Expert Navigators Deploy Rational Hierarchical Priorization Over Predictive Maps For Large-Scale Real-World Planning

Pablo Fernandez Velasco[1,2]*, Eva-Maria Griesbauer[1]*, Iva Brunec[3], Jeremy Morley[4], Ed Manley[5,6], Daniel C. McNamee[7†], Hugo J. Spiers[1†]

1 Institute of Behavioural Neuroscience, Department of Experimental Psychology, University College London, London, UK.
2 Department of Philosophy, University of York, York, UK
3 Department of Psychology, University of Pennsylvania, USA
4 Ordnance Survey Ltd, Southampton, UK.
5 School of Geography, University of Leeds, Leeds, UK
6 Centre for Advanced Spatial Analysis, University College London, London, UK
7 Neuroscience Programme, Champalimaud Research, Centre for the Unknown, Lisbon, Portugal
*equal contribution
†equal contribution

Abstract

Efficient planning is a distinctive hallmark of human intelligence. Computational analyses of human behavior during planning provide evidence for hierarchically segmented representations of the state space of options. However, such evidence derives from simplistic tasks with small state-spaces that belie the complexity of real-world planning. Here, we examine the street-by-street route plans of London taxi drivers navigating across more than 26,000 streets in London (UK). Response times were faster for states with higher successor representations, providing evidence for predictive mapping. We also find an effect for the interaction between a state's successor representation and local transition entropy, which indicates hierarchical chunking of transition sequences, and thus provides real-world support to existing theories of hierarchical state-space representations and planning. Finally, we explored how planning unfolded dynamically over different phases of the constructed journey and identify theoretic principles by which expert planners rationally prioritize specific states during the planning process. Overall, our findings provide real-world evidence for predictive maps and rational hierarchical prioritization in expert route planning.

Introduction

In the novel *Invisible Cities*, the author Italo Calvino plays on an analogy between navigating through an urban street network and making moves in a game of chess. Both require considering future actions and their outcomes in order to decide which way to move, that is, planning (Calvino, 1978). This analogy between the real and the abstract is a common presupposition in the study of cognitive mechanisms of expert human reasoning (Newell & Simon, 1972). Indeed, to date, computational modeling studies of human planning have relied on abstract games such as chess, or Go, or the tower of Hanoi problem for experimental verification (Allen et al, 2023). Thus, the degree to which theoretic principles developed in the fields of cognitive science and artificial intelligence can be extrapolated to the cognitive

computations of human domain experts solving large-scale planning problems in the real world is largely unknown.

The classic approach to planning, both in the cognitive sciences (Daw, Niv, & Dayan, 2005; Huys et al., 2012; van Opheusden et al., 2023) and in artificial intelligence (Shannon, 1950; Silver et al., 2016; Russell and Norvig 2010), is to implement a tree search algorithm. This approach involves a representation of possible trajectories across a state-space, with potential actions as branching points. Each sampled trajectory along the decision tree results in an estimated expected reward, and action plans are decided on based on those expected rewards (Streeter & Vitello, 1986; Elliott & Lesk, 1982; Miller & Venditto, 2020). The fundamental difficulty with algorithms based on tree search is that they become computationally intractable in complex state-spaces. This so-called curse of dimensionality (Bellman, 1957) presents a puzzle for modeling human cognition, given that humans have the capability to plan in vast state-spaces in which humans operate under substantial time pressures and with limited working memory capacity (McNamee, Wolpert, and Lengyel, 2016; Griffiths, 2020).

And yet, humans routinely defy the curse of dimensionality, by planning successfully in extremely complex environments such as cities. Recent approaches in computational modeling have sought to specify how humans do this. Existing studies have modeled how people allocate limited computational resources (Snider et al., 2015; Callaway et al., 2018), modulate the depth of their planning (Keramati et al., 2016; van Opheusden et al., 2023), prune initially unpromising paths (Huys et al., 2012), and identify how much planning to perform (Russek et al., 2022). Some of this work has also investigated possible modifications of the problem representation itself, through different forms of simplified representations (Ho et al., 2022) and task decomposition (Solway et al., 2014; Huys et al., 2015; Correa et al., 2023) leading to hierarchical planning (Lashley, 1951; Botvinick, 2012; McNamee, Wolpert & Lengyel, 2016; O'Doherty, Lee & McNamee, 2015; Botvinick, Niv & Barto, 2009; Tomov, Yagati, Kumar, Yang & Gershman, 2020; Bast, et al., 2015).

In particular, hierarchical state-space models segment a state-space into distinct sets of connected states (which are referred to as clusters). Global planning can then be carried out in order to identify a trajectory through critical junction states which interface between clusters. More localized planning is then implemented within a cluster and thus is restricted to a subset of the states. The hierarchical representation of the state-space combined with global and local planning counters the curse of dimensionality by reducing the number of states which must be considered (McNamee, Wolpert & Lengyel, 2016). Furthermore, local planning may make use of chunked action sequences for further efficiency gains as in the options framework of hierarchical reinforcement learning (Botvinick, Niv and Barto, 2009; Jinnai et al 2019). Hierarchical planning has been studied in a broad range of experimental and theoretic settings, such as problem-solving games (Ward & Allport, 1997; Knoblock, 1990), regionalised virtual environments (Wiener & Mallot, 2003; Wiener, Schnee & Mallot, 2004), the linguistic impact of semantics and preferences for regional crossings (Schick et al., 2019; Schick, 2018), latent graphs governing transitions between object stimuli (e.g. Wiener, Ehbauer & Mallot, 2009; Hommel, Gehrke & Knuf, 2000; Solway et al., 2014) or a virtual subway network task (Balaguer et al., 2016). Evidence from these studies suggests not only the existence of regionalised representations of different environments, but also that humans exploit hierarchical levels to minimize planning demands (Balaguer et al., 2016; Schick et al., 2019; McNamee, Wolper & Lengyel, 2016; Correa et al 2023).

However, most studies to date have employed either virtual small-scale or abstract environments (Wiener & Mallot, 2009; Wiener, Schnee & Mallot, 2004; Schick et al., 2019; Büchner, Hölscher & Strube, 2007; Badre, Kayser & D'Esposito, 2010). As a result, it remains unclear how humans process and exploit state-space hierarchies in large, real-world environments. Cities epitomize the intricacy of planning in real-world environments. If we considered planning a path across a large city that between start and

destination involved 30 streets, using a minimal tree-search algorithm would involve an evaluation of over 1 billion potential street sequences (Huys et al., 2012). The complexity of cities also lies in the fact that they lack definitive regional segregation. For example, in London (UK) there is no definite boundary between *Farringdon* and *City of London*, despite both regions being well known to Londoners (Griesbauer et al. 2022a). In contrast, the state-spaces used in experimental work are often designed so as to accentuate particular hierarchical decompositions (Balaguer et al., 2016; Arnold, Iaria & Ekstrom, 2016; Wiener & Mallot, 2009; Wiener, Schnee & Mallot, 2004; Schick et al., 2019). A case in point is the study by Balaguer and colleagues (2016), in which a virtual subway network was divided into subway lines, which facilitated hierarchically organized planning. Whether the expert human planning in large real-world environments without artificially imposed state-space hierarchy also involves hierarchical processing, and potentially other strategies to increase computational efficiency, remains to be seen.

