High-fidelity social learning via shared episodic memories can improve collaborative foraging

Ismael T. Freire
Donders Institute for Brain, Cognition and Behaviour - Centre for Neuroscience (DCN-FNWI)
Radboud University, Nijmegen, the Netherlands
ismael.freire@donders.ru.nl

Paul F.M.J. Verschure
Donders Institute for Brain, Cognition and Behaviour - Centre for Neuroscience (DCN-FNWI)
Radboud University, Nijmegen, the Netherlands
paul.verschure@donders.ru.nl

Abstract

Social learning, a cornerstone of cultural evolution, allows individuals to acquire knowledge by observing and imitating others. Central to its efficacy is episodic memory, which records specific behavioral sequences to facilitate learning. This study examines the interrelation between social learning and episodic memory in the context of collaborative foraging. Specifically, we examine how variations in the frequency and fidelity of social learning impact collaborative foraging, and how the length of behavioral sequences preserved in agents' episodic memory modulates these factors. To this end, we deploy Sequential Episodic Control agents capable of sharing among them behavioral sequences stored in their episodic memories. Our findings indicate that high-frequency, high-fidelity social learning promotes more distributed and efficient resource collection, a benefit that remains consistent regardless of the length of the shared episodic memories. In contrast, low-fidelity social learning shows no advantages over non-social learning in terms of resource acquisition. In addition, storing and disseminating longer episodic memories contribute to enhanced performance up to a certain threshold, beyond which increased memory capacity does not yield further benefits. Our findings emphasize the crucial role of high-fidelity social learning in collaborative foraging, and illuminate the intricate relationship between episodic memory capacity and the quality and frequency of social learning. This work aims to highlight the potential of neuro-computational models like episodic control algorithms in understanding social learning and offers a new perspective for investigating the cognitive mechanisms underlying open-ended cultural evolution.

1 Introduction

Social learning refers to the acquisition of knowledge, skills, and behaviors through observation and imitation within a group, underpins many of the behaviors seen across various species, facilitating the transmission of crucial information from one individual to another [1, 2]. It manifests in both direct actions and subtler forms like verbal cues or media [3, 4, 5]. Several compelling examples of social learning are evident within the domain of foraging, as illustrated by behaviors ranging from the potato-washing of Japanese monkeys to the hunting methods of killer whales, and even the foraging traditions in wild birds like the great tits [6, 7, 8, 9]. These examples are evidence of how learning and behavior are molded by social factors across the animal kingdom.

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

While many species exhibit collective foraging, humans stand out for their uniquely sophisticated collaborative skills, notably sharing knowledge and resources extensively within their social groups [10]. The genesis of such uniquely human collaboration is postulated by the interdependence hypothesis, suggesting that ecological conditions experienced by early humans made them obligate collaborative foragers, deepening their interdependence and fostering social learning [11, 12]. This ability to learn socially and transfer knowledge inter-generationally has been central to human success, leading to cumulative cultural evolution [13]. The essence of this evolution lies in its focus on high fidelity storage and transmission of knowledge, since loss of accumulated information can lead to significant challenges in rediscovery [14, 15]. Still, the precise cognitive mechanisms fueling this high-fidelity social learning are an active area of exploration [16, 17, 18].

Within this context, we hypothesize that episodic memory plays a potentially pivotal role in the cognitive scaffolding of high-fidelity social learning. Episodic memory, characterized by its capacity to record and recall specific events and experiences, is crucial for the accurate imitation and adaptation of observed behaviors [19]. This type of memory enables individuals to not only observe and replicate actions but also to communicate them to others, enhancing the precision and adaptability of social learning [20, 21, 22].

