thesis could be further reconciled by appropriate application of

computational modelling methodologies that, when fit to observed behaviour, can be used to derive latent decision variables that can be investigated in distributed neural codes, not only in humans but other species as well.

Chapter 7

References

Aghajan, Z. M., Acharya, L., Moore, J. J., Cushman, J. D., Vuong, C., & Mehta, M. R. (2015). Impaired spatial selectivity and intact phase precession in two-dimensional virtual reality. Nature neuroscience, 18(1), 121-128.

Akrami, A., Kopec, C. D., Diamond, M. E., & Brody, C. D. (2018). Posterior parietal cortex represents sensory history and mediates its effects on behaviour. Nature, 554(7692), 368–372. https://doi.org/10.1038/nature25510

Ambrose, R. E., Pfeiffer, B. E., & Foster, D. J. (2016). Reverse Replay of Hippocampal Place Cells Is Uniquely Modulated by Changing Reward. Neuron, 91(5), 1124–1136. https://doi.org/10.1016/j.neuron.2016.07.047

Andersen P, Morris R, Amaral D, et al. (eds) (2006) The Hippocampus Book. Oxford: Oxford University Press.

Anggraini, D., Glasauer, S., & Wunderlich, K. (2018). Neural signatures of reinforcement learning correlate with strategy adoption during spatial navigation. Scientific Reports, 8(1), 10110. https://doi.org/10.1038/s41598-018-28241-z

Aoki, Y., Igata, H., Ikegaya, Y., & Sasaki, T. (2019). The Integration of Goal-Directed Signals onto Spatial Maps of Hippocampal Place Cells. Cell Reports, 27(5), 1516-1527.e5. https://doi.org/10.1016/j.celrep.2019.04.002

Auer, P., Cesa-Bianchi, N., & Fischer, P. (2002). Finite-time analysis of the multiarmed bandit problem. Machine Learning, 47(2-3), 235-256. doi: 10.1023/A:1013689704352

Badre, D., Frank, M. J., & Moore, C. I. (2015). Interactionist Neuroscience. Neuron, 88(5), 855–860. https://doi.org/10.1016/j.neuron.2015.10.021

Balaguer, J., Spiers, H., Hassabis, D., & Summerfield, C. (2016). Neural Mechanisms of Hierarchical Planning in a Virtual Subway Network. Neuron, 90(4), 893–903. https://doi.org/10.1016/j.neuron.2016.03.037

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The Role of the Dorsal Striatum in Reward and Decision-Making. Journal of Neuroscience, 27(31), 8161–8165. https://doi.org/10.1523/JNEUROSCI.1554-07.2007

Balleine, B. W., Dezfouli, A., Ito, M., & Doya, K. (2015). Hierarchical control of goal-directed action in the cortical–basal ganglia network. Current Opinion in Behavioral Sciences, 5, 1–7. https://doi.org/10.1016/j.cobeha.2015.06.001

Balleine, B. W., & O'Doherty, J. P. (2010). Human and Rodent Homologies in Action Control: Corticostriatal Determinants of Goal-Directed and Habitual Action. Neuropsychopharmacology, 35(1), 48–69. https://doi.org/10.1038/npp.2009.131

Banerjee, A., Parente, G., Teutsch, J., Lewis, C., Voigt, F. F., & Helmchen, F. (2020). Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. Nature, 585(7824), 245–250. https://doi.org/10.1038/s41586-020-2704-z

Banino, A., Barry, C., Uria, B., Blundell, C., Lillicrap, T., Mirowski, P., Pritzel, A., Chadwick, M. J., Degris, T., Modayil, J., Wayne, G., Soyer, H., Viola, F., Zhang, B., Goroshin, R., Rabinowitz, N., Pascanu, R., Beattie, C., Petersen, S., … Kumaran, D. (2018). Vector-based navigation using grid-like representations in artificial agents. Nature, 557(7705), 429–433. https://doi.org/10.1038/s41586-018-0102-6

Barhorst-Cates, E. M., Meneghetti, C., Zhao, Y., Pazzaglia, F., & Creem-Regehr, S. H. (2021). Effects of home environment structure on navigation preference and performance: A comparison in Veneto, Italy and Utah, USA. Journal of Environmental Psychology, 74, 101580. https://doi.org/10.1016/j.jenvp.2021.101580

Barron, H. C., Reeve, H. M., Koolschijn, R. S., Perestenko, P. V., Shpektor, A., Nili, H., Rothaermel, R., Campo-Urriza, N., O'Reilly, J. X., Bannerman, D. M., Behrens, T. E. J., & Dupret, D. (2020). Neuronal Computation Underlying Inferential Reasoning in Humans and Mice. Cell, 183(1), 228-243.e21. https://doi.org/10.1016/j.cell.2020.08.035

Barry, C., Lever, C., Hayman, R., Hartley, T., Burton, S., O'Keefe, J., Jeffery, K., & Burgess, N. (2006). The Boundary Vector Cell Model of Place Cell Firing and Spatial Memory. Reviews in the Neurosciences, 17(1–2).

https://doi.org/10.1515/REVNEURO.2006.17.1-2.71

Barto, A. G. and Mahadevan, S. (2003), 'Recent advances in hierarchical reinforcement learning', Discrete event dynamic systems 13(1-2), 41–77.

Basu, R., Gebauer, R., Herfurth, T., Kolb, S., Golipour, Z., Tchumatchenko, T., & Ito, H. T. (2021). The orbitofrontal cortex maps future navigational goals. Nature, 599(7885), 449–452. https://doi.org/10.1038/s41586-021-04042-9

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018b). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. Neuron, 100(2), 490–509. https://doi.org/10.1016/j.neuron.2018.10.002

Bellmund, J. L. S., Gardenfors, P., Moser, E. I., & Doeller, C. F. (2018). Navigating cognition: Spatial codes for human thinking. Science, 362(6415), eaat6766. https://doi.org/10.1126/science.aat6766

Bendor, D., & Spiers, H. J. (2016). Does the Hippocampus Map Out the Future? Trends in Cognitive Sciences, 20(3), 167–169. https://doi.org/10.1016/j.tics.2016.01.003

Benhamou, S. (1996). No evidence for cognitive mapping in rats. Animal Behaviour, 52(1), 201-212.

Bennett, A. T. (1996). Do animals have cognitive maps?. The journal of experimental biology, 199(1), 219-224.

Bhattarai, B., Lee, J. W., & Jung, M. W. (2020). Distinct effects of reward and navigation history on hippocampal forward and reverse replays. Proceedings of the National Academy of Sciences, 117(1), 689–697. https://doi.org/10.1073/pnas.1912533117

Bicanski, A., & Burgess, N. (2020). Neuronal vector coding in spatial cognition. Nature Reviews Neuroscience. https://doi.org/10.1038/s41583-020-0336-9

Boccara, C. N., Nardin, M., Stella, F., O'Neill, J., & Csicsvari, J. (2019). The entorhinal cognitive map is attracted to goals. Science, 363(6434), 1443–1447. https://doi.org/10.1126/science.aav4837

Boccara, C. N., Sargolini, F., Thoresen, V. H., Solstad, T., Witter, M. P., Moser, E. I., & Moser, M.-B. (2010). Grid cells in pre- and parasubiculum. Nature Neuroscience, 13(8), 987–994. https://doi.org/10.1038/nn.2602

Boeing, G. (2017). OSMnx: A Python package to work with graph-theoretic OpenStreetMap street networks. The Journal of Open Source Software, 2(12), 215. https://doi.org/10.21105/joss.00215

Bohbot, V. D., Lerch, J., Thorndycraft, B., Iaria, G., & Zijdenbos, A. P. (2007). Gray Matter Differences Correlate with Spontaneous Strategies in a Human Virtual Navigation Task. Journal of Neuroscience, 27(38), 10078–10083. https://doi.org/10.1523/JNEUROSCI.1763-07.2007

Boorman, E. D., Rajendran, V. G., O'Reilly, J. X., & Behrens, T. E. (2016). Two Anatomically and Computationally Distinct Learning Signals Predict Changes to Stimulus-Outcome Associations in Hippocampus. Neuron, 89(6), 1343–1354. https://doi.org/10.1016/j.neuron.2016.02.014

Bostock, E., Muller, R. U., & Kubie, J. L. (1991). Experience-dependent modifications of hippocampal place cell firing. Hippocampus, 1(2), 193–205. https://doi.org/10.1002/hipo.450010207

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. Trends in Cognitive Sciences, 8(12), 539–546. https://doi.org/10.1016/j.tics.2004.10.003

Botvinick, M. M., Niv, Y., & Barto, A. G. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. Cognition, 113(3), 262–280. https://doi.org/10.1016/j.cognition.2008.08.011

Bradfield, L. A., & Hart, G. (2020). Rodent medial and lateral orbitofrontal cortices represent unique components of cognitive maps of task space. Neuroscience & Biobehavioral Reviews, 108, 287–294. https://doi.org/10.1016/j.neubiorev.2019.11.009

Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., & Zaremba, W. (2016). Openai gym. arXiv preprint arXiv:1606.01540.

Bronowski, J. (1973). The ascent of man. Random House.