Here, we sought real-world evidence of theoretic principles underlying human expert planning in ecologically valid, large-scale, state-spaces. Such an effort faces two key challenges, one from an analysis perspective, and the other from an experimental perspective.

From an analysis perspective, the size of the state-space presents a significant challenge for fitting and evaluating mechanistic planning models, thus, in this work, we pursued a descriptive statistical approach based on prior theoretic analyses of trajectory search efficiency resulting in optimized hierarchies for planning (McNamee, Wolpert and Lengyel, 2016). Indeed, this approach was previously used to infer the existence of hierarchical representations in laboratory experiments involving the mental simulation of real-world routes (e.g. Bonasia, Blommesteyn & Moscovitch, 2015; Arnold, Iaria & Ekstrom, 2016). A recent successful approach to testing spatial planning has been to examine the reaction times of choices when navigating step by step in a fictitious subway network (Balaguer et al., 2016). In this approach, increased reaction times imply higher computational demands on the brain to make the next choice, and can be employed to infer the computational structure underlying a planning process. Here we extend this approach from a state-space of 35 states to a state space of over 26,000 states: the street network of London (UK).

From an experimental perspective, the challenge was to identify a large and ecologically valid real-world state-space, and then to recruit a human participant population with expert domain knowledge thereof. To explore how people navigate, ideally participants would tell you their planned route step-by-step. The issue is that the larger the street network, the harder it will be to find participants who can plan routes and recall the names of all the streets between any two points. We propose that licensed London taxi drivers provide an exceptional opportunity to meet this challenge. London taxi drivers are unique in the world for their knowledge of street networks and navigational expertise (Fernandez Velasco and Spiers 2024). They are unique with respect to their vast experience and standardized expert training of the London street network (Griesbauer, 2022b), and they have been well characterized in prior studies exploring cognition, brain function and structural changes as a result of expertise (Maguire et al., 2000; Maguire, Woollett & Spiers, 2006; Maguire, Nannery & Spiers, 2006; Spiers & Maguire, 2006a,b, 2007a,b,c, 2008). Indeed, in order to obtain their license, London taxi drivers are specifically examined on their ability to mentally navigate through London by naming all the streets and turns in a planned route between a given origin and destination. This familiar task paradigm forms the basis of our experimental approach (Griesbauer, 2022b). Participants were asked to plan routes between a set of origin-destination pairs, and then “call out” the routes step by step. The collected verbal data of the route “call outs” was transcribed with regards to street names and to response times. We then employ information-theoretic models to provide predictor variables for the statistical analysis of data collected from forty-three taxi drivers. Specifically, we used linear models to examine the effect of environmental,

cognitive, and theoretical metrics on route planning, with the aim of testing for hierarchical planning and for the impact of different variable classes, spatial and theoretic, on real-world large-scale expert planning.

Results

Data extraction and preparation. Response times were extracted from transcribed verbal reports of London routes (Figure 1). Participants were given an origin and a destination, and they had to "call out" the route step by step. For each route, the response time data consisted of an offline planning time (the time from presentation to the first call-out) and subsequent call-out times (the times between each call-out). Mean *offline planning time* was M = 13.83s (*SD* = 13.40) over N = 315 routes. On average, taxi drivers recalled 9.1 routes out of 12 during the first and 7.5 routes out of 8 during the second experiment. The average *total response duration* across tasks was M = 17.53s (*SD* = 16.47). The total number of pauses between two consecutively named streets was 3398, with a mean *call-out time* of M = 1.82s (*SD* = 3.24). For more details of the task see Griesbauer et al (2024).

FIGURE 1 caption next page

FIGURE 1. Route planning task. London taxi drivers received pre-recorded audio instructions of the origin and destination locations and 'called out' the names of all the streets they would navigate between the two points, after which they were presented with the next origin-destination pair. All instructions and route plans were audio recorded and transcribed to extract response times. The time lapsed between the instructions and the first recalled street was defined as the initial response time (corresponding to offline planning). Time lapses between "called" streets of the route were defined as response times between streets (online planning). **A**) is an overlay of the responses from all participants for all routes. For the same origin-destination pair, different participants called different routes. The white rectangle indicates the area covered in route 8 (i.e. as in Fig 1B-1D). **B)** is an example of route 8 (Joe Allen's restaurant to Courthouse Hotel) for a single subject. The subject "called" the route turn by turn. In the map is the resulting route as it evolved, indicating the corresponding time (top-right of each map) and streets called (left). **C**) represents the whole route mapped for steps shown in B, but with the size of the circumferences corresponding to the length of the time lapses between recalled streets. **D**) shows the "call out" times (Response times, average z-scores). Notice the long pause before Shaftesbury Avenue that is consistent across all taxi drivers (corresponding to the largest circle in B). See Supplementary Movie 1 for an audio recording of a taxi driver calling out route 8 and an animation of the corresponding route evolving turn-by-turn.

Linear multivariate analysis of response times. Fixed effects linear regression analysis revealed that the following variables had significant effects on response times: *Euclidean distance, segment length, remaining streetwise distance, remaining direct distance, successor representation, segment angular deviation, destination angular deviation,* and *SR-LTE interaction* (Table 1, see Figure 2 for a graphical illustration of SR-LTE interaction).

Variable	Coefficient	Standard Error	t	P>|t|	Confidence interval [0.025	0.975]
Intercept	0.0251	0.051	0.494	0.621	-0.075	0.125
Initial planning time	-0.0073	0.022	-0.333	0.739	-0.05	0.036
Plan step number	0.0025	0.001	1.816	0.069	0	0.005
Number of planning steps	-0.0063	0.005	-1.234	0.217	-0.016	0.004
Direct distance [route total]	0.1805	0.123	1.471	0.141	-0.06	0.421
Streetwise distance [route total]	**-0.5201**	**0.156**	**-3.333**	**0.001**	**-0.826**	**-0.214**
Street segment length	**0.1974**	**0.025**	**7.82**	**0.000**	**0.148**	**0.247**
Streetwise distance [remaining]	**0.328**	**0.107**	**3.068**	**0.002**	**0.118**	**0.538**
Direct distance [remaining]	-0.11	0.105	-1.049	0.294	-0.316	0.096
Segment angular deviation	-0.013	0.06	-0.217	0.828	-0.131	0.105
Destination angular deviation	**0.2233**	**0.087**	**2.575**	**0.01**	**0.053**	**0.393**

Successor representation (SR)	**-0.3414**	**0.113**	**-3.031**	**0.002**	**-0.562**	**-0.12**
Local transition entropy (LTE)	0.0325	0.022	1.459	0.145	-0.011	0.076
SR-LTE interaction	**0.2092**	**0.071**	**2.957**	**0.003**	**0.07**	**0.348**

Table 1. Ordinary least-squares regression results. The dependent variable is the log-transformed and z-standardised call-out times. Variables that reached significance are highlighted in bold. Consistent with theoretic analysis of planning complexity, the abstract state-space variables, namely SR and local transition entropy, significantly impacted call-out times independent of environmental variables which characterize the veridical spatial context of the London routes.

