Flexible Hippocampal Representation of Abstract Boundaries Supports Memory-guided Choice

Mariachiara Esposito* (mesposit@uji.es)
Universitat Jaume I, Department of Basic Psychology, Av. Vicent Sos Baynat
Castellón de la Plana, Castellón 12071 Spain

Lubna Abdul* (abdul@uji.es)
Universitat Jaume I, Department of Basic Psychology, Av. Vicent Sos Baynat
Castellón de la Plana, Castellón

Ameeruddin Ghouse (ghouse@uji.es)
Universitat Jaume I, Department of Basic Psychology, Av. Vicent Sos Baynat
Castellón de la Plana, Castellón

Marta Rodriguez Aramendía (aramendi@uji.es)
Universitat Jaume I, Department of Basic Psychology, Av. Vicent Sos Baynat
Castellón de la Plana, Castellón

Raphael Kaplan (kaplan@uji.es)
Universitat Jaume I, Department of Basic Psychology, Av. Vicent Sos Baynatt
Castellón de la Plana, Castellón

* Contributed equally

Abstract

Cognitive maps represent the location of different spatial cues relative to environmental boundaries. Map-like neural representations in the hippocampus are particularly sensitive to boundary changes, which highlights how cognitive maps can facilitate flexible learning in dynamically changing environments. Notably, cognitive maps in the hippocampus and medial prefrontal cortex (mPFC) also represent abstract knowledge. Yet it's unclear whether map-like knowledge representations in these regions are sensitive to boundary changes in abstract spaces. Here, we use a memory-guided decision-making task to test whether the human hippocampus and mPFC flexibly learn abstract boundary changes in distinct two-dimensional(2D) decision spaces. Despite being unnecessary to accurately make decisions, participants conserve a 2D map-like representation of abstract boundaries after the task, where the precision of their representation relates to prior choice accuracy. Additionally, we find that mPFC and hippocampus represent the euclidean distance to the relevant boundary during decision-making. Testing whether there are brain regions sensitive to changes to the limits of the decision space, we observe flexible hippocampal representation of abstract boundaries that relates to choice accuracy. Taken together, these results highlight the importance of hippocampal boundary representations in facilitating flexible knowledge retrieval across diverse spatial and abstract contexts.

Keywords: cognitive maps; learning; memory; hippocampus

Introduction

Environmental boundaries are known to exert a notable influence on spatial cognition across species (Poulter, Hartley, & Lever, 2018). In particular, the hippocampal formation dynamically represents the geometrical features of an environment, which facilitates flexible contextual processing (Lever, Burton, Jeewajee, O'Keefe, & Burgess, 2009; Wills, Lever, Cacucci, Burgess, & O'Keefe, 2005) in the service of spatial memory formation (Bellmund et al., 2020; Julian & Doeller, 2021; Steemers et al., 2016). Notably, hippocampal cognitive maps are thought to extend more generally to episodic memory (O'Keefe & Dostrovsky, 1971), putatively serving as a format for flexibly representing abstract knowledge (Schiller et al., 2015; Kaplan, Schuck, & Doeller, 2017; Whittington, McCaffary, Bakermans, & Behrens, 2022) The theoretical importance of cognitive maps in flexibly coding abstract knowledge raises the possibility that learning abstract boundaries supports map-like knowledge representations.

Addressing whether learning abstract boundaries is processed by the human hippocampal-prefrontal circuit in a similar way as physical borders, we developed a two alternative forced-choice(2AFC) fMRI task where participants made similarity judgements on either price or freshness levels of various fruits and vegetables (Figure 1A). Unbeknownst to participants, the continuous price and freshness levels for each produce good formed two-dimensional(2D) abstract spaces, with four of the goods placed at the extreme coordinates(boundary goods), one good fixed near the center(landmark good), and sixteen cued goods that were located within the boundaries. In the 2AFC, participants saw a produce good and decided whether it was more similar in price or freshness to either the landmark good or the most proximal boundary good(in 2D Euclidean distance). Crucially, the boundary produce goods had two distinct sets of coordinates depending on which run they were featured, where each run was defined by one of two distinct abstract spaces(square shape and distorted shape). In contrast, the landmark coordinates were consistent in the two spaces(Figure 1A, bottom left).
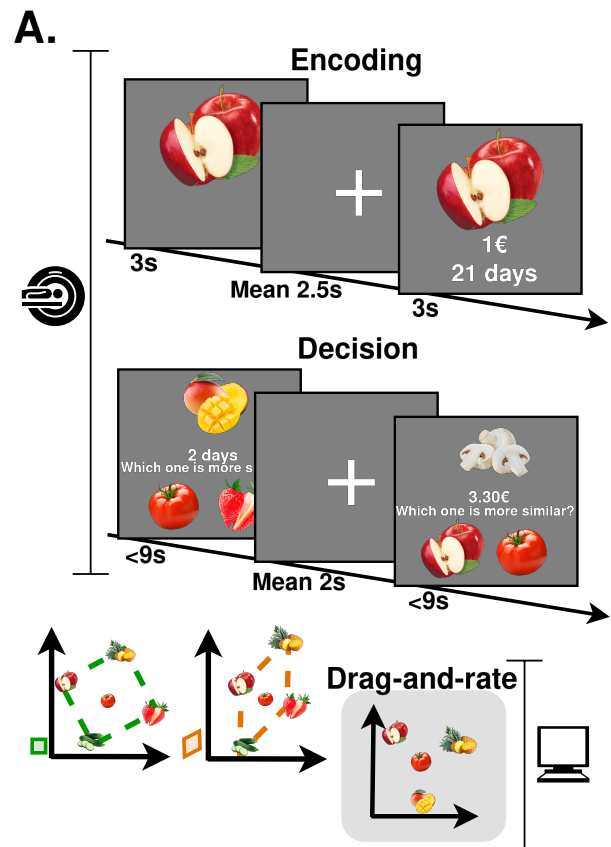