Despite their widespread use in the study of collective foraging and social learning, agent-based models often lack the resolution to investigate specific cognitive processes such as episodic memory [23, 24, 25, 26, 27]. Neuro-computational models, anchored in human and animal cognition, offer a promising avenue for modeling individual and social learning, while also being able to quantify cognitive aspects that remain elusive in conventional human studies or agent-based models, such as the content of episodic memory [28, 29, 30, 31]. Computational models that integrate episodic memory, such as episodic control algorithms, are particularly skilled at storing past events and rapidly learning from these experiences [32, 33, 34]. This capability makes them especially suitable for examining the dynamics of transmission and storage of information integral to cumulative cultural evolution.

In this paper, we investigate the relationship between episodic memory and social learning in a collaborative foraging task. We explore (Q1) the impact of episodic memory on social learning; (Q2) the outcomes of variations in social learning frequency and fidelity on foraging effectiveness; and (Q3) the influence of social learning fidelity on individual agent performance. We propose the following hypotheses: (H1) the depth of episodic memory influences an agent's behavioral efficiency; (H2) high-fidelity social learning boosts collective foraging, but its low-fidelity counterpart may not help or even hinder performance as frequency rises; (H3) sharing accurate information enhances individual agent efficiency, ensuring equitable reward distribution, whereas misinformation offers no advantage. To address these questions, we conduct experiments on a collaborative foraging task using groups of Sequential Episodic Control agents [35, 36]. We evaluate the impact of episodic memory length in social learning by manipulating key factors, including the transmission interval and transmission noise. This approach helps us understand their effects on both individual and collective resource gathering capacities. Our results shed light on the interplay between individual cognitive mechanisms, social learning dynamics, and collective behavior.

2 Methods

2.1 Experimental setup

In our study, we simulate a collaborative foraging scenario on an 11x15 2D grid-world, containing four agents, four fruits, and a nest (refer to Figure 1). Agents are rewarded when they deposit fruits in the nest and can transfer fruits to other agents. An episode ends when all fruits are collected or at 1000 timesteps, and each simulation runs a total of 5000 episodes. Agents possess limited visibility, seeing only a 3x3 grid segment depending on their location and direction. The grid cells are coded by attributes like object type, color, and information about other agents. An agent's state s_t is formed from these observed grid cell details, and their actions a_t are represented as integer numbers indicating specific actions, such as turning or picking objects.

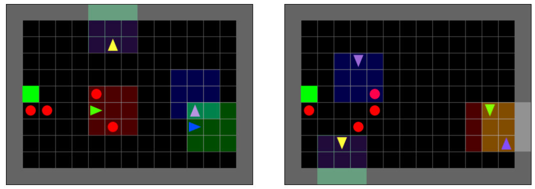

Figure 1: Collaborative foraging task modelled in a 2D grid-world environment. The environment contains four agents, four fruits (red circles), and a nest (green square). Agents are color-coded triangles with a 3x3 viewing area.

2.2 Sequential Episodic Control

This study employs the Sequential Episodic Control (SEC) algorithm to model foraging agents capable of storing past experienced events and learning from them [35, 36]. SEC is a type of episodic control model designed to guide an agent's behavior based on previously rewarding state-action sequences, as opposed to classical episodic control algorithms that store state-action couplets in isolation [34, 37, 38]. In turn, SEC considers state-action pairs as integrated representational primitives and stores the complete sequence of state-action pairs leading to goal states, conserving their serial order, a key feature of the hippocampal function [39, 40]. SEC functions in two phases: storage and retrieval. During storage, recent state-action pairs are held in the STM, with the LTM storing rewarding sequences. For retrieval, SEC accesses memories from LTM based on their relevance to the current state. This relevance is determined by an eligibility score, which is calculated based on how similar the current state is to the stored states in the LTM and how recently the memory was retrieved (for a detailed mathematical description of the algorithm, see [36]). Then, algorithm calculates the value of each potential action based on the eligibility score of the selected memories and their associated discounted rewards. Finally, the algorithm generates a probability distribution over the action space $Q(s, a)$, from which it samples the action to be performed. Initially, the algorithm explores actions randomly, but it becomes more selective as it accumulates more episodic memories.