Brown, T. I., Carr, V. A., LaRocque, K. F., Favila, S. E., Gordon, A. M., Bowles, B., Bailenson, J. N., & Wagner, A. D. (2016a). Prospective representation of navigational goals in the human hippocampus. Science, 352(6291), 1323–1326. https://doi.org/10.1126/science.aaf0784

Brown, T. I., Carr, V. A., LaRocque, K. F., Favila, S. E., Gordon, A. M., Bowles, B., Bailenson, J. N., & Wagner, A. D. (2016b). Prospective representation of navigational goals in the human hippocampus. Science, 352(6291), 1323–1326. https://doi.org/10.1126/science.aaf0784

Brown, T. I., Carr, V. A., LaRocque, K. F., Favila, S. E., Gordon, A. M., Bowles, B., Bailenson, J. N., & Wagner, A. D. (2016c). Prospective representation of navigational goals in the human hippocampus. Science, 352(6291), 1323–1326. https://doi.org/10.1126/science.aaf0784

Brown, T. I., Ross, R. S., Tobyne, S. M., & Stern, C. E. (2012). Cooperative interactions between hippocampal and striatal systems support flexible navigation. NeuroImage, 60(2), 1316–1330. https://doi.org/10.1016/j.neuroimage.2012.01.046

Browne, C. B., Powley, E., Whitehouse, D., Lucas, S. M., Cowling, P. I., Rohlfshagen, P., Tavener, S., Perez, D., Samothrakis, S., & Colton, S. (2012). A Survey of Monte Carlo Tree Search Methods. IEEE Transactions on Computational Intelligence and AI in Games, 4(1), 1–43. https://doi.org/10.1109/TCIAIG.2012.2186810

Brunec, I. K., Bellana, B., Ozubko, J. D., Man, V., Robin, J., Liu, Z.-X., Grady, C., Rosenbaum, R. S., Winocur, G., Barense, M. D., & Moscovitch, M. (2018). Multiple Scales of Representation along the Hippocampal Anteroposterior Axis in Humans. Current Biology, 28(13), 2129-2135.e6. https://doi.org/10.1016/j.cub.2018.05.016

Brunec, I. K., Robin, J., Patai, E. Z., Ozubko, J. D., Javadi, A.-H., Barense, M. D., Spiers, H. J., & Moscovitch, M. (2019). Cognitive mapping style relates to posterior-anterior hippocampal volume ratio. Hippocampus, 29(8), 748–754. https://doi.org/10.1002/hipo.23072

Brunton, B. W., Botvinick, M. M., & Brody, C. D. (2013). Rats and Hu-

mans Can Optimally Accumulate Evidence for Decision-Making. Science, 340(6128), 95–98. https://doi.org/10.1126/science.1233912

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. Neuron, 35(4), 625-641.

Bush, D., Barry, C., Manson, D., & Burgess, N. (2015). Using Grid Cells for Navigation. Neuron, 87(3), 507–520. https://doi.org/10.1016/j.neuron.2015.07.006

Butler, W. N., Hardcastle, K., & Giocomo, L. M. (2019). Remembered reward locations restructure entorhinal spatial maps. Science, 363(6434), 1447–1452. https://doi.org/10.1126/science.aav5297

Buzsaki, G. (2015). Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning: HIPPOCAMPAL SHARP WAVE-RIPPLE. Hippocampus, 25(10), 1073–1188. https://doi.org/10.1002/hipo.22488

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior Cingulate Cortex, Error Detection, and the Online Monitoring of Performance. Science, 280(5364), 747–749. https://doi.org/10.1126/science.280.5364.747

Chadwick, M. J., Jolly, A. E. J., Amos, D. P., Hassabis, D., & Spiers, H. J. (2015). A Goal Direction Signal in the Human Entorhinal/Subicular Region. Current Biology, 25(1), 87–92. https://doi.org/10.1016/j.cub.2014.11.001

Clark, I. A., Monk, A. M., Hotchin, V., Pizzamiglio, G., Liefgreen, A., Callaghan, M. F., & Maguire, E. A. (2020). Does hippocampal volume explain performance differences on hippocampal-dependant tasks?. NeuroImage, 221, 117211.

Constantinescu, A. O., OReilly, J. X., & Behrens, T. E. J. (2016). Organizing conceptual knowledge in humans with a gridlike code. Science, 352(6292), 1464–1468. https://doi.org/10.1126/science.aaf0941

Corrado, G. S., Sugrue, L. P., Brown, J. R., & Newsome, W. T. (2009). The Trouble with Choice: Studying Decision Variables in the Brain. In Neuroeconomics (pp. 463–480). Elsevier. https://doi.org/10.1016/B978-0-12-374176-9.00030-0

Coutrot, A., Manley, E., Yesiltepe, D., Dalton, R. C., Wiener, J. M., Holscher, C., Hornberger, M., & Spiers, H. J. (2020). Cities have a negative impact on navigation ability: Evidence from 38 countries [Preprint]. Neuroscience. https://doi.org/10.1101/2020.01.23.917211

Coutrot, A., Schmidt, S., Coutrot, L., Pittman, J., Hong, L., Wiener, J. M., Holscher, C., Dalton, R. C., Hornberger, M., & Spiers, H. J. (2019). Virtual navigation tested on a mobile app is predictive of real-world wayfinding navigation performance. PLOS ONE, 14(3), e0213272. https://doi.org/10.1371/journal.pone.0213272

Coutrot, A., Silva, R., Manley, E., de Cothi, W., Sami, S., Bohbot, V. D., Wiener, J. M., Holscher, C., Dalton, R. C., Hornberger, M., & Spiers, H. J. (2018). Global Determinants of Navigation Ability. Current Biology, 28(17), 2861-2866.e4. https://doi.org/10.1016/j.cub.2018.06.009

Cox, J., & Witten, I. B. (2019). Striatal circuits for reward learning and decision-making. Nature Reviews Neuroscience, 20(8), 482–494. https://doi.org/10.1038/s41583-019-0189-2

Dahmani, L., & Bohbot, V. D. (2015). Dissociable contributions of the prefrontal cortex to hippocampus- and caudate nucleus-dependent virtual navigation strategies. Neurobiology of Learning and Memory, 117, 42–50. https://doi.org/10.1016/j.nlm.2014.07.002

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. Neuron, 69(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nature Neuroscience, 8(12), 1704–1711. https://doi.org/10.1038/nn1560

Dayan, P. (1993). Improving Generalization for Temporal Difference Learning: The Successor Representation. Neural Computation, 5(4), 613–624. https://doi.org/10.1162/neco.1993.5.4.613

Dayan, P., & Abbott, L. F. (n.d.). THEORETICAL NEUROSCIENCE. 432.

Dayan, P., & Niv, Y. (2008). Reinforcement learning: The Good, The Bad and The Ugly. Current Opinion in Neurobiology, 18(2), 185–196. https://doi.org/10.1016/j.conb.2008.08.003

de Cothi, W., Nyberg, N., Griesbauer, E.-M., Ghanamé, C., Lefort, J., Fletcher, L., Newton, C., Renaudineau, S., Bendor, D., Grieves, R., Duvelle, É., & Barry, C. (n.d.). Predictive Maps in Rats and Humans for Spatial Navigation. 27.

Devan, B. D., & White, N. M. (1999). Parallel Information Processing in the Dorsal Striatum: Relation to Hippocampal Function. The Journal of Neuroscience, 19(7), 2789–2798. https://doi.org/10.1523/JNEUROSCI.19-07-02789.1999

Dickinson, A. (1985). Actions and Habits: The Development of Behavioural Autonomy. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 308(1135,), 67–78.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. Numerische mathematik, 1(1), 269-271.

Doeller, C. F., King, J. A., & Burgess, N. (2008). Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. Proceedings of the National Academy of Sciences, 105(15), 5915–5920. https://doi.org/10.1073/pnas.0801489105

Dolan, R. J., & Dayan, P. (2013). Goals and Habits in the Brain. Neuron, 80(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. Current Opinion in Neurobiology, 22(6), 1075–1081. https://doi.org/10.1016/j.conb.2012.08.003

Dragoi, G., & Tonegawa, S. (2011). Preplay of future place cell sequences by hippocampal cellular assemblies. Nature, 469(7330), 397–401. https://doi.org/10.1038/nature09633

Dragoi, G., & Tonegawa, S. (2014). Selection of preconfigured cell assemblies for representation of novel spatial experiences. Philosophical Trans-

actions of the Royal Society B: Biological Sciences, 369(1635), 20120522. https://doi.org/10.1098/rstb.2012.0522

Dupret, D., O'Neill, J., Pleydell-Bouverie, B., & Csicsvari, J. (2010). The reorganization and reactivation of hippocampal maps predict spatial memory performance. Nature Neuroscience, 13(8), 995–1002. https://doi.org/10.1038/nn.2599

Duvelle, É., Grieves, R. M., Liu, A., Jedidi-Ayoub, S., Holeniewska, J., Harris, A., Nyberg, N., Donnarumma, F., Lefort, J. M., Jeffery, K. J., Summerfield, C., Pezzulo, G., & Spiers, H. J. (2020). Hippocampal place cells encode global location but not changes in environmental connectivity in a 4-room navigation task. BioRxiv, 2020.10.20.346130. https://doi.org/10.1101/2020.10.20.346130

Eichenbaum, H. (2017). On the integration of space, time, and memory. Neuron, 95(5), 1007-1018.

