Use of Vector- and Transition-based Strategies is Modulated by Knowledge of the Environment in Human Spatial Planning

Denis C. L. Lan (denis.lan@psy.ox.ac.uk)
Department of Experimental Psychology, University of Oxford
Oxford, United Kingdom

Laurence T. Hunt (laurence.hunt@psy.ox.ac.uk) *
Department of Experimental Psychology, University of Oxford
Oxford, United Kingdom

Christopher Summerfield (christopher.summerfield@psy.ox.ac.uk) *
Department of Experimental Psychology, University of Oxford
Oxford, United Kingdom

(* **Contributed equally as senior authors)**

Abstract:

Planning and navigation may depend on 'transition-based' strategies that focus on connectivity between states and 'vector-based' strategies that focus on their relative spatial locations. We hypothesised that relative use of these two strategies might depend on our knowledge of the environment, such as the number of known 'landmarks'. We developed a novel behavioural paradigm to dissociate use of the two strategies. We find that participants use transition-based strategies to navigate to well-learned 'landmark' locations, especially when there are fewer such locations. Moreover, we find that performance is best at an intermediate number of 'landmarks', suggesting a trade-off between coverage of the environment and the strength of memory. A computational model with memory constraints that considers its uncertainty during navigation mimics these behavioural patterns.
Keywords: planning, navigation, cognitive map, vector, transitions

Introduction

Humans and other animals are skilled at flexibly planning routes, leading to suggestions that we possess a 'cognitive map' that affords us flexibility in navigation (Tolman, 1948). Researchers have proposed several strategies by which we might navigate such a map, including transition-based strategies, such as successor representation or model-based reinforcement learning, that learn from experienced state transitions (de Cothi et al., 2022; Momennejad et al., 2017) and vector-based strategies that focus on the geometry and relative spatial locations of states in the environment (Banino et al., 2018; Bush et al., 2015).

However, we do not understand how people might combine the two strategies. Given that transition-based (but not vector-based) strategies require knowledge of associations between individual states, we might rely on vector-based strategies in unfamiliar regions to head in the general goal direction, and transition-based strategies in well-learned locations to fine-tune our navigation. This implies that the strategies have complementary costs and benefits that depend on the sparsity of our knowledge on the environment and memory constraints. Knowing more 'landmarks' might facilitate more accurate navigation, but encoding a few landmarks strongly might be more beneficial than encoding many landmarks weakly.

In this study, we develop a novel paradigm that allows us to dissociate the use of these two strategies and test these predictions. We predict that (1) people are more prone to using transition-based strategies near well-learned 'landmarks' and (2) there is some intermediate number of landmarks that is optimal for performance, as it balances environmental coverage with memory constraints.

Method

40 participants performed an online experiment in which they navigated through an 8x8 grid of 'objects'. Participants completed 4 blocks, each consisting of 6 trials. In the learning phase of each trial, participants viewed a grid with its constituent objects obscured (Figure 1A). They were required to click on a sequence of grid squares, highlighted blue, successively revealing the 'landmark' objects that were at the corresponding locations. There were a different number of landmarks in each block (2, 4, 8, or 16, in random order) but the number of total presentations was kept the same (2 landmarks would be shown 8 times, 16 landmarks would be shown once each, etc). After clicking on the blue squares, participants were required to click on a yellow grid square to reveal the 'goal' object location for the upcoming trial. Every trial involved different objects and landmarks at different locations.

In the 'test' phase of each trial, participants started in a random location on the grid that they had not experienced during the 'learning' phase. They were required to navigate to the 'goal' object (always 4 steps away from the starting location (Figure 1B)). The object corresponding to the current state was displayed in the centre of the screen. Crucially, participants could navigate the grid in two ways. They could either click on arrows located on one side of the screen, which took them one step in the corresponding direction (arrow strategy). This involves choosing an action without knowing the subsequent state, as in vector-based navigation. Alternatively, they could click one of the adjacent objects (displayed in random order), to move them to the grid square corresponding to this object. This involves choosing a state transition without knowing the action that was taken to perform that transition (object strategy). Every step they took cost 50 points and reaching the goal state earned participants 1000 points.

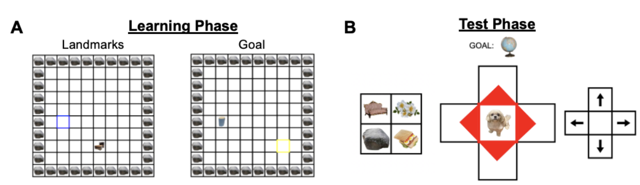

Figure 1: A: In the learning phase, participants clicked on blue and yellow squares to reveal the landmarks and goal respectively. B: in the test phase, participants navigated to the goal using either the objects (left) or arrows (right). After making a step, red triangles animated to show the direction participants move in.

