

Multi-goal spatial navigation is mediated by predictive representations with episodic replay in the human brain

Christoffer J. Gahnstrom (cgahn@sas.upenn.edu)

University of Pennsylvania, Department of Psychology
Philadelphia, Pennsylvania, USA

Russell A. Epstein (epstein@psych.upenn.edu)

University of Pennsylvania, Department of Psychology
Philadelphia, Pennsylvania, USA

Abstract

What are the neural and computational mechanisms underlying human spatial navigation? Previous studies have suggested that reward prediction and replay might underlie key navigational components such as credit assignment, memory consolidation, and planning. However, these mechanisms are usually tested with relatively simple paradigms, making it unclear what role they might play in ecologically realistic navigational tasks involving rapidly changing goal locations. To investigate this issue, we scanned participants (N=15) with fMRI while they performed a “taxi-cab” task in a virtual city with multiple possible goals. We found that a successor representation model incorporating episodic replay (SR-DYNA) best fit the observed human behavior. To identify the possible neural systems underlying SR-DYNA, we analyzed BOLD activity in terms of several components of the model. We observed parametric tracking of successor state values in anterior hippocampus, parametric tracking of successor prediction error in a network of cortical regions previously implicated in visuospatial memory, and evidence for remote context-dependent episodic replay in the posterior hippocampus. Our results provide behavioral and neural evidence for predictive representations imbued with episodic reactivations as a plausible mechanism of human flexible navigation.

Keywords: spatial navigation; reinforcement learning; successor representation; decision making; computational modeling

Introduction

People are able to efficiently navigate through large and complex spatial environments encompassing many different goals and goal locations. For instance, upon arrival to a new city, one might explore the immediate surroundings, encounter several relevant locations, such as museums, cafes, and restaurants, store these locations in memory, and then navigate to each one over the course of the following days. How do we accomplish this feat?

Reinforcement learning (RL) models have been proposed as possible algorithms for navigation (Foster, Morris, & Dayan, 2000; Stachenfeld, Botvinick, & Gershman, 2017; De Cothi et al., 2022; He, Liu, Eschapaspe, Beveridge, & Brown, 2022; Simon & Daw, 2011; Anggraini, Glasauer, & Wunderlich, 2018). However, we do not have a good understanding of how these algorithms might apply to situations involving interleaved navigation to multiple goals. In this paper, we address this lacuna by analyzing behavioral and fMRI data obtained from human subjects navigating through a virtual city with multiple goal locations.

Our hypothesis was that behavior and neural signals would reveal evidence that participants use the successor representation (SR) algorithm to navigate (Gershman, 2018). The SR is an extension of temporal-difference learning that has been 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 . M is updated using a process similar to temporal difference updating, but using successor state prediction error instead of reward prediction error (Russek, Momennejad, Botvinick, Gershman, & Daw, 2017). We implemented both standard SR, and also a version in which previously experienced episodes of successful navigation to the current goal were replayed at the onset of new trials (SR-DYNA).

To simplify the computations required, we abstracted the physical space into a graph-like representation of places and connections (Fig. 1), and to accommodate the existence of multiple goals, we incorporated a cue signal corresponding to the current goal of the participants.

Experimental Procedure

Fifteen healthy individuals participated in 7 or 8 experimental sessions over the course of 2 weeks. Participants performed a taxi driver task, in which they alternated between randomly exploring the environment to find a passenger and taking the passenger to two cued goals locations in sequence. The first goal was cued upon pickup of the passenger, and the second goal was cued upon arrival at the first goal location. Following a training phase, participants underwent up to 6 additional sessions of the taxi driver task within an fMRI scanner, each comprising four to six scan runs of approximately 11 minutes each.

Results

Human performance most closely matches SR-DYNA

As described in our previous report (CCN 2023), we compared model performance to human performance in sequences of goal-directed trials for which start and goal locations were exactly matched (Fig. 1). Human performance was not significantly different from SR-DYNA using 1 or 2 replay episodes per trial, respectively (SR-DYNA1 $t(14) = 0.69$, $p = 0.49$ and SR-DYNA2 $t(14) = -1.43$, $p = 0.16$). In contrast, human performance was significantly better than a model-free algorithm (Q-learner $t(14) = -13.8$, $p = 5.15e-14$) and SR without replay ($t(14) = -4.27$, $p = 0.0002$). Lastly, humans were significantly worse than a model-based valuation agent ($t(14) = 5.54$, $p = 6.36e-6$). Thus, an SR model incorporating replay was the best match to the human performance level.