Eckstein, M. K., Wilbrecht, L., & Collins, A. G. (2021). What do reinforcement learning models measure? Interpreting model parameters in cognition and neuroscience. Current Opinion in Behavioral Sciences, 41, 128-137.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proceedings of the national academy of sciences, 113(28), 7900-7905.

Ekstrom, A. D. (2015). Why vision is important to how we navigate. Hippocampus, 25(6), 731-735.

Ekstrom, A. D. (2021). Regional variation in neurovascular coupling and why we still lack a Rosetta Stone. Philosophical Transactions of the Royal Society B, 376(1815), 20190634.

Ekstrom, A. D., Harootonian, S. K., & Huffman, D. J. (2020). Grid coding, spatial representation, and navigation: Should we assume an isomorphism? Hippocampus, 30(4), 422–432. https://doi.org/10.1002/hipo.23175

Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., & Fried, I. (2003). Cellular networks underlying human

spatial navigation. Nature, 425(6954), 184-188.

Ekstrom, A. D., Huffman, D. J., & Starrett, M. (2017). Interacting networks of brain regions underlie human spatial navigation: a review and novel synthesis of the literature. Journal of neurophysiology, 118(6), 3328-3344.

Epstein, R. A., Patai, E. Z., Julian, J. B., & Spiers, H. J. (2017). The cognitive map in humans: Spatial navigation and beyond. Nature Neuroscience, 20(11), 1504–1513. https://doi.org/10.1038/nn.4656

Esteban, O., Ciric, R., Finc, K., Blair, R., Markiewicz, C. J., Moodie, C. A., Kent, J. D., Goncalves, M., DuPre, E., Gomez, D. E. P., Ye, Z., Salo, T., Valabregue, R., Amlien, I. K., Liem, F., Jacoby, N., Stojić, H., Cieslak, M., Urchs, S., ... Gorgolewski, K. J. (2019). Analysis of task-based functional MRI data preprocessed with fMRIPrep [Preprint]. Bioinformatics. https://doi.org/10.1101/694364

Fassihi, A., Akrami, A., Esmaeili, V., & Diamond, M. E. (2014). Tactile perception and working memory in rats and humans. Proceedings of the National Academy of Sciences, 111(6), 2331–2336. https://doi.org/10.1073/pnas.1315171111

Ferbinteanu, J. (2019). Memory systems 2018 – Towards a new paradigm. Neurobiology of Learning and Memory, 157, 61–78. https://doi.org/10.1016/j.nlm.2018.11.005

Foster, D. J. (2017). Replay Comes of Age. Annual Review of Neuroscience, 40(1), 581–602. https://doi.org/10.1146/annurev-neuro-072116-031538

Foster, D. J., Morris, R. G. M., & Dayan, P. (n.d.-a). A model of hippocampally dependent navigation, using the temporal difference learning rule. 16.

Foster, D. J., Morris, R. G. M., & Dayan, P. (n.d.-b). A Model of Hippocampally Dependent Navigation, Using the Temporal Difference Learning Rule. 17.

Foster, D. J., & Wilson, M. A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature, 440(7084), 680–683. https://doi.org/10.1038/nature04587

Foster, D. J., & Wilson, M. A. (2007). Hippocampal theta sequences. Hip-

pocampus, 17(11), 1093–1099. https://doi.org/10.1002/hipo.20345

Fyhn, M., Hafting, T., Treves, A., Moser, M.-B., & Moser, E. I. (2007). Hippocampal remapping and grid realignment in entorhinal cortex. Nature, 446(7132), 190–194. https://doi.org/10.1038/nature05601

Fyhn, M., Hafting, T., Witter, M. P., Moser, E. I., & Moser, M.-B. (2008). Grid cells in mice. Hippocampus, 18(12), 1230–1238. https://doi.org/10.1002/hipo.20472

Gahnstrom, C. J., & Spiers, H. J. (2020). Striatal and hippocampal contributions to flexible navigation in rats and humans. Brain and Neuroscience Advances, 4, 239821282097977. https://doi.org/10.1177/2398212820979772

Gallistel, C. R. and King, A. P. (2011), Memory and the computational brain: Why cognitive science will transform neuroscience, Vol. 6, John Wiley & Sons.

Gardner, M. P. H., & Schoenbaum, G. (2021). The orbitofrontal cartographer. Behavioral Neuroscience, 135(2), 267–276. https://doi.org/10.1037/bne0000463

Gasser, J., Pereira de Vasconcelos, A., Cosquer, B., Boutillier, A.-L., & Cassel, J.-C. (2020). Shifting between response and place strategies in maze navigation: Effects of training, cue availability and functional inactivation of striatum or hippocampus in rats. Neurobiology of Learning and Memory, 167, 107131. https://doi.org/10.1016/j.nlm.2019.107131

Gauthier, J. L., & Tank, D. W. (2018). A Dedicated Population for Reward Coding in the Hippocampus. Neuron, 99(1), 179-193.e7. https://doi.org/10.1016/j.neuron.2018.06.008

Geerts, J. P., Chersi, F., Stachenfeld, K. L., & Burgess, N. (n.d.). A general model of hippocampal and dorsal striatal learning and decision making. COMPUTATIONAL BIOLOGY, 11.

Gershman, S. J. (2018). The Successor Representation: Its Computational Logic and Neural Substrates. The Journal of Neuroscience, 38(33), 7193–7200. https://doi.org/10.1523/JNEUROSCI.0151-18.2018

Gershman, S. J., & Daw, N. D. (2017). Reinforcement Learning and Episodic

Memory in Humans and Animals: An Integrative Framework. Annual Review of Psychology, 68(1), 101–128. https://doi.org/10.1146/annurev-psych-122414-033625

Gerstner, W., Sprekeler, H., & Deco, G. (2012). Theory and Simulation in Neuroscience. Science, 338(6103), 60–65. https://doi.org/10.1126/science.1227356

Glascher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning. Neuron, 66(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016

Goodroe, S. C., Starnes, J., & Brown, T. I. (2018). The Complex Nature of Hippocampal-Striatal Interactions in Spatial Navigation. Frontiers in Human Neuroscience, 12, 250. https://doi.org/10.3389/fnhum.2018.00250

Grieves, R. M., & Jeffery, K. J. (2017). The representation of space in the brain. Behavioural Processes, 135, 113–131. https://doi.org/10.1016/j.beproc.2016.12.012

Grieves, R. M., & Dudchenko, P. A. (2013). Cognitive maps and spatial inference in animals: Rats fail to take a novel shortcut, but can take a previously experienced one. Learning and Motivation, 44(2), 81-92.

Gupta, A. S., van der Meer, M. A. A., Touretzky, D. S., & Redish, A. D. (2010). Hippocampal Replay Is Not a Simple Function of Experience. Neuron, 65(5), 695–705. https://doi.org/10.1016/j.neuron.2010.01.034

Haber, S. N., Fudge, J. L., & McFarland, N. R. (2000). Striatonigrostriatal Pathways in Primates Form an Ascending Spiral from the Shell to the Dorsolateral Striatum. The Journal of Neuroscience, 20(6), 2369–2382. https://doi.org/10.1523/JNEUROSCI.20-06-02369.2000

Haber, S. N., & Knutson, B. (2010). The Reward Circuit: Linking Primate Anatomy and Human Imaging. Neuropsychopharmacology, 35(1), 4–26. https://doi.org/10.1038/npp.2009.129

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. Nature, 436(7052), 801–806. https://doi.org/10.1038/nature03721

Haklay, M., & Weber, P. (2008). OpenStreetMap: User-Generated Street Maps. IEEE Pervasive Computing, 7(4), 12–18. https://doi.org/10.1109/MPRV.2008.80

Hales, J. B., Schlesiger, M. I., Leutgeb, J. K., Squire, L. R., Leutgeb, S., & Clark, R. E. (2014). Medial Entorhinal Cortex Lesions Only Partially Disrupt Hippocampal Place Cells and Hippocampus-Dependent Place Memory. Cell Reports, 9(3), 893–901. https://doi.org/10.1016/j.celrep.2014.10.009

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the Role of the Orbitofrontal Cortex and the Striatum in the Computation of Goal Values and Prediction Errors. Journal of Neuroscience, 28(22), 5623–5630. https://doi.org/10.1523/JNEUROSCI.1309-08.2008

Harten, L., Katz, A., Goldshtein, A., Handel, M., & Yovel, Y. (2020). The ontogeny of a mammalian cognitive map in the real world. Science, 369(6500), 194-197.

Hartley, T., Maguire, E. A., Spiers, H. J., & Burgess, N. (n.d.). The Well-Worn Route and the Path Less Traveled: Distinct Neural Bases of Route Following and Wayfinding in Humans. 12.

Harvey, C. D., Collman, F., Dombeck, D. A., & Tank, D. W. (2009). Intracellular dynamics of hippocampal place cells during virtual navigation. Nature, 461(7266), 941–946. https://doi.org/10.1038/nature08499

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. Proceedings of the National Academy of Sciences, 104(5), 1726-1731.