Figure 1: A. Encoding phase: Encoding of boundary and landmark goods attributes. Decision phase: In a two alternative forced-choice task, participants needed to choose whether a given good is "more similar" in price or freshness to a boundary or to the landmark good. Bottom Left: Coordinates of boundary goods in the two abstract spaces forming square and distorted shapes. Bottom Right: Post-task drag-and-rate. Participants were presented with a blank map and asked to place the goods seen during the task according to the position they most strongly associated with that particular good.

Results and Discussion

Using the drag-and-rate after scanning (Kriegeskorte & Mur, 2012), we tested whether participants retained a 2D representation of one of the decision spaces. Participants consistently organized boundary goods in the correct order(each boundary good shared a side with its neighboring boundary good), surpassing chance expectations(p< .001). Conducting a Monte Carlo simulation to determine the likelihood of all cued goods falling within the shapes, we found that they were positioned within the boundaries of the abstract space significantly above chance level(p< .001). This outcome highlights that the boundaries delimited the placement of observed products within the abstract space. Moreover, post-task fMRI debriefing indicated that none of the participants were aware that the goods formed any kind of 2D space. Next, we asked whether the precision during the placement was related to performance during the 2AFC. We then correlated the Procrustes distance from each original space with performance on task, and found when participants better reconstructed the original distorted shape, they consequently achieved greater task performance(ρ=-0.42, p=.026). However, this relationship was not found when participants needed to reconstruct the square shape (ρ=-0.08, p=.70).

We investigated whether the similarity of neural representations of the goods in the abstract spaces were modulated by the 2D Euclidean distance to the landmark good and the most proximal boundary good. We observed a significant effect of 2D Euclidean distance to the closest boundary in our bilateral hippocampus (t(25)=2.60, p=.015) and mPFC masks(t(25)=2.39, p=.024), suggesting that sharing similar Euclidean distance from the boundary is a significant predictor of neural similarity across products. Conversely, we didn't observe any significant effect of Euclidean distance of the products to the landmark good in mPFC(t(25) = 0.62, p=.54) or the hippocampus(t(25)= 0.50, p=.62).

Asking whether the hippocampus and mPFC flexibly represented abstract spaces consisting of the same goods with different boundary coordinates(price and freshness values), we tested whether boundary-defined contextual identity(the shape of abstract spaces) was decodable from hippocampal and mPFC signals. We conducted multivariate pattern analysis(MVPA) on encoding phase boundary good trials to test whether the patterns elicited during encoding trials in the square shape were distinguishable from trials in the distorted shape(i.e., same visual stimuli in both abstract spaces). We used a linear support vector machine(SVM), and applied a leave-one-subject-out cross-validation procedure to test whether the classifier could decode the shape of the abstract space above chance. In other words, we used hippocampal and mPFC fMRI data to classify whether an encoding trial for a boundary good belonged to a run that was either in a square or distorted shaped context. The classifier yielded significant results in the hippocampus(p=.02, accuracy:60%), but not in the mPFC(p=.29, accuracy:53%). We tested the abstract shape classifier on the landmark good