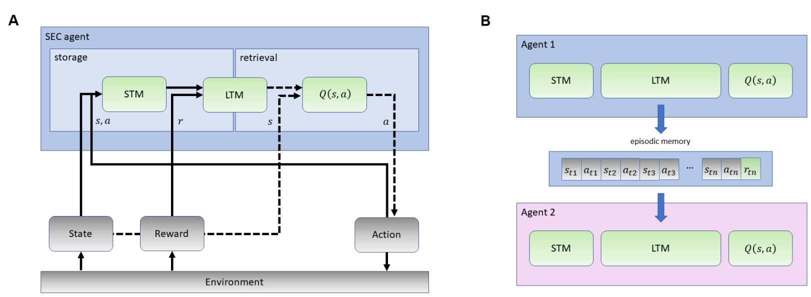

Figure 2: Panel A: The SEC model diagram for the 2D foraging task. SEC has storage and retrieval phases. In storage, agents store state-action (s, a) in short-term memory (STM). On receiving a reward (r), STM content moves to long-term memory (LTM). In retrieval, agents use LTM to compute the state-action value function $(Q(s, a))$ and select actions. Panel B: Social learning in SEC agents: Agent 1 (blue) shares an episodic memory with Agent 2 (pink), who saves it in LTM.

We consider a set of four SEC agents interacting together and with the capacity to share past experienced events with each other. We model social learning in episodic control as the transmission of episodic memories between agents. In this context, the transmission of social information is inherently local and decentralized, as agents can only share information with other agents present within their field of view (see Figure 1). The episodic memories that are socially transmitted follow a process of 'prioritized experience sharing'[31] (similar to prioritized experience replay [41]), whereby memory sequences associated with higher reward values are shared more often. In other words, there is a prioritization in the retrieval of sequential memories proportional to the reward value associated with episodic memory. There is evidence of such a process of prioritization taking place during memory retrieval and replay in the rodent hippocampus [42, 43, 44].

Two main factors influence social learning in our model: transmission interval T_i and transmission noise T_n. The transmission interval, T_i, represents the number of episodes between each transmission of information between agents, thus dictating the frequency at which agents share episodic memories. If $T_i = 1$, it means that agents can share information at every episode, while $T_i = 50$ implies that they can only interact with each other every 50 episodes. On the other hand, transmission noise, T_n, measures the likelihood of information distortion or loss during these transmissions. In other words, a value of $T_n = 0.1$ means there is a 10% probability of information loss for each element composing the episodic memory, following a process akin to mutation in genetic algorithms.

3 Results

In the context of this study, the ability to store longer episodic memories translates into more complex behavioral sequences being shared and socially learned. Results show that episodic memory capacity significantly impacts agent performance, a result consistent with previous research [35, 36]. For instance, agents with their memory constrained to 10 units (STM = 10) earn fewer rewards, regardless of their ability to socially learn from their peers (see Figure 3). However, more memory capacity does not always lead to better results. Interestingly, agents with larger memory capacity (STM = 30) do not perform better than those with an average capacity (STM = 20). These results suggest that there is an optimal size for episodic memories that might vary with the particular task and environmental configuration. Furthermore, the results show that memory capacity does not influence the distribution of rewards among agents (3, right panels).