Neural correlates of predictive state values, successor prediction errors, and episodic replay

The behavioral analyses suggest that humans use predictive representations and replay to perform the task, consistent with other behavioral work (Momennejad et al., 2017). We thus proceeded to test for three sources of neural evidence for the use of these mechanisms..

First, we looked for fMRI signals related to the current SR state value, by testing a GLM with a parametric regressor for

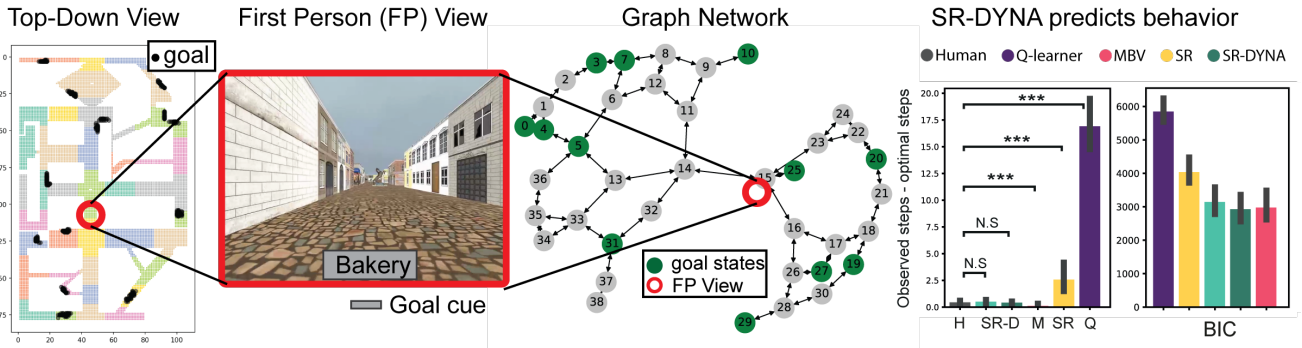


Figure 1: Task environment and behavioral performance. Left: Map and first-person view of the virtual city, downsampled into a graph network of 39 discrete states defined by behaviorally salient regions (hallways and junctions). Right: Simulations of SR-DYNA and other candidate models shows that SR-DYNA is the best match to human performance, and model fitting procedure using maximum likelihood estimation identified SR-DYNA as having the best fit (lowest bayesian information criterion score).

the current state value after each state transition. We observed such a BOLD response in the bilateral hippocampus, consistent with previous work implicating the hippocampus in the formation of predictive state relationships (Schapiro, Turk-Browne, Norman, & Botvinick, 2016) (Figure 2 Left).

Second, we looked for fMRI signals related to the magnitude of successor prediction errors (schematic in figure 2 Middle). These prediction errors varied substantially over the course of the experiment, given that the current policy changed on each trial as a different goal location was queued. This analysis revealed a network of brain regions whose BOLD activity parametrically varied with the successor prediction errors (Figure 2 Middle), including retrosplenial complex, occipital place area, parahippocampal place area, and superior parietal cortex. Notably, this network has been previously implicated in visuospatial memory (Steel et al. 2023).

Third, we looked for fMRI signals related to replay. The replay (DYNA) component of the SR model was based on episodic reactivation of recent experiences (Momennejad,

2020); specifically reactivation at the beginning of each trial of the most recent previous trial to the same goal. We modeled this as the total successor prediction error for the putative reactivation. We found evidence that posterior right hippocampal BOLD activity scaled with the magnitude of replay-induced prediction errors (Figure 2 Right).

Conclusion

A crucial but understudied feature of human navigation is the ability to learn multiple goal locations and switch between them in a flexible manner. We found behavioral and fMRI evidence for hypothesized neurocomputational mechanisms by fitting reinforcement learning models to behavior and BOLD responses. Our results suggest that predictive representations with episodic reactivation are plausible algorithms underlying flexible human navigation. Future work may also investigate normative replay mechanisms and their ability to explain behavior and neural activity during multi-goal navigation (Mattar & Daw, 2018).