trials and observed no significant effect in the hippocampus(p=.47, accuracy:51%), suggesting that the hippocampal classification effect was specific to the representation of boundary goods. Classifier accuracy didn't correlate with general task performance(ρ=0.22, p=.30). Relating the difference in task performance in the distorted versus square shape to the hippocampal classifier bias, we observed a significant correlation with classifier accuracy(ρ=0.47,p=.022). This result means that classifier accuracy was highest in the hippocampus for participants that performed better in the distorted shape compared to the square shape. Taken together, our data provide important clues on how the hippocampus can guide decision making across diverse spatial and abstract contexts.
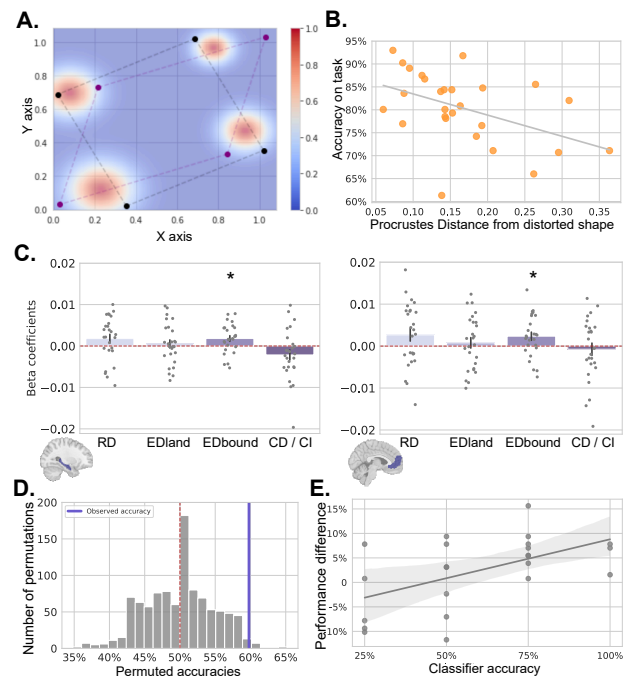

Figure 2: A. Participants' average placement of boundary goods in the drag-and-rate. B. Correlation between choice accuracy and drag-and-rate task precision. Plot displays correlation between Procrustes distance of 29 participants' reconstruction from the original distorted shape with overall accuracy on task. C. Beta coefficients of RSA GLM in hippocampus and mPFC. The bar plots show beta coefficients resulting from the GLM for each predictor; every participant (N=26) is an individual dot. Error bars represent SEM. Left: Significant bilateral hippocampal effect of euclidean distance to the closest boundary; Right: Significant mPFC effect of euclidean distance to the closest boundary. D.Hippocampal classifier results after one thousand permutations with mean accuracy 50% (chance level, indicated by red dashed line). Bold purple line indicates observed hippocampal classifier accuracy. E. Correlation between difference in task performance and hippocampal classifier accuracy (N=24). Positive differences indicate higher performance in distorted versus square shape. Asterisks showing significant differences: *= p< .05.

Acknowledgments

References

Bellmund, J. L., De Cothi, W., Ruiter, T. A., Nau, M., Barry, C., & Doeller, C. F. (2020). Deforming the metric of cognitive maps distorts memory. *Nature Human Behaviour*, *4*(2), 177–188.

Julian, J. B., & Doeller, C. F. (2021). Remapping and realignment in the human hippocampal formation predict context-dependent spatial behavior. *Nature Neuroscience*, *24*(6), 863–868.

Kaplan, R., Schuck, N. W., & Doeller, C. F. (2017). The role of mental maps in decision-making. *Trends in Neurosciences*, *40*(5), 256–259.

Kriegeskorte, N., & Mur, M. (2012). Inverse mds: Inferring dissimilarity structure from multiple item arrangements. *Frontiers in psychology*, *3*, 28167.

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.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*.

Poulter, S., Hartley, T., & Lever, C. (2018). The neurobiology of mammalian navigation. *Current Biology*, *28*(17), R1023–R1042.

Schiller, D., Eichenbaum, H., Buffalo, E. A., Davachi, L., Foster, D. J., Leutgeb, S., & Ranganath, C. (2015). Memory and space: Towards an understanding of the cognitive map. *Journal of Neuroscience*, *35*(41), 13904–13911.

Steemers, B., Vicente-Grabovetsky, A., Barry, C., Smulders, P., Schröder, T. N., Burgess, N., & Doeller, C. F. (2016). Hippocampal attractor dynamics predict memory-based decision making. *Current Biology*, *26*(13), 1750–1757.

Whittington, J. C., McCaffary, D., Bakermans, J. J., & Behrens, T. E. (2022). How to build a cognitive map. *Nature neuroscience*, *25*(10), 1257–1272.

Wills, T. J., Lever, C., Cacucci, F., Burgess, N., & O'Keefe, J. (2005). Attractor dynamics in the hippocampal representation of the local environment. *Science*, *308*(5723), 873–876.