Since the product of SR and LTE provides a normative localized metric of planning complexity, we review in detail how they impacted call-out times (McNamee, Wolpert, and Lengyel, 2016). There was a highly significant and negative main effect of SR on call-out times such that more predictable states were associated with faster COTs. This effect did not interact with planning step number ($p=0.16$) indicating that it was consistently present through the planning process. There was no significant main effect of LTE. In the following sections, we highlight several highly significant interaction effects involving LTE consistent with distinct rational strategies for efficient planning.

Successor representation and local transition entropy measures interact to influence call-out times consistent with hierarchical chunking.

The significant effect of the interaction of SR and LTE on call-out times throughout the route descriptions is driven by a selective reduction in response time for low LTE and high SR states (Figure 2C). The high SR measurement reflects a relatively high probability that this state is typically traversed while the low LTE implies a relatively simplified transition to be processed at that state (Figure 2A, top-right). For example, a single adjacent state (excluding the previous state in the trajectory) has no associated planning cost and corresponds to an LTE of zero (though may have an associated memory retrieval cost). Therefore, such state transitions may be automatically generated conditional on the previous state. Indeed, this reasoning provides the fundamental principle for several classical hierarchical state-space models which suggest more broadly that such transitions should be targeted for concatenation, or “chunking”, into extended automated sequences (Jinnai et al 2019).

Hierarchical chunking of fast automated transitions is a feature of an efficient hierarchical state-space representation (McNamee, Wolpert, and Lengyel, 2016) and thus should manifest in reduced response times independently of the planning dynamics. That is, the retrieval and expression of chunked transitions should result in faster response times throughout the planning process. In order to examine whether this interaction effect between SR measurements and LTE was present throughout the planning process, we ran the same statistical tests on the early call-outs (early online planning phase) and the late call-outs (late online planning phase) separately under a median split of the planning step number (Figure 3A). Notably, the main interaction effect whereby relatively faster call-outs are observed for low transition entropy states on a SR-dependent basis was present in both the early planning ($p<0.005$) and late planning ($p<0.008$) phases suggesting that this mechanism functions independently of online planning dynamics which we explore in the next section.

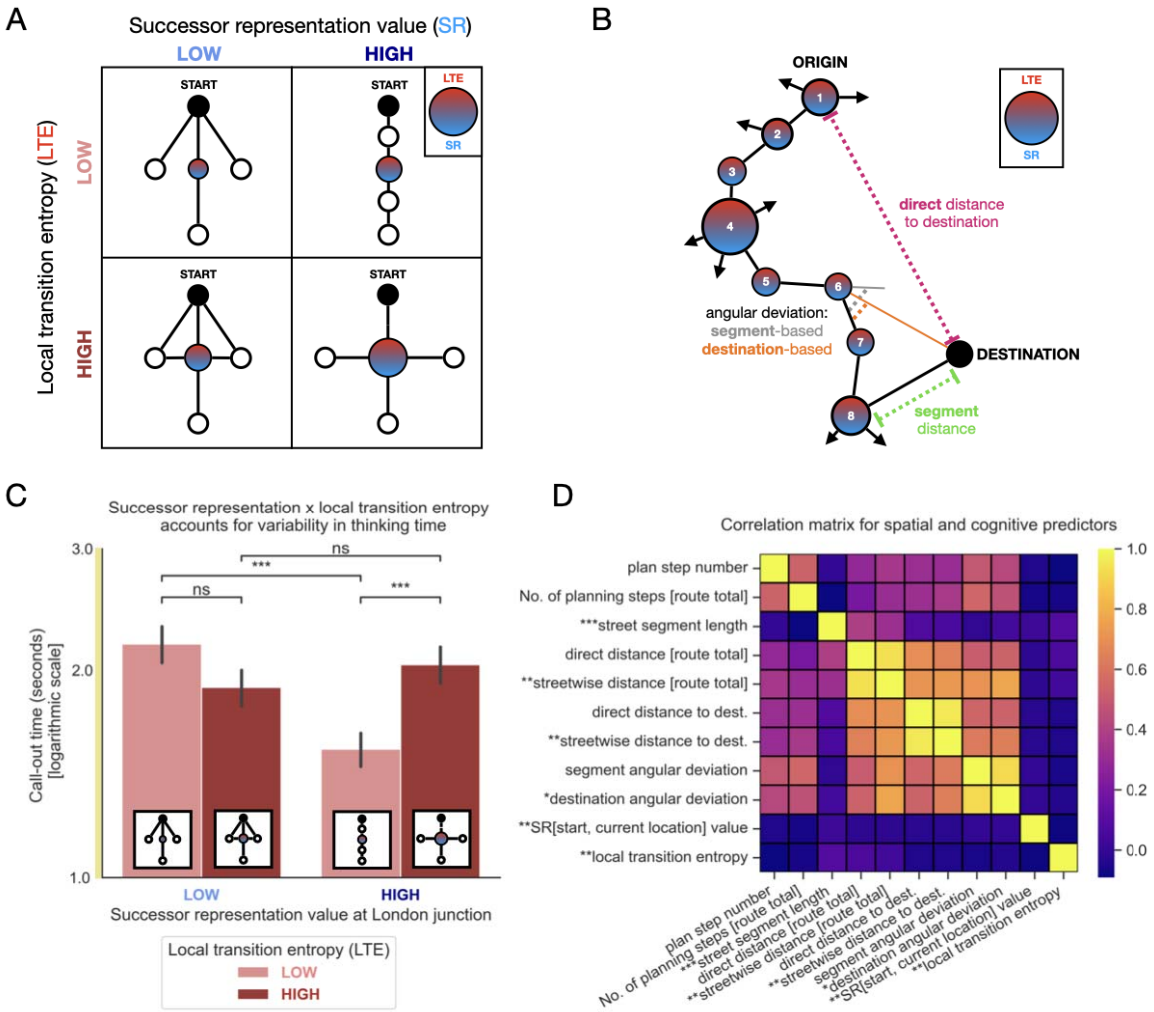

FIGURE 2. An interaction between variables from an information-theoretic model of planning predicts call-out times. A. B. C. In addition to a main effect of successor representation (SR) on call-out times (COTs) (p=0.002), there was a significant interaction between SR and local transition entropy (LTE) (p=0.003). This effect was driven by a reduction in COT specifically for low LTE street transitions with high SR measurements. That is, streets with relatively uncomplicated local structure and high throughput under random walk dynamics. Successive street segments of this type are normatively chunked together under hierarchical control models. Error bars reflect standard error of the mean. **D.** This correlation matrix demonstrates that the theoretic predictors, successor representation measurement and local transition entropy, varied independently with respect to each other and the spatial predictor variables.

Multi-phase dynamics in response times consistent with efficient non-sequential priorization in planning.

In our experiment, planning can take place before the first call-out (the *offline planning phase*) and then, may continue in parallel while the route is verbally reported. This latter *online planning phase* we divide into *early* and *late* periods (Figure 3A). We sought to examine how SR and LTE influence behavioral response times from a dynamic multi-phase perspective over the course of planning. LTE measures the local processing cost at a particular street segment, and it is weighted by SR in determining its contribution to global planning complexity (see Methods, Eqn. 1). In a state-space as

complex as London, there are inevitably too many states to consider and search assuming limited capacities for memory and planning. Given this, our theoretic analysis prescribes a rational strategy whereby high LTE states with high SR are prioritized for planning (Figure 3B).