Hok, V., Lenck-Santini, P.-P., Roux, S., Save, E., Muller, R. U., & Poucet, B. (2007). Goal-Related Activity in Hippocampal Place Cells. Journal of Neuroscience, 27(3), 472–482. https://doi.org/10.1523/JNEUROSCI.2864-06.2007

Hollup, S. A., Molden, S., Donnett, J. G., Moser, M.-B., & Moser, E. I. (2001). Accumulation of Hippocampal Place Fields at the Goal Location in an Annular Watermaze Task. The Journal of Neuroscience, 21(5), 1635–1644.

https://doi.org/10.1523/JNEUROSCI.21-05-01635.2001

Horner, A. J., Bisby, J. A., Zotow, E., Bush, D., & Burgess, N. (2016). Grid-like Processing of Imagined Navigation. Current Biology, 26(6), 842–847. https://doi.org/10.1016/j.cub.2016.01.042

Howard, L. R., Javadi, A. H., Yu, Y., Mill, R. D., Morrison, L. C., Knight, R., Loftus, M. M., Staskute, L., & Spiers, H. J. (2014). The Hippocampus and Entorhinal Cortex Encode the Path and Euclidean Distances to Goals during Navigation. Current Biology, 24(12), 1331–1340. https://doi.org/10.1016/j.cub.2014.05.001

Hoydal, O. A., Skytoen, E. R., Andersson, S. O., Moser, M.-B., & Moser, E. I. (2019). Object-vector coding in the medial entorhinal cortex. Nature, 568(7752), 400–404. https://doi.org/10.1038/s41586-019-1077-7

Huang, Y., Yaple, Z. A., & Yu, R. (2020). Goal-oriented and habitual decisions: Neural signatures of model-based and model-free learning. NeuroImage, 215, 116834. https://doi.org/10.1016/j.neuroimage.2020.116834

Huffman, D. J., & Ekstrom, A. D. (2019). A Modality-Independent Network Underlies the Retrieval of Large-Scale Spatial Environments in the Human Brain. Neuron, 104(3), 611-622.e7. https://doi.org/10.1016/j.neuron.2019.08.012

Huffman, D. J., & Ekstrom, A. D. (2019). Which way is the bookstore? A closer look at the judgments of relative directions task. Spatial Cognition & Computation, 19(2), 93-129.

Iaria, G., Petrides, M., Dagher, A., Pike, B., & Bohbot, V. D. (n.d.). Cognitive Strategies Dependent on the Hippocampus and Caudate Nucleus in Human Navigation: Variability and Change with Practice. 8.

Igata, H., Ikegaya, Y., & Sasaki, T. (2021). Prioritized experience replays on a hippocampal predictive map for learning. Proceedings of the National Academy of Sciences, 118(1), e2011266118. https://doi.org/10.1073/pnas.2011266118

Igloi, K., Zaoui, M., Berthoz, A., & Rondi-Reig, L. (2009). Sequential egocentric strategy is acquired as early as allocentric strategy: Parallel acquisi-

tion of these two navigation strategies. Hippocampus, 19(12), 1199–1211. https://doi.org/10.1002/hipo.20595

Javadi, A.-H., Emo, B., Howard, L. R., Zisch, F. E., Yu, Y., Knight, R., Pinelo Silva, J., & Spiers, H. J. (2017). Hippocampal and prefrontal processing of network topology to simulate the future. Nature Communications, 8, 14652. https://doi.org/10.1038/ncomms14652

Javadi, A.-H., Patai, E. Z., Marin-Garcia, E., Margois, A., Tan, H.-R. M., Kumaran, D., Nardini, M., Penny, W., Duzel, E., Dayan, P., & Spiers, H. J. (2019a). Backtracking during navigation is correlated with enhanced anterior cingulate activity and suppression of alpha oscillations and the 'default-mode' network. Proceedings of the Royal Society B: Biological Sciences, 286(1908), 20191016. https://doi.org/10.1098/rspb.2019.1016

Javadi, A.-H., Patai, E. Z., Marin-Garcia, E., Margois, A., Tan, H.-R. M., Kumaran, D., Nardini, M., Penny, W., Duzel, E., Dayan, P., & Spiers, H. J. (2019b). Backtracking during navigation is correlated with enhanced anterior cingulate activity and suppression of alpha oscillations and the 'default-mode' network. Proceedings of the Royal Society B: Biological Sciences, 286(1908), 20191016. https://doi.org/10.1098/rspb.2019.1016

Javadi, A.-H., Patai, E. Z., Marin-Garcia, E., Margolis, A., Tan, H.-R. M., Kumaran, D., Nardini, M., Penny, W., Duzel, E., Dayan, P., & Spiers, H. J. (2019). Prefrontal Dynamics Associated with Efficient Detours and Shortcuts: A Combined Functional Magnetic Resonance Imaging and Magnetoencenphalography Study. Journal of Cognitive Neuroscience, 31(8), 1227–1247. https://doi.org/10.1162/jocn_a_01414

Johnson, A., van der Meer, M. A., & Redish, A. D. (2007). Integrating hippocampus and striatum in decision-making. Current Opinion in Neurobiology, 17(6), 692–697. https://doi.org/10.1016/j.conb.2008.01.003

Juechems, K., & Summerfield, C. (2019). Where Does Value Come From? Trends in Cognitive Sciences, 23(10), 836–850. https://doi.org/10.1016/j.tics.2019.07.012

Julian, J. B., Keinath, A. T., Frazzetta, G., & Epstein, R. A. (2018). Human

entorhinal cortex represents visual space using a boundary-anchored grid. Nature Neuroscience, 21(2), 191–194. https://doi.org/10.1038/s41593-017-0049-1

Jung, M. W., Lee, H., Jeong, Y., Lee, J. W., & Lee, I. (2018). Remembering rewarding futures: A simulation-selection model of the hippocampus. Hippocampus, 28(12), 913–930. https://doi.org/10.1002/hipo.23023

Kaplan, R., King, J., Koster, R., Penny, W. D., Burgess, N., & Friston, K. J. (2017). The Neural Representation of Prospective Choice during Spatial Planning and Decisions. PLOS Biology, 15(1), e1002588. https://doi.org/10.1371/journal.pbio.1002588

Kaplan, R., Tauste Campo, A., Bush, D., King, J., Principe, A., Koster, R., ... & Friston, K. J. (2020). Human hippocampal theta oscillations reflect sequential dependencies during spatial planning. Cognitive Neuroscience, 11(3), 122-131.

Keiflin, R., Pribut, H. J., Shah, N. B., & Janak, P. H. (2019). Ventral Tegmental Dopamine Neurons Participate in Reward Identity Predictions. Current Biology, 29(1), 93-103.e3. https://doi.org/10.1016/j.cub.2018.11.050

Keller, J., Strasburger, H., Cerutti, D. T., & Sabel, B. A. (2000). Assessing spatial vision—Automated measurement of the contrast-sensitivity function in the hooded rat. Journal of Neuroscience Methods, 97(2), 103–110. https://doi.org/10.1016/S0165-0270(00)00173-4

Killian, N. J., Jutras, M. J., & Buffalo, E. A. (2012). A map of visual space in the primate entorhinal cortex. Nature, 491(7426), 761–764. https://doi.org/10.1038/nature11587

Kimchi, E. Y., Torregrossa, M. M., Taylor, J. R., & Laubach, M. (2009). Neuronal Correlates of Instrumental Learning in the Dorsal Striatum. Journal of Neurophysiology, 102(1), 475–489. https://doi.org/10.1152/jn.00262.2009

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A Neostriatal Habit Learning System in Humans. Science, 273(5280), 1399–1402. https://doi.org/10.1126/science.273.5280.1399

Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural Mechanisms of Foraging. Science, 336(6077), 95–98. https://doi.org/10.1126/science.1216930

Komorowski, R. W., Manns, J. R., & Eichenbaum, H. (2009). Robust Conjunctive Item-Place Coding by Hippocampal Neurons Parallels Learning What Happens Where. Journal of Neuroscience, 29(31), 9918–9929. https://doi.org/10.1523/JNEUROSCI.1378-09.2009

Kriegeskorte, N., & Douglas, P. K. (2018). Cognitive computational neuroscience. Nature Neuroscience, 21(9), 1148–1160. https://doi.org/10.1038/s41593-018-0210-5

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. Frontiers in systems neuroscience, 2, 4.

Kropff, E., Carmichael, J. E., Moser, M.-B., & Moser, E. I. (2015). Speed cells in the medial entorhinal cortex. Nature, 523(7561), 419–424. https://doi.org/10.1038/nature14622

Krupic, J., Bauza, M., Burton, S., Barry, C., & O'Keefe, J. (2015). Grid cell symmetry is shaped by environmental geometry. Nature, 518(7538), 232–235. https://doi.org/10.1038/nature14153

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. The annals of mathematical statistics, 22(1), 79-86.

Kunz, L., Brandt, A., Reinacher, P. C., Staresina, B. P., Reifenstein, E. T., Weidemann, C. T., ... & Jacobs, J. (2021). A neural code for egocentric spatial maps in the human medial temporal lobe. Neuron, 109(17), 2781-2796.