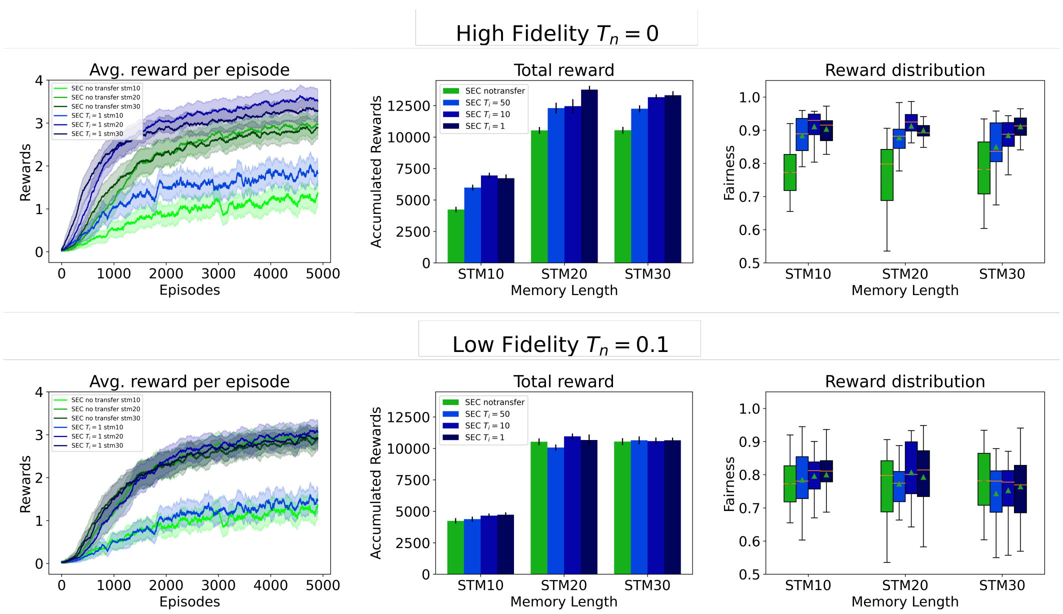

Figure 3: Performance results of SEC agents across different social learning conditions. Top: high-fidelity social learning ($T_n = 0$), Bottom: low-fidelity social learning ($T_n = 0.1$). Panels show average reward over time, total reward, and reward distribution ratio. Colors indicate social learning interval: green (none, $T_i = 0$), blue (every 50 episodes, $T_i = 50$), royal blue (every 10 episodes, $T_i = 10$), and continuous ($T_i = 1$). Left panels highlight $T_i = 0$ and $T_i = 1$ results.

During high-fidelity social interactions, as depicted in the top panels of Figure 3, we observed that increasing the frequency of social transmissions generally led to higher reward acquisition. This is true for the amount of rewards agents are able to collect each episode (Figure 3, top left) as well as for the total accumulated reward across episodes (Figure 3, top center). Regarding total accumulated reward, performance steadily increases along with the frequency of social interactions. This effect is consistent within each memory condition. In other words, more frequent high-fidelity social learning is beneficial for agents regardless of their memory capacity. In addition, social learning frequency also affects how evenly distributed is the recollection of rewards are among the population of agents

(Figure 3, top right). This frequency effect over the reward distribution remains constant across all memory conditions.

Regarding the performance metrics associated with low-fidelity social learning, the data suggests that such learning does not enhance the agents' ability to obtain rewards, both within individual episodes and overall (refer to Figure 3, bottom panels). Essentially, the outcomes from low-fidelity social learning mirror those of non-social learning agents with similar memory capacities, irrespective of how often social learning occurred. This starkly contrasts with the generally positive effect observed in reward acquisition as the frequency of high-fidelity social learning increased (see Figure 3, top panels). Furthermore, reward distribution is also negatively affected by low-fidelity social learning, aligning it with levels observed in agents under a non-social learning paradigm across all frequency conditions (Figure 3, bottom right).

4 Discussion

In this paper, we have investigated the interplay of social learning and episodic memory in collaborative foraging using sequential episodic control agents. Our results indicate that high-fidelity social learning leads to more efficient distribution of information between the group, resulting in improved collective performance, as well in a more equitable contribution of each agent to the collective effort. On the other hand, low-fidelity social learning offers no advantages over non-social learning in terms of individual and collective resource recollection. A crucial insight of this study is that while high-fidelity social learning is overall beneficial, its benefits are constrained by the cognitive capacity of individual agents to store and share longer episodic memories. In future steps,we plan to delve deeper into the agents' mnemonic content and examine how it correlates with other studied factors. To achieve this, we will use mnemonic metrics like diversity and alignment, as proposed in [31]. We will also explore intermediate levels of transmission noise to understand better the transition between low-fidelity and high-fidelity regimes in social learning.