In order to test this, we focused on the relationship between the length of the offline planning phase and the subsequent effects of SR and LTE on call-out times. We reasoned that a longer offline planning phase should be optimized to prospectively plan specifically at high SR and high LTE states and thus reduce their subsequent call-out times (Figure 3B). This hypothesis was tested using median-splits of the data according to the length of the offline planning period and comparing high and low SR (LTE) call-outs across this split. For longer offline planning periods, we found a significant and selective reduction in call-out time for both the SR (Figure 3C) and LTE (Figure 3D) variables. In contrast, a serial non-hierarchical planning model, whereby states are prioritized in the order they occur along the route, would predict an interaction effect of planning step number and the length of the offline planning period (Figure 3B). In particular, the call-out times would be relatively reduced for states closer to the origin; however such an effect was not present in the data ($p=0.648$, $t(43)=-0.457$).

The interaction between planning phase and LTE continued into the online planning period while verbal call-outs were being made by the taxi drivers. In particular, we found that call-out times selectively decreased for high LTE transitions from the early to late online planning phases (Figure 3E). These results suggest that SR and LTE measures may steer the rational use of cognitive resources (in this case, time) in prioritizing planning at critical states in a non-sequential order globally across the entire state-space (Figure 3B).

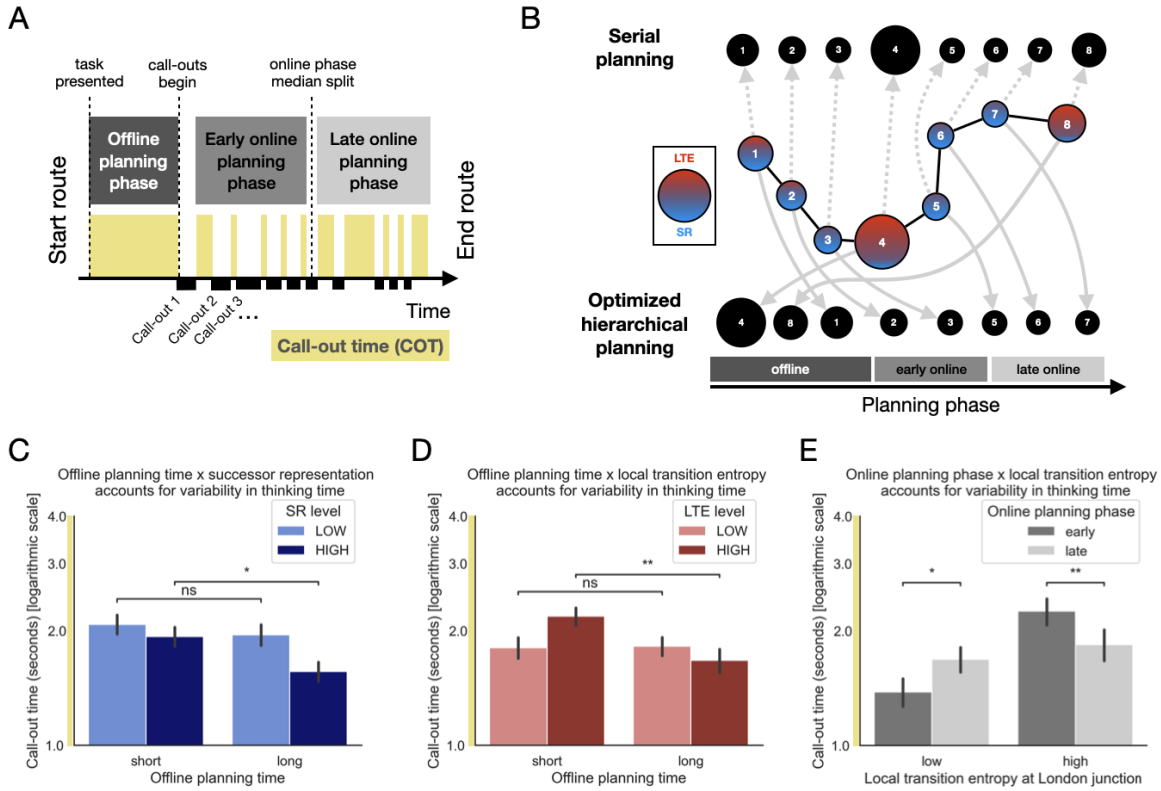

FIGURE 3. The local complexity of the street network, measured by local transition entropy, interacts with planning phase in accounting for variability in call-out time data. A. B. C. D. E. In particular, COTs

specifically for high complexity transitions become faster as planning progresses (p=0.006). Furthermore, the offline planning time (OPT) differentially interacts with theoretic variables (SR and LTE) in determining call-out times (COTs) between planning phases. In particular, an interaction between IPT and LTE (p=0.0001) and a three-way interaction between planning step number, IPT, and LTE (p=0.015) were observed. A longer initial planning period resulted in a significantly decreased COT specifically for complex local transition structures (high LTE). This effect was particularly pronounced in the early planning phase consistent with the continuous evolution of planning throughout the initial, early, and late planning phases. Error bars reflect standard error of the mean.

Discussion

In the present study, we set out to examine planning in a large-scale, ecologically valid, setting by testing licensed taxi drivers performing route planning tasks across London. London taxi drivers are uniquely positioned for this investigation because they undergo rigorous training and have a vast, first-hand experience of the street network. In our study, we asked forty-three taxi drivers to plan routes between a set of origin-destination pairs and describe those routes street by street. We then transcribed their responses and measured the response times before and while calling out the names of streets and turns along routes through the city. Using a linear multivariate regression analysis, we examined how a range of spatial, environmental, and cognitive variables impacted planning with response times serving as the behavioral measure. This approach exposed three principles of efficient expert planning, namely predictive mapping (Stachenfeld et al., 2017), hierarchical chunking (Jinnai et al., 2019; Correa et al 2023), and global non-sequential prioritization (McNamee, Wolpert, and Lengyel, 2016).

There was a main effect of successor representation whereby call-out times were reduced for states with higher SR measurements. The successor representation (SR) was initially developed as an efficient state-space representation strategy for reinforcement learning and planning as it facilitates the rapid evaluation of a given policy in an sequential decision-making environment (Dayan, 1993). Given the current state, the SR encodes the likelihood, or expected number of times, of being in another state in the future. This summarises the long-time relational structure of states based on a state transition policy (for a review of behavioural and neural evidence of *SR* in human cognition, see Gershman, 2018). That the SR of a given state (with respect to the origin state on a route) is correlated with faster responses from the taxi drivers is consistent with the internal use of a predictive map while navigating London. It is also in line with recent work showing the use of predictive maps in both human and non-human animals (de Cothi et al., 2022; Momennejad et al., 2017; Momennejad, 2024).