Kurth-Nelson, Z., Economides, M., Dolan, R. J., & Dayan, P. (2016). Fast Sequences of Non-spatial State Representations in Humans. Neuron, 91(1), 194–204. https://doi.org/10.1016/j.neuron.2016.05.028

Lak, A., Costa, G. M., Romberg, E., Koulakov, A. A., Mainen, Z. F., & Kepecs, A. (2014). Orbitofrontal Cortex Is Required for Opti-

mal Waiting Based on Decision Confidence. Neuron, 84(1), 190–201. https://doi.org/10.1016/j.neuron.2014.08.039

Langdon, A. J., Sharpe, M. J., Schoenbaum, G., & Niv, Y. (2018). Model-based predictions for dopamine. Current Opinion in Neurobiology, 49, 1–7. https://doi.org/10.1016/j.conb.2017.10.006

Lawton, C. A. (2001). Gender and Regional Differences in Spatial Referents Used in Direction Giving. Sex Roles, 17.

Lever, C., Burton, S., Jeewajee, A., O'Keefe, J., & Burgess, N. (2009). Boundary Vector Cells in the Subiculum of the Hippocampal Formation. Journal of Neuroscience, 29(31), 9771–9777. https://doi.org/10.1523/JNEUROSCI.1319-09.2009

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. nature, 412(6843), 150-157.

Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the BOLD signal. Annu. Rev. Physiol., 66, 735-769.

Ludvig, N., Tang, H. M., Gohil, B. C., & Botero, J. M. (2004). Detecting location-specific neuronal firing rate increases in the hippocampus of freely-moving monkeys. Brain research, 1014(1-2), 97-109.

Maguire, E. A., Frackowiak, R. S. J., & Frith, C. D. (1997). Recalling Routes around London: Activation of the Right Hippocampus in Taxi Drivers. The Journal of Neuroscience, 17(18), 7103–7110. https://doi.org/10.1523/JNEUROSCI.17-18-07103.1997

Maguire, E. A., Nannery, R., & Spiers, H. J. (2006). Navigation around London by a taxi driver with bilateral hippocampal lesions. Brain, 129(11), 2894-2907.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. Proceedings of the National Academy of Sciences, 97(8), 4398–4403.

https://doi.org/10.1073/pnas.070039597

Maguire, E. A., Spiers, H. J., Good, C. D., Hartley, T., Frackowiak, R. S. J., & Burgess, N. (2003). Navigation expertise and the human hippocampus: A structural brain imaging analysis. Hippocampus, 13(2), 250–259. https://doi.org/10.1002/hipo.10087

Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: A structural MRI and neuropsychological analysis. Hippocampus, 16(12), 1091–1101. https://doi.org/10.1002/hipo.20233

Mao, D., Avila, E., Caziot, B., Laurens, J., Dickman, J. D., & Angelaki, D. E. (2021). Spatial modulation of hippocampal activity in freely moving macaques. Neuron, 109(21), 3521-3534.

Marchette, S. A., Bakker, A., & Shelton, A. L. (2011). Cognitive Mappers to Creatures of Habit: Differential Engagement of Place and Response Learning Mechanisms Predicts Human Navigational Behavior. Journal of Neuroscience, 31(43), 15264–15268. https://doi.org/10.1523/JNEUROSCI.3634-11.2011

Marr, D. (1982), 'Vision: A computational investigation into the human representation and processing of visual information. mit press', Cambridge, Massachusetts.

Mattar, M. G., & Daw, N. D. (2017). Prioritized memory access explains planning and hippocampal replay [Preprint]. Neuroscience. https://doi.org/10.1101/225664

McDonald, R. J., & White, N. M. (1994). Parallel information processing in the water maze: Evidence for independent memory systems involving dorsal striatum and hippocampus. Behavioral and Neural Biology, 61(3), 260–270. https://doi.org/10.1016/S0163-1047(05)80009-3

McNamara, C. G., Tejero-Cantero, Á., Trouche, S., Campo-Urriza, N., & Dupret, D. (2014). Dopaminergic neurons promote hippocampal reactivation and spatial memory persistence. Nature Neuroscience, 17(12), 1658–1660. https://doi.org/10.1038/nn.3843

McNamee, D., Wolpert, D. M., & Lengyel, M. (n.d.). Efficient state-space modularization for planning: Theory, behavioral and neural signatures. 9.

McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., & Moser, M.-B. (2006). Path integration and the neural basis of the 'cognitive map'. Nature Reviews Neuroscience, 7(8), 663–678. https://doi.org/10.1038/nrn1932

McNaughton, B. L., Chen, L. L., & Markus, E. J. (1991). "Dead Reckoning," Landmark Learning, and the Sense of Direction: A Neurophysiological and Computational Hypothesis. Journal of Cognitive Neuroscience, 3(2), 190–202. https://doi.org/10.1162/jocn.1991.3.2.190

Meier, P. M., & Reinagel, P. (2013). Rats and humans differ in processing collinear visual features. Frontiers in Neural Circuits, 7. https://doi.org/10.3389/fncir.2013.00197

Miller, K. J., Botvinick, M. M., & Brody, C. D. (2017). Dorsal hippocampus contributes to model-based planning. Nature Neuroscience, 20(9), 1269–1276. https://doi.org/10.1038/nn.4613

Miller, K. J., & Venditto, S. J. C. (2021). Multi-step planning in the brain. Current Opinion in Behavioral Sciences, 38, 29–39. https://doi.org/10.1016/j.cobeha.2020.07.003

Mittelstaedt, M.-L., & Mittelstaedt, H. (1980). Homing by path integration in a mammal. Naturwissenschaften, 67(11), 566–567. https://doi.org/10.1007/BF00450672

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. nature, 518(7540), 529-533.

Mok, R. M., & Love, B. C. (2018). A non-spatial account of place and grid cells based on clustering models of concept learning [Preprint]. Neuroscience. https://doi.org/10.1101/421842

Momennejad, I. (2020). Learning Structures: Predictive Representations, Replay, and Generalization. Current Opinion in Behavioral Sciences, 32, 155–166. https://doi.org/10.1016/j.cobeha.2020.02.017

Momennejad, I., Otto, A. R., Daw, N. D., & Norman, K. A. (2018). Offline replay supports planning in human reinforcement learning. ELife, 7, e32548.

https://doi.org/10.7554/eLife.32548

Momennejad, I., Russek, E. M., Cheong, J. H., Botvinick, M. M., Daw, N. D., & Gershman, S. J. (2017). The successor representation in human reinforcement learning. Nature Human Behaviour, 1(9), 680–692. https://doi.org/10.1038/s41562-017-0180-8

Morris, R. G. M., Garrud, P., Rawlins, J. N. P., & O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. Nature, 297(5868), 681–683. https://doi.org/10.1038/297681a0

Morris, R. (1984). Developments of a water-maze procedure for studying spatial learning in the rat. Journal of neuroscience methods, 11(1), 47-60.

Moser, E. I., Moser, M.-B., & McNaughton, B. L. (2017). Spatial representation in the hippocampal formation: A history. Nature Neuroscience, 20(11), 1448–1464. https://doi.org/10.1038/nn.4653

Moser, M.-Brit. (1995). Spatial learning with a minislab in the dorsal hippocampus. Proc. Natl. Acad. Sci. USA, 5.

Newcombe, N. S. (2018). Individual variation in human navigation. Current Biology, 28(17), R1004–R1008. https://doi.org/10.1016/j.cub.2018.04.053

Newman, E. L., Caplan, J. B., Kirschen, M. P., Korolev, I. O., Sekuler, R., & Kahana, M. J. (2007). Learning your way around town: How virtual taxicab drivers learn to use both layout and landmark information. Cognition, 104(2), 231-253.

Niv, Y. (n.d.). Reinforcement learning in the brain. 38.

Noonan, M. P., Walton, M. E., Behrens, T. E. J., Sallet, J., Buckley, M. J., & Rushworth, M. F. S. (2010). Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. Proceedings of the National Academy of Sciences, 107(47), 20547–20552. https://doi.org/10.1073/pnas.1012246107

O'Doherty, J. (2004). Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. Science, 304(5669), 452–454. https://doi.org/10.1126/science.1094285

O'Keefe, J., & Conway, D. H. (1978). Hippocampal place units in the freely moving rat: Why they fire where they fire. Experimental Brain Research, 31(4). https://doi.org/10.1007/BF00239813

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Research, 34(1), 171–175. https://doi.org/10.1016/0006-8993(71)90358-1

O'Keefe, J., & Nadel, L. (1978). The hippocampus as a cognitive map. Clarendon Press; Oxford University Press.

Ólafsdóttir, H. F., Barry, C., Saleem, A. B., Hassabis, D., & Spiers, H. J. (2015). Hippocampal place cells construct reward related sequences through unexplored space. ELife, 4, e06063. https://doi.org/10.7554/eLife.06063

Olthof, A., Sutton, J. E., Slumskie, S. V., D'Addetta, J., & Roberts, W. A. (1999). In search of the cognitive map: Can rats learn an abstract pattern of rewarded arms on the radial maze?. Journal of Experimental Psychology: Animal Behavior Processes, 25(3), 352.