Our study shows that high-fidelity social learning can improve individual and group resource recollection in a collaborative foraging task. However, it is important to note that our findings might be context-specific, derived from the particular properties of the presented foraging task, and therefore may not encompass all aspects of collaborative foraging. Recent studies suggest that communication noise and transmission errors in social learning can enhance collective problem-solving by maintaining diversity within a group [45]. The diversity produced by low-fidelity social learning could be advantageous in scenarios where resources are distributed across distant patches, as a group with lower fidelity might explore more areas, potentially leading to better resource utilization. These findings suggest that the relationship between the fidelity of social learning and the exploration-exploitation trade-off of collective knowledge might be more complex, and might also depend on the environmental constraints of each foraging context.

Similarly, in contexts resembling social dilemmas, as explored by [46], lower fidelity in social learning could prevent overexploitation of resources and encourage a more sustainable approach. The complexity of this issue is further illustrated in [25], where the balance between exploiting social information and engaging in individual exploration is highlighted. Their findings suggest that while social learning is beneficial in certain environments, its advantages are maximized when coupled with significant individual exploration.

In light of these perspectives, our conclusions should not be directly applied to other foraging contexts. Future research should explore how different levels of social learning fidelity impact group dynamics in varied foraging scenarios and social dilemmas, to better understand the balance between fidelity, exploration, and exploitation in collective foraging. Our study contributes to this ongoing discourse by highlighting the nuanced interplay of individual cognitive mechanisms and social learning dynamics, and have implications for enhancing coordination and efficiency in multi-agent systems. Furthermore, neuro-computational models like SEC, offer a promising perspective to study open-ended cultural evolution dynamics by capitalizing on the potential of episodic control models to bridge individual cognitive properties with individual and collective behavior.

References

[1] A. Whiten, Primate culture and social learning, Cognitive Science 24 (3) (2000) 477–508.

[2] K. N. Laland, Social learning strategies, Animal Learning & Behavior 32 (1) (2004) 4–14.

[3] A. Bandura, R. H. Walters, Social learning theory, Vol. 1, Englewood cliffs Prentice Hall, 1977.

[4] D. Munkenbeck Fragaszy, E. Visalberghi, Social processes affecting the appearance of innovative behaviors in capuchin monkeys, Folia primatologica 54 (3-4) (1990) 155–165.

[5] J. E. Grusec, Social learning theory and developmental psychology: The legacies of robert r. sears and albert bandura., Developmental Psychology 28 (1992) 776–786.

[6] S. Kawamura, The process of sub-culture propagation among japanese macaques, Primates 2 (1959) 43–60.

[7] L. M. Aplin, D. R. Farine, J. Morand-Ferron, A. Cockburn, A. Thornton, B. C. Sheldon, Experimentally induced innovations lead to persistent culture via conformity in wild birds, Nature 518 (7540) (2015) 538–541.

[8] N. A. Black, C. M. Talty, A. Schulman-Janiger, M. Srinivasan, Mammal hunting killer whales off monterey, california: A 30-year synthesis, in: Social Strategies of Carnivorous Mammalian Predators: Hunting and Surviving as Families, Springer, 2023, pp. 283–333.

[9] A. Whiten, C. P. Van Schaik, The evolution of animal 'cultures' and social intelligence, Philosophical Transactions of the Royal Society B: Biological Sciences 362 (1480) (2007) 603–620.

[10] A. B. Migliano, L. Vinicius, The origins of human cumulative culture: from the foraging niche to collective intelligence, Philosophical Transactions of the Royal Society B 377 (1843) (2022) 20200317.