Furthermore, the data indicated an interaction between a state's SR measurement and local transition entropy (LTE). With respect to planning, local transition entropy provides a local measure of the degree of complexity in identifying the best route to take. This is because the overall trajectory entropy, which serves as a global measure of route complexity, is a linear sum of local transition entropies (see Methods, Eqn. 1). The higher the local transition entropy, the more the transition at the given state increases the global complexity and consequently the number of possible routes that need to be evaluated during planning. It was found that for low LTE states (i.e. those with relatively simple local transition structure, Fig. 2), higher SR resulted in faster response times relative to low SR states. This effect is consistent with a hierarchical chunking of transition sequences which are rapidly retrieved (and described by the taxi drivers) without intermediate planning (McNamee, Wolpert, and Lengyel, 2016; Jinnai et al 2019). Conversely, in terms of hierarchical segmentations of the state-space, high LTE states with high SR tend to form critical junction states which interface between adjacent clusters and thus the relatively slow call-outs at these states may be caused by internal

switches between distinct local state-space representations. Thus, expert navigators, such as taxi drivers, may exploit segmentations of London into hierarchically organised regions in order to optimise for planning demands (Balaguer et al., 2016; Schick et al., 2019). This result provides real-world support to previous theories of hierarchical planning coming from both computational and laboratory work (Botvinick, 2012; McNamee, Wolpert & Lengyel, 2016; O'Doherty, Lee & McNamee, 2015; Botvinick, Niv & Barto, 2009; Tomov et al., 2020; Bast et al.,, 2015). Figure 2A illustrates the interaction of SR and local transition entropy in abstract network structures. Notice how when we compare the two states with high LTE, high SR results in a considerably larger response time, with the opposite being the case for the state with LTE.

We also sought to determine how planning dynamically unfolded over different phases of each trial. Specifically, by examining the influence of the time length of the offline planning period on call-outs, we aimed to identify how expert planners rationally prioritize planning at specific states. According to the normative trajectory entropy minimization perspective, states should be prioritized according to the SR and LTE measures as knowledge of the correct transitions at such states provide the maximal reduction in the overall uncertainty regarding a particular route (McNamee, Wolpert & Lengyel, 2016; McNamee and Wolpert 2019). In accordance with this idea, the data we collected from expert London taxi drivers reflected selective reductions in response times at high SR and high LTE states for longer offline planning periods. Previous theoretic work based on the Q-DYNA variant of model-based RL characterized the prioritization of state sampling during planning as being dependent on so-called *need* and *gain* terms (Mattar and Daw 2018). Our analysis aligns well with these concepts in that the need measure is precisely the SR as used in our model. The gain term refers to the amount of reward that can be gained by implementing a Bellman backup update to the policy. This requires that a particular reward function be specified (Sutton 1991). Given that our task, and the real-world experience of taxi drivers, corresponds to the multi-task RL setting with different starts and goals defining a huge number of possible reward functions, we suggest that the LTE measure provides an appropriate analogous measure agnostic to any particular reward function. Indeed, the conceptual usage of LTE in our model is consistent with a number of information-theoretic studies investigating efficient generic state-space representations for exploration and planning in artificial agents (Salge et al 2013; Solway et al 2014; Gregor et al 2016; Goyal et al 2019; Archer et al 2021). An important distinction with respect to meta-optimized models based on Q-DYNA and tree search (Mattar and Daw 2018; van Opheusden, 2023), is the apparent global non-sequential nature of the prioritization effect we observe. In our data, there was no effect of planning step number. Therefore, states were not prioritized in the order they could be sequentially generated from the origin but out-of-order with respect to the theoretic variables of SR and LTE. We suggest that such a global re-organization of planning from local simulations to non-sequential prioritization of critical states throughout the London state-space may reflect extremely refined planning strategies for expert planners. This predicts alternative forms of neural replay which may be testable in data and motivates the investigation of possible generative models of planning which exhibit such prioritization dynamics (McNamee et al 2021).

One of the environmental metrics that has a significant effect on response times is segment length, i.e. the length of the current segment at the state-level. This effect means that taxi drivers are slower at transitioning through physically longer segments. A potential explanation for this effect is that taxi drivers employ some form of veridical simulation of the route based on an internal model of the physical properties of London (Battaglia et al 2013; McNamee and Wolpert 2019). This is consistent with existing work suggesting that taxi drivers simulate in-street views of the trajectory when computing route plans (Griesbauer 2021). Further details of our analyses support this view. First, as the destination becomes closer in physical space, which would reduce the amount of internal simulation required, it was observed that response times get faster. Second, this effect was stronger than the insignificant effect of step

number, a graph-based measure of proximity to the destination without a spatial embedding, on response times. Third, the remaining direct Euclidean distance to the destination ("as the crow flies") did not have a significant effect on response times suggesting that the route itself was being generated internally across the path distance. Notably, this path distance to target, rather than Euclidean distance, is correlated with hippocampal activity (Howard et al., 2014; Spiers et al., 2018), consistent with the hippocampal simulation of future paths (Widloski and Foster, 2022; Olafsdottir et al., 2015; Pezzulo et al. 2017).

Further evidence supporting the hypothesis that taxi drivers planned using internal models which reflected the embodied spatial structure of London comes from the fact that destination angular deviation (i.e. angular deviation with respect to the final destination) also has an effect on response times. We can see this metric as a measure of the curvature of the route within a real-world spatial embedding as opposed to a pure graph-based representation of the state-space which lacks angular relationships. Within this framework, the slowing down of response times as a function of this variable is consistent with vector-based navigation (Le Moël et al., 2019). Using deep reinforcement learning, Banino and colleagues provided support for the idea that entorhinal grid cells contribute the spatial metric enabling the calculation of vectors between the current location and the goal (Banino et al., 2018). Another recent computational approach employed a hybrid model that combined a map-based strategy reliant on boundary cells and place cells and a vector-based strategy reliant on grid-cells (Edvardsen et al., 2020; for a review, see Nyberg et al., 2022). Future work could investigate in more mechanistic detail regarding how the use of vector-based navigation, predictive map use, and hierarchical planning vary from task to task, from environment to environment, and from participant to participant.

In conclusion, we examined the step-by-step city-wide route planning of London taxi drivers, and analyzed which factors drove their response times using a mix of environmental, cognitive, and information-theoretic variables. This provided real-world evidence for the use of distinct computational resource-rational strategies supporting expert planning in large state-spaces. Furthermore, we found several convergent effects supporting the hypothesis that features of the spatial structure of London are internally represented and influence the process of planning.

Methods

Experimental protocol. Taxi drivers were presented with an origin-destination pair and then were required to verbally describe a route from the origin to the destination within London. These verbal descriptions of the routes were recorded and transcribed from the taxi drivers (Griesbauer 2022a). Response times were calculated as pauses between two consecutively named streets (see **Figure 1** highlighted in yellow). The response time to the first named street, which tended to be significantly longer, is referred to as the *offline planning time* while the subsequent response times are referred to as *(online) call-out times*. Response times and thus were log-transformed prior to linear regression analyses and presentation. Furthermore, these data was standardized via z-scoring on a per-participant basis.