O'Reilly, J. X., Schuffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. Proceedings of the National Academy of Sciences, 110(38), E3660–E3669. https://doi.org/10.1073/pnas.1305373110

Packard, M. G., & McGaugh, J. L. (1996). Inactivation of Hippocampus or Caudate Nucleus with Lidocaine Differentially Affects Expression of Place and Response Learning. Neurobiology of Learning and Memory, 65(1), 65–72. https://doi.org/10.1006/nlme.1996.0007

Packard, M., Hirsh, R., & White, N. (1989). Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: Evidence for multiple memory systems. The Journal of Neuroscience, 9(5), 1465–1472. https://doi.org/10.1523/JNEUROSCI.09-05-01465.1989

Padoa-Schioppa, C., & Conen, K. E. (2017). Orbitofrontal Cortex: A Neural Circuit for Economic Decisions. Neuron, 96(4), 736–754.

https://doi.org/10.1016/j.neuron.2017.09.031

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. Journal of Behavioral and Experimental Finance, 17, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Palminteri, S., Wyart, V., & Koechlin, E. (2017). The Importance of Falsification in Computational Cognitive Modeling. Trends in Cognitive Sciences, 21(6), 425–433. https://doi.org/10.1016/j.tics.2017.03.011

Parker, N. F., Cameron, C. M., Taliaferro, J. P., Lee, J., Choi, J. Y., Davidson, T. J., Daw, N. D., & Witten, I. B. (2016). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. Nature Neuroscience, 19(6), 845–854. https://doi.org/10.1038/nn.4287

Patai, E. Z., Javadi, A.-H., Ozubko, J. D., O'Callaghan, A., Ji, S., Robin, J., Grady, C., Winocur, G., Rosenbaum, R. S., Moscovitch, M., & Spiers, H. J. (2019). Hippocampal and Retrosplenial Goal Distance Coding After Long-term Consolidation of a Real-World Environment. Cerebral Cortex, 29(6), 2748–2758. https://doi.org/10.1093/cercor/bhz044

Patai, E. Z., Javadi, A.-H., Ozubko, J. D., O'Callaghan, A., Ji, S., Robin, J., Grady, C., Winocur, G., Rosenbaum, S. R., Moscovitch, M., & Spiers, H. J. (2017). Long-term consolidation switches goal proximity coding from hippocampus to retrosplenial cortex [Preprint]. Neuroscience. https://doi.org/10.1101/167882

Patai, E. Z., & Spiers, H. J. (2021). The Versatile Wayfinder: Prefrontal Contributions to Spatial Navigation. Trends in Cognitive Sciences, 25(6), 520–533. https://doi.org/10.1016/j.tics.2021.02.010

Pearce, J. M., Roberts, A. D. L., & Good, M. (1998). Hippocampal lesions disrupt navigation based on cognitive maps but not heading vectors. Nature, 396(6706), 75–77. https://doi.org/10.1038/23941

Peer, M., Brunec, I. K., Newcombe, N. S., & Epstein, R. A. (2020). Structuring Knowledge with Cognitive Maps and Cognitive Graphs. Trends in Cognitive Sciences, S1364661320302503. https://doi.org/10.1016/j.tics.2020.10.004

Peer, M., & Epstein, R. A. (2021). The human brain uses spatial schemas to represent segmented environments. Current Biology, 31(21), 4677-4688.e8. https://doi.org/10.1016/j.cub.2021.08.012

Pennartz, C. M. A. (2004). The Ventral Striatum in Off-Line Processing: Ensemble Reactivation during Sleep and Modulation by Hippocampal Ripples. Journal of Neuroscience, 24(29), 6446–6456. https://doi.org/10.1523/JNEUROSCI.0575-04.2004

Pezzulo, G., van der Meer, M. A. A., Lansink, C. S., & Pennartz, C. M. A. (2014). Internally generated sequences in learning and executing goal-directed behavior. Trends in Cognitive Sciences, 18(12), 647–657. https://doi.org/10.1016/j.tics.2014.06.011

Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. Nature, 497(7447), 74–79. https://doi.org/10.1038/nature12112

Piray, P., & Daw, N. D. (2019). Linear reinforcement learning: Flexible reuse of computation in planning, grid fields, and cognitive control [Preprint]. Neuroscience. https://doi.org/10.1101/856849

Poldrack, R. A., Mumford, J. A. and Nichols, T. E. (2011), Handbook of functional MRI data analysis, Cambridge University Press.

Porter, B. S., Hillman, K. L., & Bilkey, D. K. (2019). Anterior cingulate cortex encoding of effortful behavior. Journal of Neurophysiology, 121(2), 701–714. https://doi.org/10.1152/jn.00654.2018

Prados, J., Chamizo, V. D., & Mackintosh, N. J. (1999). Latent inhibition and perceptual learning in a swimming-pool navigation task. Journal of Experimental Psychology: Animal Behavior Processes, 25(1), 37.

Ragozzino, M. E., Ragozzino, K. E., Mizumori, S. J. Y., & Kesner, R. P. (2002). Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. Behavioral Neuroscience, 116(1), 105–115. https://doi.org/10.1037/0735-7044.116.1.105

Ranck Jr, J. (1984), 'Head direction cells in the deep layer of dorsal presubicu-

lum in freely moving rats', 10, 599.

Redish, A. D. (2016). Vicarious trial and error. Nature Reviews Neuroscience, 17(3), 147–159. https://doi.org/10.1038/nrn.2015.30

Regier, P. S., Amemiya, S., & Redish, A. D. (2015). Hippocampus and subregions of the dorsal striatum respond differently to a behavioral strategy change on a spatial navigation task. Journal of Neurophysiology, 114(3), 1399–1416. https://doi.org/10.1152/jn.00189.2015

Rich, E. L., & Wallis, J. D. (2016). Decoding subjective decisions from orbitofrontal cortex. Nature Neuroscience, 19(7), 973–980. https://doi.org/10.1038/nn.4320

Rinaldi, A., De Leonibus, E., Cifra, A., Torromino, G., Minicocci, E., De Sanctis, E., López-Pedrajas, R. M., Oliverio, A., & Mele, A. (2020). Flexible use of allocentric and egocentric spatial memories activates differential neural networks in mice. Scientific Reports, 10(1), 11338. https://doi.org/10.1038/s41598-020-68025-y

Rolls, E. T. (n.d.). Spatial View Cells and the Representation of Place in the Primate Hippocampus. 14.

Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J., & Daw, N. D. (2017). Predictive representations can link model-based reinforcement learning to model-free mechanisms. PLOS Computational Biology, 13(9), e1005768. https://doi.org/10.1371/journal.pcbi.1005768

Russell, S., & Norvig, P. (2010). Artificial intelligence: a modern approach.

Rusu, S. I., & Pennartz, C. M. A. (2020). Learning, memory and consolidation mechanisms for behavioral control in hierarchically organized cortico-basal ganglia systems. Hippocampus, 30(1), 73–98. https://doi.org/10.1002/hipo.23167

Santoro, A., Frankland, P. W., & Richards, B. A. (2016). Memory Transformation Enhances Reinforcement Learning in Dynamic Environments. Journal of Neuroscience, 36(48), 12228–12242. https://doi.org/10.1523/JNEUROSCI.0763-16.2016

Sarel, A., Finkelstein, A., Las, L., & Ulanovsky, N. (2017). Vectorial representation of spatial goals in the hippocampus of bats. Science, 355(6321), 176–180. https://doi.org/10.1126/science.aak9589

Save, E., & Sargolini, F. (2017). Disentangling the Role of the MEC and LEC in the Processing of Spatial and Non-Spatial Information: Contribution of Lesion Studies. Frontiers in Systems Neuroscience, 11, 81. https://doi.org/10.3389/fnsys.2017.00081

Savelli, F., Yoganarasimha, D., & Knierim, J. J. (2008). Influence of boundary removal on the spatial representations of the medial entorhinal cortex. Hippocampus, 18(12), 1270–1282. https://doi.org/10.1002/hipo.20511

Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. Analytical chemistry, 36(8), 1627-1639.

Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. Neuron, 91(6), 1402–1412. https://doi.org/10.1016/j.neuron.2016.08.019

Schuck, N. W., & Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. Science, 364(6447), eaaw5181. https://doi.org/10.1126/science.aaw5181

Schultz, W., Dayan, P., & Montague, P. R. (n.d.). A Neural Substrate of Prediction and Reward. 8.

Scoville, W. B., & Milner, B. (n.d.). LOSS OF RECENT MEMORY AFTER BILATERAL HIPPOCAMPAL LESIONS. 11.

Shadlen, M. N., & Shohamy, D. (2016). Decision Making and Sequential Sampling from Memory. Neuron, 90(5), 927–939. https://doi.org/10.1016/j.neuron.2016.04.036

Shamash, P., Olesen, S. F., Iordanidou, P., Campagner, D., Banerjee, N., & Branco, T. (2021). Mice learn multi-step routes by memorizing subgoal locations. Nature Neuroscience, 24(9), 1270-1279.

Sharpe, M. J., Batchelor, H. M., Mueller, L. E., Yun Chang, C., Maes, E. J. P., Niv, Y., & Schoenbaum, G. (2020). Dopamine transients do not act as

model-free prediction errors during associative learning. Nature Communications, 11(1), 106. https://doi.org/10.1038/s41467-019-13953-1

Sharpe, M. J., Stalnaker, T., Schuck, N. W., Killcross, S., Schoenbaum, G., & Niv, Y. (2019). An Integrated Model of Action Selection: Distinct Modes of Cortical Control of Striatal Decision Making. Annual Review of Psychology, 70(1), 53–76. https://doi.org/10.1146/annurev-psych-010418-102824

Sherrill, K. R., Erdem, U. M., Ross, R. S., Brown, T. I., Hasselmo, M. E., & Stern, C. E. (2013). Hippocampus and retrosplenial cortex combine path integration signals for successful navigation. Journal of Neuroscience, 33(49), 19304-19313.