Results

Despite being trained on just a subset of the object locations, participants achieved an average choice accuracy of 80.76% (SD = 15.15%, chance differs based on location in the grid but averaged 47.97%). While participants had a strong bias towards using the arrow strategy, we found that the response strategy differed depending on the type of destination state (Figure 2A): participants were significantly more likely to use the object on the step that took them to the goal (generalised logistic mixed effects model: $\beta = 5.17$, $p < .001$) or a landmark ($\beta = 4.91$, $p < .001$). This was moderated by the number of landmarks (object type x number of landmarks interaction: $\beta = -0.17$, $p < .001$), with participants more likely to use the object strategy when there were fewer landmarks.

Performance (steps taken to goal) was also modulated by the number of landmarks (ANOVA: $F(3, 117) = 4.53$, $p = .005$), with performance being best at 8 landmarks (Figure 2B). This was likely due to the trade-off between coverage of the environment and familiarity with individual landmarks: when fewer (strongly encoded) landmarks were shown during learning, each landmark was more helpful for navigation (higher likelihood of making correct step after encountering landmark; Figure 2C; logistic mixed effects model: $\beta = -0.07$, $p < .001$) but rarely encountered, but when there were more (weakly encoded) landmarks, they were often encountered but less helpful due to uncertainty about their position. Eight landmarks offered a good trade-off between these extremes.

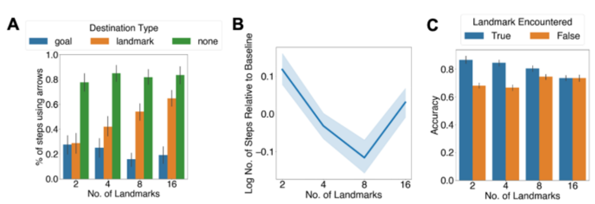

Figure 2: A: percentage of steps made using arrows as a function of destination type and number of landmarks. B: log steps to goal (relative to average across conditions) as a function of the number of landmarks. C: accuracy as a function of number of landmarks and whether a landmark has been encountered. All error bars represent +/- 1 SE.

Lastly, to better understand these results, we constructed a computational simulation of participant performance. During learning, our model used a multilayer perceptron (MLP) trained with dropout (i.e., hidden units deactivated with $p = 0.2$) to output x- and y-coordinates given a one-hot encoded state as an input. The model was trained on the same landmarks as human participants, and the number of total training iterations was kept the same for all numbers of landmarks. For each input state, we sampled 100 outputs with dropout and took the variance of these x and y estimates as a proxy for the model's confidence (Gal & Ghahramani, 2016). During navigation, the model used (1) its estimate of its current location from the MLP to compute the cosine similarity between the estimated vector to goal and the four cardinal directions (for the vector-based component) and (2) its estimate of the location of the adjacent objects to estimate the distance to goal of the four adjacent objects (for the transition-based component). It then weights these values with its confidence in its current location (for the vector-based component) or the confidence in the location of the objects (for the transition-based component) and uses these weighted values to choose which strategy and action to take.

Our model approximates participants' main behavioural patterns: it uses the objects more often at landmarks (Figure 3A) and performs best at an intermediate number of landmarks (Figure 3B). We only observe the former pattern when the model uses its confidence in choosing the two strategies (Figure 3C), suggesting the importance of uncertainty in determining the balance of these two strategies, and the latter pattern when the MLP has twice as many hidden units (Figure 3D), suggesting the importance of memory constraints in determining the U-shaped relationship between performance and number of landmarks.

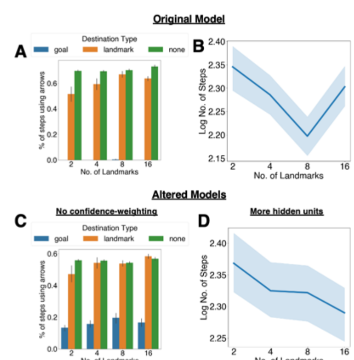

Figure 3: Same as Figure 2A but for A: our computational model and C: our model without confidence-weighting. Same as Figure 2D but for B: our model and D: our model with twice as many hidden units

Discussion

Overall, we have shown that participants are more prone to using transition-based strategies at well-learnt locations. Moreover, we observe a U-shaped relationship between the number of landmarks and steps-to-goal, likely due to our memory constraints in remembering the locations of states.

Acknowledgments

D.L. is supported by the Clarendon Fund and the Christopher Welch Trust. L.H. is supported by the Wellcome Trust and Royal Society. C.S. is supported by an ERC Consolidator Award.

References

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

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

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

Gal, Y., & Ghahramani, Z. (2016). *Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning* (arXiv:1506.02142). arXiv. https://doi.org/10.48550/arXiv.1506.02142

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), Article 9. https://doi.org/10.1038/s41562-017-0180-8

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