Origin-destination combinations were selected such that the corresponding routes varied in their environmental and structural properties (e.g. Euclidean distance, path length, direction of travel, and complexity). The experimental data was collected in two phases. In the first phase, tasks consisted of 12 origin-destination pairs. In the second data collection phase, tasks consisted of 8 origin-destination pairs, including two origin-destination pairs from the original set to allow for a comparison across the two groups. To keep traffic conditions consistent, participants were asked to imagine they were carrying out the planning task at 11.00 AM on a typical Monday. The data from both experiments was treated as one

data set, because the Wilcoxon Signed-rank test indicated no group differences between the two sets of taxi drivers for tasks 7 and 8 for log-transformed (Mdn(S1) = -0.22, IQR = 1.18, Mdn(S2) = -0.16, IQR = 1.24, p = .252, r = .046) or z-standardised (Mdn(S1) = -0.29, IQR = 1.11, Mdn(S2) = -0.33, IQR = 1.24, p = .688, r = .016) call-out times.

Environmental measures of London spatial context. To test the impact of spatial features of the street network on route planning behaviour, *call-out times* were analysed in association with the following environmental metrics: direct (*Euclidean) distance, streetwise (Euclidean) distance, segment length, remaining direct (Euclidean) distance, remaining streetwise (Euclidean) distance, segment angular deviation, destination angular deviation*. *Direct distance* corresponds to the Euclidean distance from origin to destination as the crow flies, at the route level. *Streetwise distance* corresponds to the distance from origin to destination along the shortest path, at the route level. *Segment length* corresponds to the length of the current segment, at the state-level. *Remaining streetwise (direct) distance* corresponds to the remaining streetwise (direct) distance to the destination. *Segment angular deviation* corresponds to the angular deviation with respect to the next segment in a route. *Destination angular deviation* corresponds to the angular deviation with respect to the destination of the route. See Figure 2B for a schematic illustration of these environmental measures.

Normative measures of planning complexity.

We use trajectory entropy as an information-theoretic measure of planning complexity under a random search model (McNamee et al., 2016). This quantifies how much information must be generated on average in order to specify a route from an origin state to a destination state. The matrix $H(\tau)$ of trajectory entropies for all (origin, destination) pairs in London can be expressed as a matrix-vector product of the *successor representation (SR)* M and the *local transition entropy* (LTE) H (Ekroot and Cover 1993; Kafsi et al 2012):

$$H(\tau) = \sum_s \quad M(o,s)H(s) \qquad (1)$$

where o is the origin state. The *successor representation (SR)* is defined as:

$$M = (I - \lambda T)^{-1} \qquad (2)$$

where T is the transition matrix defining the probability of transitioning between states of the environment under a random walk (i.e. exploratory) policy. Specifically, $T(s,s')$ is the probability moving to state s' from state s. The SR measures how likely different states (in our case, street segments) are to be sampled under a given transition policy and a model of the environment (Dayan, 1993). To a state s, we associate the SR measure $M(o,s)$ where o is the origin state for a particular route task. *Local transition entropy (LTE)* is defined as

$$H(s) = -\sum_{s' \in S} \quad T(s,s') \, log \, T(s,s') \qquad (3)$$

In our analyses, the LTE associated with a reported London position is computed with respect to the transition structure at the previously reported position. That is, for a transition $s_i \rightarrow s_j$ then the LTE $H(s_i)$ is aligned with s_j since this call-out time for s_j is expected to reflect the complexity of the planning process which produced state s_j (see Figure 1). Since the overall complexity of planning may be expressed as a product of SR and LTE, we reasoned that these variables should determine the mental labour of planning and, consequently, guide how planning is efficiently structured in expert planners (Gershman et al 2015; Kool and Bokvinik 2018; Lieder and Griffiths 2019). For example, regarding the first hypothesis, we expect that call-out times scale with LTE. With regards to the second hypothesis, we suggest that expert planners prioritize high SR-LTE situations for processing. From a

normative perspective, we suggest that plans are processed at states in the order by which they reduce the overall uncertainty in the route trajectory (McNamee, Wolpert, and Lengyel, 2016). We refer to this as a *trajectory entropy minimization* strategy. For example, taxi drivers may establish an approximate transition plan for a critical junction such as Charing Cross before establishing the sub-route from the origin to Charing Cross.

Network structural motifs of theoretic variables.

The SR and LTE variables have graphical interpretation in terms of the hierarchical structure of the state-space (Figure 2A) (Stachenfeld et al., 2017). For example, from the graph-theoretic perspective, high SR-LTE states tend to form critical junction states (Figure 2A, bottom right) while high SR low-LTE correspond to highly connected that may be concatenated into sub-routes (Figure 2A, top right) as in the options framework in hierarchical reinforcement learning (RL) (Jinnai et al 2019; Botvkink et al 2009). The network structure interpretation of SR and LTE is related to their role as normative metrics of planning complexity. Indeed, the correspondence between the network structure motifs described by these theoretic variables and their normative computational implications has been studied previously (Solway et al 2014; Archer et al 2021). Taken together, these complementary perspectives suggest resource-rational principles by which planning may be re-organized from a localized serial search approach to a globalized multi-scale process with hierarchical dynamics guided by SR and LTE metrics in extremely large state-spaces (Figure 3B).

Statistical analysis. To test the effects of variables from the street network on route planning, we employed a multivariate linear regression model assuming fixed effects across taxi drivers. The dependent variable was the log-transformed response times from the "call out" of routes. There were thirteen independent variables of the model. These included the spatial metrics outlined above: *Euclidean distance, streetwise distance, segment length, remaining streetwise distance, segment angular deviation, destination angular deviation*. The independent variables in the model also included the following: *offline planning time, plan step number, number of planning steps, successor representation measurement* (SR), *local transition entropy (LTE)*, and *SR-LTE interaction*. *Plan step number* corresponded to the callout number within a run, at the state-level. *Number of planning steps* corresponded to the total number of callouts in a run, at the route-level.

References

Allen, K. R., Brändle, F., Botvinick, M. M., Fan, J., Gershman, S. J., Gopnik, A., Griffiths, T. L., Hartshorne, J., Hauser, T. U., Ho, M. K., Leeuw, J. de, Ma, W. J., Murayama, K., Nelson, J. D., Opheusden, B. van, Pouncy, H. T., Rafner, J., Rahwan, I., Rutledge, R., … Schulz, E. (2023). *Using Games to Understand the Mind*. https://doi.org/10.31234/osf.io/hbsvj

Archer, K., Volpi, N. C., Bröker, F., & Polani, D. (2022). *A space of goals: The cognitive geometry of informationally bounded agents* (arXiv:2111.03699). arXiv. https://doi.org/10.48550/arXiv.2111.03699

Arnold, A. E. G. F., Iaria, G., & Ekstrom, A. D. (2016). Mental simulation of routes during navigation involves adaptive temporal compression. *Cognition*, *157*, 14–23. https://doi.org/10.1016/j.cognition.2016.08.009

Badre, D., Kayser, A. S., & D'Esposito, M. (2010). Frontal Cortex and the Discovery of Abstract Action Rules. *Neuron*, *66*(2), 315–326. https://doi.org/10.1016/j.neuron.2010.03.025

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

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), Article 7705. https://doi.org/10.1038/s41586-018-0102-6

Bast, H., Delling, D., Goldberg, A., Müller-Hannemann, M., Pajor, T., Sanders, P., Wagner, D., & Werneck, R. F. (2016). Route Planning in Transportation Networks. In L. Kliemann & P. Sanders (Eds.), *Algorithm Engineering: Selected Results and Surveys* (pp. 19–80). Springer International Publishing. https://doi.org/10.1007/978-3-319-49487-6_2

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327–18332. https://doi.org/10.1073/pnas.1306572110

Bellman, R. (1957). *Dynamic programming*. Princeton University Press.