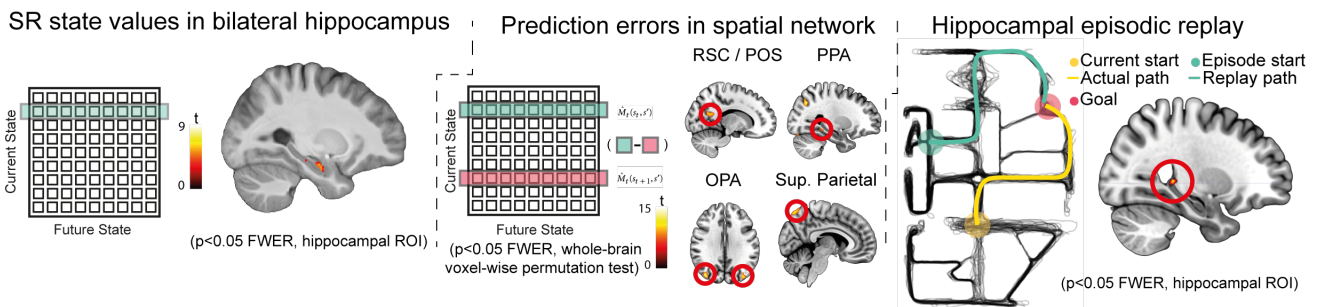


Figure 2: SR-DYNA internal computations are reflected in fMRI responses. Left: SR matrix with one row highlighted in green representing the successor values for one state. The SR-DYNA model was used to extract SR values following each state transition. BOLD activity parametrically modulated with these values in the anterior hippocampus. Middle: The SR model computes prediction errors after each state transition which are used to update the successor matrix. We observed BOLD activity that parametrically modulated with these errors in a spatial memory network (Steel et al., 2023) encompassing RSC, PPA, OPA, and superior parietal cortex. Right: Each replay episode generated a series of prediction errors. We observed BOLD activity that parametrically varied with the total error of the series in the right posterior hippocampus.

Acknowledgments

We thank Zhengang Lu for sharing the behavioral dataset and helpful discussions. We also thank William De Cothi for comments on the reinforcement learning approaches and Sarah Goodroe for comments on the abstract. This work was supported by NIH Grant EY022350 to RAE.

References

- Anggraini, D., Glasauer, S., & Wunderlich, K. (2018). Neural signatures of reinforcement learning correlate with strategy adoption during spatial navigation. *Scientific reports*, *8*(1), 10110.
- Dayan, P. (1993). Improving generalization for temporal difference learning: The successor representation. *Neural computation*, *5*(4), 613–624.
- De Cothi, W., Nyberg, N., Griesbauer, E.-M., Ghanamé, C., Zisch, F., Lefort, J. M., ... others (2022). Predictive maps in rats and humans for spatial navigation. *Current Biology*, *32*(17), 3676–3689.
- Foster, D. J., Morris, R. G., & Dayan, P. (2000). A model of hippocampally dependent navigation, using the temporal difference learning rule. *Hippocampus*, *10*(1), 1–16.
- Gershman, S. J. (2018). The successor representation: its computational logic and neural substrates. *Journal of Neuroscience*, *38*(33), 7193–7200.
- He, Q., Liu, J. L., Eschapaspe, L., Beveridge, E. H., & Brown, T. I. (2022). A comparison of reinforcement learning models of human spatial navigation. *Scientific Reports*, *12*(1), 13923.
- Mattar, M. G., & Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nature neuroscience*, *21*(11), 1609–1617.
- Momennejad, I. (2020). Learning structures: predictive representations, replay, and generalization. *Current Opinion in Behavioral Sciences*, *32*, 155–166.
- 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.
- 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.
- Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., & Botvinick, M. M. (2016). Statistical learning of temporal community structure in the hippocampus. *Hippocampus*, *26*(1), 3–8.
- 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.
- Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature neuroscience*, *20*(11), 1643–1653.
- Steel, A., Garcia, B. D., Goyal, K., Mynick, A., & Robertson, C. E. (2023). Scene perception and visuospatial memory converge at the anterior edge of visually responsive cortex. *Journal of Neuroscience*, *43*(31), 5723–5737.