Shin, J. D., Tang, W., & Jadhav, S. P. (2019). Awake hippocampal-prefrontal replay mediates spatial learning and decision making [Preprint]. Neuroscience. https://doi.org/10.1101/632042

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., ... & Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science, 362(6419), 1140-1144.

Simon, D. A., & Daw, N. D. (2011). Neural Correlates of Forward Planning in a Spatial Decision Task in Humans. Journal of Neuroscience, 31(14), 5526–5539. https://doi.org/10.1523/JNEUROSCI.4647-10.2011

Solstad, T., Boccara, C. N., Kropff, E., Moser, M.-B., & Moser, E. I. (2008). Representation of Geometric Borders in the Entorhinal Cortex. Science, 322(5909), 1865–1868. https://doi.org/10.1126/science.1166466

Souman, J. L., Frissen, I., Sreenivasa, M. N., & Ernst, M. O. (2009). Walking Straight into Circles. Current Biology, 19(18), 1538–1542. https://doi.org/10.1016/j.cub.2009.07.053

Spiers, H. J. (2001). Hippocampal Amnesia. https://doi.org/10.1076/neur.7.5.357.16245

Spiers, H. J. (2020). The Hippocampal Cognitive Map: One Space or Many? Trends in Cognitive Sciences, 24(3), 168–170. https://doi.org/10.1016/j.tics.2019.12.013

Spiers, H. J., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe,

J. (2001). Bilateral hippocampal pathology impairs topographical and episodic memory but not visual pattern matching. Hippocampus, 11(6), 715–725. https://doi.org/10.1002/hipo.1087

Spiers, H. J., & Gilbert, S. J. (2015). Solving the detour problem in navigation: A model of prefrontal and hippocampal interactions. Frontiers in Human Neuroscience, 9. https://doi.org/10.3389/fnhum.2015.00125

Spiers, H. J., & Maguire, E. A. (2006a). Spontaneous mentalizing during an interactive real world task: An fMRI study. Neuropsychologia, 44(10), 1674–1682. https://doi.org/10.1016/j.neuropsychologia.2006.03.028

Spiers, H. J., & Maguire, E. A. (2006b). Thoughts, behaviour, and brain dynamics during navigation in the real world. NeuroImage, 31(4), 1826–1840. https://doi.org/10.1016/j.neuroimage.2006.01.037

Spiers, H. J., & Maguire, E. A. (2007). A navigational guidance system in the human brain. Hippocampus, 17(8), 618–626. https://doi.org/10.1002/hipo.20298

Spiers, H. J., Olafsdottir, H. F., & Lever, C. (2018). Hippocampal CA1 activity correlated with the distance to the goal and navigation performance. Hippocampus, 28(9), 644–658. https://doi.org/10.1002/hipo.22813

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. Cold Spring Harbor perspectives in biology, 7(8), a021766.

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. Nature Neuroscience, 20(11), 1643–1653. https://doi.org/10.1038/nn.4650

Stalnaker, T. A., Berg, B., Aujla, N., & Schoenbaum, G. (2016). Cholinergic Interneurons Use Orbitofrontal Input to Track Beliefs about Current State. The Journal of Neuroscience, 36(23), 6242–6257. https://doi.org/10.1523/JNEUROSCI.0157-16.2016

Stangl, M., Topalovic, U., Inman, C. S., Hiller, S., Villaroman, D., Aghajan, Z. M., Christov-Moore, L., Hasulak, N. R., Rao, V. R., Halpern, C. H., Eliashiv, D., Fried, I., & Suthana, N. (2021). Boundary-anchored neural mechanisms of location-encoding for self and others. Nature, 589(7842),

420–425. https://doi.org/10.1038/s41586-020-03073-y

Starkweather, C. K., Babayan, B. M., Uchida, N., & Gershman, S. J. (2017). Dopamine reward prediction errors reflect hidden-state inference across time. Nature Neuroscience, 20(4), 581–589. https://doi.org/10.1038/nn.4520

Steele, R. J., & Morris, R. G. M. (n.d.). Delay-dependent impairment of a matching-to-place task with chronic and intrahippocampal infusion of the NMDA-antagonist D-AP5. 19.

Stella, F., Baracskay, P., O'Neill, J., & Csicsvari, J. (2019). Hippocampal Reactivation of Random Trajectories Resembling Brownian Diffusion. Neuron, 102(2), 450-461.e7. https://doi.org/10.1016/j.neuron.2019.01.052

Summerfield, C., & Tsetsos, K. (2012). Building Bridges between Perceptual and Economic Decision-Making: Neural and Computational Mechanisms. Frontiers in Neuroscience, 6. https://doi.org/10.3389/fnins.2012.00070

Sutherland, R. J., Whishaw, I. Q., & Kolb, B. (1983). A behavioural analysis of spatial localization following electrolytic, kainate- or colchicine-induced damage to the hippocampal formation in the rat. Behavioural Brain Research, 7(2), 133–153. https://doi.org/10.1016/0166-4328(83)90188-2

Sutton, R. S., Precup, D., & Singh, S. (1999). Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. Artificial Intelligence, 112(1–2), 181–211. https://doi.org/10.1016/S0004-3702(99)00052-1

Sutton, R. S., Barto, A. G. et al. (1998), Reinforcement learning: An introduction.

Sutton, R. S., Barto, A. G. et al. (2018), Reinforcement learning: An introduction.

Sutton, R. S., Precup, D. and Singh, S. (1999), 'Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning', Artif. Intell. 112(1-2), 181–211.

Taube, J. S. (2007). The Head Direction Signal: Origins and Sensory-Motor Integration. Annual Review of Neuroscience, 30(1), 181–207.

https://doi.org/10.1146/annurev.neuro.29.051605.112854

Tavares, R. M., Mendelsohn, A., Grossman, Y., Williams, C. H., Shapiro, M., Trope, Y., & Schiller, D. (2015). A Map for Social Navigation in the Human Brain. Neuron, 87(1), 231–243. https://doi.org/10.1016/j.neuron.2015.06.011

Tessereau, C., O'Dea, R., Coombes, S., & Bast, T. (2020). Reinforcement Learning approaches to hippocampus-dependant flexible spatial navigation [Preprint]. Neuroscience. https://doi.org/10.1101/2020.07.30.229005

The caudate nucleus egocentric localization system.pdf. (n.d.).

Thorn, C. A., Atallah, H., Howe, M., & Graybiel, A. M. (2010). Differential Dynamics of Activity Changes in Dorsolateral and Dorsomedial Striatal Loops during Learning. Neuron, 66(5), 781–795. https://doi.org/10.1016/j.neuron.2010.04.036

Toledo, S., Shohami, D., Schiffner, I., Lourie, E., Orchan, Y., Bartan, Y., & Nathan, R. (2020). Cognitive map–based navigation in wild bats revealed by a new high-throughput tracking system. Science, 369(6500), 188–193. https://doi.org/10.1126/science.aax6904

Toledo, S., Shohami, D., Schiffner, I., Lourie, E., Orchan, Y., Bartan, Y., & Nathan, R. (2020). Cognitive map–based navigation in wild bats revealed by a new high-throughput tracking system. Science, 369(6500), 188-193.

Tolman, E. C. (1948). Cognitive maps in rats and men. Psychological Review, 55(4), 189–208. https://doi.org/10.1037/h0061626

Tolman, E. C., Ritchie, B. F., & Kalish', D. (n.d.). STUDIES IN SPATIAL LEARNING. I. ORIENTATION AND THE SHORT-GUT. 12.

Ulanovsky, N., & Moss, C. F. (2007). Hippocampal cellular and network activity in freely moving echolocating bats. Nature Neuroscience, 10(2), 224–233. https://doi.org/10.1038/nn1829

van der Meer, M. A. A., Johnson, A., Schmitzer-Torbert, N. C., & Redish, A. D. (2010). Triple Dissociation of Information Processing in Dorsal Striatum, Ventral Striatum, and Hippocampus on a Learned Spatial Decision Task. Neuron, 67(1), 25–32. https://doi.org/10.1016/j.neuron.2010.06.023

van der Meer, M. A., & Redish, A. D. (2011). Ventral striatum: A critical look at models of learning and evaluation. Current Opinion in Neurobiology, 21(3), 387–392. https://doi.org/10.1016/j.conb.2011.02.011

Vaz, A. P., Wittig, J. H., Inati, S. K., & Zaghloul, K. A. (2020). Replay of cortical spiking sequences during human memory retrieval. Science, 367(6482), 1131–1134. https://doi.org/10.1126/science.aba0672

Vertechi, P., Lottem, E., Sarra, D., Godinho, B., Treves, I., Quendera, T., Oude Lohuis, M. N., & Mainen, Z. F. (2020). Inference-Based Decisions in a Hidden State Foraging Task: Differential Contributions of Prefrontal Cortical Areas. Neuron, 106(1), 166-176.e6. https://doi.org/10.1016/j.neuron.2020.01.017

Viard, A., Doeller, C. F., Hartley, T., Bird, C. M., & Burgess, N. (2011). Anterior Hippocampus and Goal-Directed Spatial Decision Making. Journal of Neuroscience, 31(12), 4613–4621. https://doi.org/10.1523/JNEUROSCI.4640-10.2011

Vikbladh, O. M., Meager, M. R., King, J., Blackmon, K., Devinsky, O., Shohamy, D., Burgess, N., & Daw, N. D. (2019). Hippocampal Contributions to Model-Based Planning and Spatial Memory. Neuron, 102(3), 683-693.e4. https://doi.org/10.1016/j.neuron.2019.02.014

Voermans, N. C., Petersson, K. M., Daudey, L., Weber, B., van Spaendonck, K. P., Kremer, H. P. H., & Fernández, G. (2004). Interaction between the Human Hippocampus and the Caudate Nucleus during Route Recognition. Neuron, 43(3), 427–435. https://doi.org/10.1016/j.neuron.2004.07.009

Walton, M. E., Behrens, T. E. J., Buckley, M. J., Rudebeck, P. H., & Rushworth, M. F. S. (2010). Separable Learning Systems in the Macaque Brain and the Role of Orbitofrontal Cortex in Contingent Learning. Neuron, 65(6), 927–939. https://doi.org/10.1016/j.neuron.2010.02.027

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing, 13(4), 600-612.