Bonasia, K., Blommesteyn, J., & Moscovitch, M. (2016). Memory and navigation: Compression of space varies with route length and turns. *Hippocampus*, *26*(1), 9–12. https://doi.org/10.1002/hipo.22539

Botvinick, M. M. (2012). Hierarchical reinforcement learning and decision making. *Current Opinion in Neurobiology*, *22*(6), 956–962. https://doi.org/10.1016/j.conb.2012.05.008

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

Büchner, S., Hölscher, C., Strube, G., Vosniadou, S., Kayser, D., & Protopapas, A. (2007). Path choice heuristics for navigation related to mental representations of a building. *Proceedings of the European Cognitive Science Conference*, 504–509.

Callaway, F., Lieder, F., Das, P., Gul, S., Krueger, P. M., & Griffiths, T. L. (2018). A resource-rational analysis of human planning. *Proceedings of the 40th Annual Meeting of the Cognitive Science Society, CogSci 2018*, 178–183. https://collaborate.princeton.edu/en/publications/a-resource-rational-analysis-of-human-planning

Calvino, I. (1978). *Invisible cities*. Houghton Mifflin Harcourt.

Correa, C. G., Ho, M. K., Callaway, F., Daw, N. D., & Griffiths, T. L. (2023). Humans decompose tasks by trading off utility and computational cost. *PLOS Computational Biology*, *19*(6), e1011087. https://doi.org/10.1371/journal.pcbi.1011087

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), Article 12. 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

de Cothi, W., Nyberg, N., Griesbauer, E.-M., Ghanamé, C., Zisch, F., Lefort, J. M., Fletcher, L., Newton, C., Renaudineau, S., Bendor, D., Grieves, R., Duvelle, É., Barry, C., & Spiers, H. J. (2022). Predictive maps in rats and humans for spatial navigation. *Current Biology*, *32*(17), 3676-3689.e5. https://doi.org/10.1016/j.cub.2022.06.090

Edvardsen, V., Bicanski, A., & Burgess, N. (2020). Navigating with grid and place cells in cluttered environments. *Hippocampus*, *30*(3), 220–232. https://doi.org/10.1002/hipo.23147

Ekroot, L., & Cover, T. M. (1993). The entropy of Markov trajectories. *IEEE Transactions on Information Theory*, *39*(4), 1418–1421. https://doi.org/10.1109/18.243461

Elliott, R. J., & Lesk, M. E. (1982). Route finding in street maps by computers and people. *Proceedings of the Second AAAI Conference on Artificial Intelligence*, 258–261.

Fernandez Velasco, P., & Spiers, H. J. (2024). Wayfinding across ocean and tundra: What traditional cultures teach us about navigation. *Trends in Cognitive Sciences*, *28*(1), 56–71.

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

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278. https://doi.org/10.1126/science.aac6076

Goyal, A., Islam, R., Strouse, D., Ahmed, Z., Botvinick, M., Larochelle, H., Bengio, Y., & Levine, S. (2023). *InfoBot: Transfer and Exploration via the Information Bottleneck* (arXiv:1901.10902). arXiv. http://arxiv.org/abs/1901.10902

Gregor, K., Rezende, D. J., & Wierstra, D. (2016). *Variational Intrinsic Control* (arXiv:1611.07507). arXiv. https://doi.org/10.48550/arXiv.1611.07507

Griesbauer, E., Manley, E., Wiener, J. M., & Spiers, H. J. (2022). London taxi drivers: A review of neurocognitive studies and an exploration of how they build their cognitive map of London. *Hippocampus*, *32*(1), 3–20. https://doi.org/10.1002/hipo.23395

Griesbauer, E.-M., Fernandez Velasco, P., Coutrot, A., Wiener, J. M., Morley, J. G., McNamee, D., Manley, E., & Spiers, H. J. (2024). London taxi drivers exploit neighbourhood boundaries for hierarchical route planning. *bioRxiv*, 2024–02.

Griesbauer, E.-M., Manley, E., McNamee, D., Morley, J., & Spiers, H. (2022). What determines a boundary for navigating a complex street network: Evidence from London taxi drivers. *The Journal of Navigation*, *75*(1), 15–34. https://doi.org/10.1017/S0373463321000679

Griffiths, T. L. (2020). Understanding Human Intelligence through Human Limitations. *Trends in Cognitive Sciences*, *24*(11), 873–883. https://doi.org/10.1016/j.tics.2020.09.001

Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., & Griffiths, T. L. (2022). People construct simplified mental representations to plan. *Nature*, *606*(7912), Article 7912. https://doi.org/10.1038/s41586-022-04743-9

Hommel, B., Gehrke, J., & Knuf, L. (2000). Hierarchical coding in the perception and memory of spatial layouts. *Psychological Research*, *64*(1), 1–10. https://doi.org/10.1007/s004260000032

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.

Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai Trees in Your Head: How the Pavlovian System Sculpts Goal-Directed Choices by Pruning Decision Trees. *PLOS Computational Biology*, *8*(3), e1002410. https://doi.org/10.1371/journal.pcbi.1002410

Huys, Q. J. M., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., Dayan, P., & Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences*, *112*(10), 3098–3103. https://doi.org/10.1073/pnas.1414219112

Jinnai, Y., Abel, D., Hershkowitz, D., Littman, M., & Konidaris, G. (2019). Finding Options that Minimize Planning Time. *Proceedings of the 36th International Conference on Machine Learning*, 3120–3129. https://proceedings.mlr.press/v97/jinnai19a.html

Jinnai, Y., Park, J. W., Abel, D., & Konidaris, G. (2019). Discovering Options for Exploration by Minimizing Cover Time. *Proceedings of the 36th International Conference on Machine Learning*, 3130–3139. https://proceedings.mlr.press/v97/jinnai19b.html

Kafsi, M., Grossglauser, M., & Thiran, P. (2013). The Entropy of Conditional Markov Trajectories. *IEEE Transactions on Information Theory*, *59*(9), 5577–5583. https://doi.org/10.1109/TIT.2013.2262497

Keramati, M., Smittenaar, P., Dolan, R. J., & Dayan, P. (2016). Adaptive integration of habits into depth-limited planning defines a habitual-goal–directed spectrum. *Proceedings of the National Academy of Sciences*, *113*(45), 12868–12873. https://doi.org/10.1073/pnas.1609094113

Knoblock, C. A. (1990). A Theory of Abstraction for Hierarchical Planning. In D. P. Benjamin (Ed.), *Change of Representation and Inductive Bias* (pp. 81–104). Springer US. https://doi.org/10.1007/978-1-4613-1523-0_5

Kool, W., & Botvinick, M. (2018). Mental labour. *Nature Human Behaviour*, *2*(12), Article 12. https://doi.org/10.1038/s41562-018-0401-9

Lashley, K. S. (1951). The problem of serial order in behavior. In *Cerebral mechanisms in behavior; the Hixon Symposium* (pp. 112–146). Wiley.

Le Moël, F., Stone, T., Lihoreau, M., Wystrach, A., & Webb, B. (2019). The Central Complex as a Potential Substrate for Vector Based Navigation. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00690

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, e1. https://doi.org/10.1017/S0140525X1900061X

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., 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., 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

Mattar, M. G., & Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nature Neuroscience*, *21*(11), Article 11. https://doi.org/10.1038/s41593-018-0232-z

McNamee, D. C., Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2021). Flexible modulation of sequence

generation in the entorhinal–hippocampal system. *Nature Neuroscience*, *24*(6), Article 6. https://doi.org/10.1038/s41593-021-00831-7

McNamee, D., & Wolpert, D. M. (2019). Internal Models in Biological Control. *Annual Review of Control, Robotics, and Autonomous Systems*, *2*(Volume 2, 2019), 339–364. https://doi.org/10.1146/annurev-control-060117-105206

McNamee, D., Wolpert, D. M., & Lengyel, M. (2016). Efficient state-space modularization for planning: Theory, behavioral and neural signatures. *Advances in Neural Information Processing Systems*, *29*. https://proceedings.neurips.cc/paper_files/paper/2016/hash/10907813b97e249163587e6246612e21-Abstract.html

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

Momennejad, I. (2024). *Memory, Space, and Planning: Multiscale Predictive Representations* (arXiv:2401.09491). arXiv. http://arxiv.org/abs/2401.09491

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.

Newell, A., & Simon, H. A. (1972). *Human problem solving* (pp. xiv, 920). Prentice-Hall.

Nyberg, N., Duvelle, É., Barry, C., & Spiers, H. J. (2022). Spatial goal coding in the hippocampal formation. *Neuron*, *110*(3), 394–422. https://doi.org/10.1016/j.neuron.2021.12.012

O'Doherty, J. P., Lee, S. W., & McNamee, D. (2015). The structure of reinforcement-learning mechanisms in the human brain. *Current Opinion in Behavioral Sciences*, *1*, 94–100. https://doi.org/10.1016/j.cobeha.2014.10.004

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.

Opheusden, B. van, Kuperwajs, I., Galbiati, G., Bnaya, Z., Li, Y., & Ma, W. J. (2023). Expertise increases planning depth in human gameplay. *Nature*, *618*(7967), Article 7967. https://doi.org/10.1038/s41586-023-06124-2

Pezzulo, G., Kemere, C., & Van Der Meer, M. A. A. (2017). Internally generated hippocampal sequences as a vantage point to probe future□oriented cognition. *Annals of the New York Academy of Sciences*, *1396*(1), 144–165. https://doi.org/10.1111/nyas.13329

Russek, E., Acosta-Kane, D., Opheusden, B. van, Mattar, M. G., & Griffiths, T. (2022). *Time spent thinking in online chess reflects the value of computation*. https://doi.org/10.31234/osf.io/8j9zx

Russell, S., & Norvig, P. (2021). Artificial Intelligence: A modern approach, 4th US ed. *University of California, Berkeley*.

Salge, C., Glackin, C., & Polani, D. (2013). *Empowerment—An Introduction* (arXiv:1310.1863). arXiv. https://doi.org/10.48550/arXiv.1310.1863

Schick, W. (2018). *Acquisition and consolidation of hierarchical representations of space*. https://doi.org/10.15496/publikation-23395

Schick, W., Halfmann, M., Hardiess, G., Hamm, F., & Mallot, H. A. (2019). Language cues in the formation of hierarchical representations of space. *Spatial Cognition & Computation*, *19*(3), 252–281. https://doi.org/10.1080/13875868.2019.1576692

Shannon, C. E. (1950). Programming a computer for playing chess. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *41*(314), 256–275. https://doi.org/10.1080/14786445008521796

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), Article 7587. https://doi.org/10.1038/nature16961

Snider, J., Lee, D., Poizner, H., & Gepshtein, S. (2015). Prospective Optimization with Limited Resources. *PLOS Computational Biology*, *11*(9), e1004501. https://doi.org/10.1371/journal.pcbi.1004501

Solway, A., Diuk, C., Córdova, N., Yee, D., Barto, A. G., Niv, Y., & Botvinick, M. M. (2014). Optimal Behavioral Hierarchy. *PLOS Computational Biology*, *10*(8), e1003779. https://doi.org/10.1371/journal.pcbi.1003779

Spiers, H. J., & Maguire, E. A. (2006a). Spontaneous mentalizing during an interactive real world task: An fMRI study. *Neuropsychologia*, *44*(10), 1674–1682.

Spiers, H. J., & Maguire, E. A. (2006b). Thoughts, behaviour, and brain dynamics during navigation in the real world. *Neuroimage*, *31*(4), 1826–1840.

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

Spiers, H. J., & Maguire, E. A. (2007b). Neural substrates of driving behaviour. *Neuroimage*, *36*(1), 245–255.

Spiers, H. J., & Maguire, E. A. (2007c). The neuroscience of remote spatial memory: A tale of two cities. *Neuroscience*, *149*(1), 7–27.

Spiers, H. J., & Maguire, E. A. (2008). The dynamic nature of cognition during wayfinding. *Journal of Environmental Psychology*, *28*(3), 232–249.

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

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

Streeter, L. A., & Vitello, D. (1986). A Profile of Drivers' Map-Reading Abilities. *Human Factors*, *28*(2), 223–239. https://doi.org/10.1177/001872088602800210

Sutton, R. S. (1991). Dyna, an integrated architecture for learning, planning, and reacting. *ACM SIGART Bulletin*, *2*(4), 160–163. https://doi.org/10.1145/122344.122377

Tomov, M. S., Yagati, S., Kumar, A., Yang, W., & Gershman, S. J. (2020). Discovery of hierarchical representations for efficient planning. *PLOS Computational Biology*, *16*(4), e1007594. https://doi.org/10.1371/journal.pcbi.1007594

van Opheusden, B., Galbiati, G., Kuperwajs, I., Bnaya, Z., Li, Y., & Ma, W. J. (2024). *Revealing the impact of expertise on human planning with a two-player board game*. https://doi.org/10.31234/osf.io/rhq5j

Ward, G., & Allport, A. (1997). Planning and Problem solving Using the Five disc Tower of London Task. *The Quarterly Journal of Experimental Psychology Section A*, *50*(1), 49–78. https://doi.org/10.1080/713755681

Widloski, J., & Foster, D. J. (2022a). Flexible rerouting of hippocampal replay sequences around changing barriers in the absence of global place field remapping. *Neuron*, *110*(9), 1547–1558.

Widloski, J., & Foster, D. J. (2022b). Flexible rerouting of hippocampal replay sequences around changing barriers in the absence of global place field remapping. *Neuron*, *110*(9), 1547–1558.

Wiener, J. M., Ehbauer, N. N., & Mallot, H. A. (2009). Planning paths to multiple targets: Memory involvement and planning heuristics in spatial problem solving. *Psychological Research PRPF*, *73*(5), 644–658. https://doi.org/10.1007/s00426-008-0181-3

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

Wiener, J. M., Schnee, A., & Mallot, H. A. (2004). Use and interaction of navigation strategies in regionalized environments. *Journal of Environmental Psychology*, *24*(4), 475–493. https://doi.org/10.1016/j.jenvp.2004.09.006