Warren, W. H. (2019). Non-Euclidean navigation. The Journal of Experimental Biology, 222(Suppl 1), jeb187971. https://doi.org/10.1242/jeb.187971

Warren, W. H., Rothman, D. B., Schnapp, B. H., & Ericson, J. D. (2017). Wormholes in virtual space: From cognitive maps to cognitive graphs. Cognition, 166, 152–163. https://doi.org/10.1016/j.cognition.2017.05.020

Weisberg, S. M., & Ekstrom, A. D. (2021). Hippocampal volume and navigational ability: The map(ping) is not to scale. Neuroscience & Biobehavioral Reviews, 126, 102–112. https://doi.org/10.1016/j.neubiorev.2021.03.012

Weisberg, S. M., & Newcombe, N. S. (2018). Cognitive Maps: Some People Make Them, Some People Struggle. Current Directions in Psychological Science, 27(4), 220–226. https://doi.org/10.1177/0963721417744521

Weisberg, S. M., Newcombe, N. S., & Chatterjee, A. (2019). Everyday taxi drivers: Do better navigators have larger hippocampi?. Cortex, 115, 280-293.

Weisberg, S. M., Schinazi, V. R., Newcombe, N. S., Shipley, T. F., & Epstein, R. A. (2014). Variations in cognitive maps: Understanding individual differences in navigation. Journal of Experimental Psychology: Learning, Memory, and Cognition, 40(3), 669–682. https://doi.org/10.1037/a0035261

Weiskopf, N., Hutton, C., Josephs, O., Turner, R., & Deichmann, R. (2007). Optimized EPI for fMRI studies of the orbitofrontal cortex: Compensation of susceptibility-induced gradients in the readout direction. Magnetic Resonance Materials in Physics, Biology and Medicine, 20(1), 39. https://doi.org/10.1007/s10334-006-0067-6

Whishaw, I. Q., Mittleman, G., Bunch, S. T., & Dunnett, S. B. (1987). Impairments in the acquisition, retention and selection of spatial navigation strategies after medial caudate-putamen lesions in rats. Behavioural Brain Research, 24(2), 125–138. https://doi.org/10.1016/0166-4328(87)90250-6

White, N. M. (1989). A functional hypothesis concerning the striatal matrix and patches: Mediation of S-R memory and reward. Life Sciences, 45(21), 1943–1957. https://doi.org/10.1016/0024-3205(89)90569-9

White, N. M., & McDonald, R. J. (2002). Multiple Parallel Memory Systems in the Brain of the Rat. Neurobiology of Learning and Memory, 77(2), 125–184. https://doi.org/10.1006/nlme.2001.4008

Wiener, J. M., & Mallot, H. A. (2003). 'Fine-to-Coarse' Route Planning and Navigation in Regionalized Environments. Spatial Cognition & Computation, 3(4), 331–358. https://doi.org/10.1207/s15427633scc0304_5

Wikenheiser, A. M., Gardner, M. P. H., Mueller, L. E., & Schoenbaum, G. (2021). Spatial Representations in Rat Orbitofrontal Cortex. The Journal of Neuroscience, 41(32), 6933–6945. https://doi.org/10.1523/JNEUROSCI.0830-21.2021

Wikenheiser, A. M., & Redish, A. D. (2015). Hippocampal theta sequences reflect current goals. Nature Neuroscience, 18(2), 289–294. https://doi.org/10.1038/nn.3909

Wills, T. J., Barry, C., & Cacucci, F. (2012). The abrupt development of adult-like grid cell firing in the medial entorhinal cortex. Frontiers in Neural Circuits, 6. https://doi.org/10.3389/fncir.2012.00021

Wilson, M., & McNaughton, B. (1994). Reactivation of hippocampal ensemble memories during sleep. Science, 265(5172), 676–679. https://doi.org/10.1126/science.8036517

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. ELife, 8, e49547. https://doi.org/10.7554/eLife.49547

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal Cortex as a Cognitive Map of Task Space. Neuron, 81(2), 267–279. https://doi.org/10.1016/j.neuron.2013.11.005

Wilson, S. P., & Wilson, P. N. (2018). Failure to demonstrate short-cutting in a replication and extension of Tolman et al.'s spatial learning experiment with humans. Plos one, 13(12), e0208794.

Wolbers, T., & Hegarty, M. (2010). What determines our navigational abilities? Trends in Cognitive Sciences, 14(3), 138–146. https://doi.org/10.1016/j.tics.2010.01.001

Woo, C. W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresh-

olding in fMRI analyses: pitfalls and recommendations. Neuroimage, 91, 412-419.

Woollett, K., & Maguire, E. A. (2010). The effect of navigational expertise on wayfinding in new environments. Journal of Environmental Psychology, 30(4), 565–573. https://doi.org/10.1016/j.jenvp.2010.03.003

Woollett, K., & Maguire, E. A. (2011). Acquiring "the Knowledge" of London's Layout Drives Structural Brain Changes. Current Biology, 21(24), 2109–2114. https://doi.org/10.1016/j.cub.2011.11.018

Woollett, K., Spiers, H. J., & Maguire, E. A. (2009). Talent in the taxi: A model system for exploring expertise. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1522), 1407–1416. https://doi.org/10.1098/rstb.2008.0288

Woolley, D. G., Laeremans, A., Gantois, I., Mantini, D., Vermaercke, B., Op de Beeck, H. P., Swinnen, S. P., Wenderoth, N., Arckens, L., & D'Hooge, R. (2013). Homologous involvement of striatum and prefrontal cortex in rodent and human water maze learning. Proceedings of the National Academy of Sciences, 110(8), 3131–3136. https://doi.org/10.1073/pnas.1217832110

Woolley, D. G., Mantini, D., Coxon, J. P., D'Hooge, R., Swinnen, S. P., & Wenderoth, N. (2015). Virtual water maze learning in human increases functional connectivity between posterior hippocampus and dorsal caudate: Spatial Learning and Functional Connectivity. Human Brain Mapping, 36(4), 1265–1277. https://doi.org/10.1002/hbm.22700

Wu, C. M., Schulz, E., Garvert, M. M., Meder, B., & Schuck, N. W. (2020). Similarities and differences in spatial and non-spatial cognitive maps. PLOS Computational Biology, 16(9), e1008149. https://doi.org/10.1371/journal.pcbi.1008149

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. Nature Human Behaviour, 2(12), 915–924. https://doi.org/10.1038/s41562-018-0467-4

Wunderlich, K., Dayan, P., & Dolan, R. J. (2012). Mapping value based planning and extensively trained choice in the human brain. Nature Neuroscience, 15(5), 786–791. https://doi.org/10.1038/nn.3068

Xu, H., Baracskay, P., O'Neill, J., & Csicsvari, J. (2019). Assembly Responses of Hippocampal CA1 Place Cells Predict Learned Behavior in Goal-Directed Spatial Tasks on the Radial Eight-Arm Maze. Neuron, 101(1), 119-132.e4. https://doi.org/10.1016/j.neuron.2018.11.015

Xu, J., Evensmoen, H. R., Lehn, H., Pintzka, C. W. S., & Haberg, A. K. (2010). Persistent posterior and transient anterior medial temporal lobe activity during navigation. NeuroImage, 52(4), 1654–1666. https://doi.org/10.1016/j.neuroimage.2010.05.074

Yartsev, M. M., Witter, M. P., & Ulanovsky, N. (2011). Grid cells without theta oscillations in the entorhinal cortex of bats. Nature, 479(7371), 103–107. https://doi.org/10.1038/nature10583

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. Nature Reviews Neuroscience, 7(6), 464–476. https://doi.org/10.1038/nrn1919

Yoder, R. M., Clark, B. J., & Taube, J. S. (2011). Origins of landmark encoding in the brain. Trends in Neurosciences, 34(11), 561–571. https://doi.org/10.1016/j.tins.2011.08.004

Yoo, S. B. M., Tu, J. C., & Hayden, B. Y. (2021). Multicentric tracking of multiple agents by anterior cingulate cortex during pursuit and evasion. Nature Communications, 12(1), 1985. https://doi.org/10.1038/s41467-021-22195-z

Zhao, M. (2018). Human spatial representation: what we cannot learn from the studies of rodent navigation. Journal of neurophysiology, 120(5), 2453-2465.