[11] M. Tomasello, A. P. Melis, C. Tennie, E. Wyman, E. Herrmann, Two key steps in the evolution of human cooperation: The interdependence hypothesis, Current anthropology 53 (6) (2012) 673–692.

[12] M. Tomasello, The ultra-social animal, European Journal of Social Psychology 44 (3) (2014) 187–194, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ejsp.2015. doi:10.1002/ejsp.2015.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/ejsp.2015

[13] R. Boyd, P. J. Richerson, J. Henrich, The cultural niche: Why social learning is essential for human adaptation, Proceedings of the National Academy of Sciences of the United States of America 108 (SUPPL. 2) (2011) 10918–10925. doi:10.1073/pnas.1100290108.

[14] H. M. Lewis, K. N. Laland, Transmission fidelity is the key to the build-up of cumulative culture, Philosophical Transactions of the Royal Society B: Biological Sciences 367 (1599) (2012) 2171–2180.

[15] M. Montrey, T. R. Shultz, The evolution of high-fidelity social learning, Proceedings of the Royal Society B: Biological Sciences 287 (1928) (2020) 20200090. doi:10.1098/rspb.2020.0090.
URL https://royalsocietypublishing.org/doi/10.1098/rspb.2020.0090

[16] A. Mesoudi, Cultural evolution: a review of theory, findings and controversies, Evolutionary biology 43 (2016) 481–497.

[17] R. Boyd, P. J. Richerson, et al., Why culture is common, but cultural evolution is rare, in: Proceedings-British Academy, Vol. 88, Oxford University Press Inc., 1996, pp. 77–94.

[18] C. Tennie, J. Call, M. Tomasello, Ratcheting up the ratchet: on the evolution of cumulative culture, Philosophical Transactions of the Royal Society B: Biological Sciences 364 (1528) (2009) 2405–2415.

[19] S. B. Klein, L. Cosmides, C. E. Gangi, B. Jackson, J. Tooby, K. A. Costabile, Evolution and episodic memory: An analysis and demonstration of a social function of episodic recollection, Social Cognition 27 (2) (2009) 283–319.

[20] J. B. Mahr, G. Csibra, Why do we remember? the communicative function of episodic memory, Behavioral and brain sciences 41 (2018) e1.

[21] R. A. Mar, R. N. Spreng, Episodic memory solves both social and nonsocial problems, and evolved to fulfill many different functions, Behavioral and Brain Sciences 41 (2018).

[22] R. Fivush, The sociocultural functions of episodic memory, Behavioral and Brain Sciences 41 (2018).

[23] I. D. Couzin, J. Krause, N. R. Franks, S. A. Levin, Effective leadership and decision-making in animal groups on the move, Nature 433 (7025) (2005) 513–516.

[24] L. Conradt, T. J. Roper, Consensus decision making in animals, Trends in ecology & evolution 20 (8) (2005) 449–456.

[25] K. Garg, C. T. Kello, P. E. Smaldino, Individual exploration and selective social learning: Balancing exploration-exploitation trade-offs in collective foraging, Journal of the Royal Society Interface 19 (189), number: 189 (2022). `doi:10.1098/rsif.2021.0915`.

[26] A. Baronchelli, The emergence of consensus: a primer, Royal Society open science 5 (2) (2018) 172189.

[27] R. X. Hawkins, N. D. Goodman, R. L. Goldstone, The emergence of social norms and conventions, Trends in cognitive sciences 23 (2) (2019) 158–169.

[28] G. Maffei, D. Santos-Pata, E. Marcos, M. Sánchez-Fibla, P. F. Verschure, An embodied biologically constrained model of foraging: from classical and operant conditioning to adaptive real-world behavior in DAC-X, Neural Networks 72 (2015) 88–108. `doi:10.1016/j.neunet.2015.10.004`.

[29] R. S. Sutton, A. G. Barto, Reinforcement learning: An introduction, MIT press, 2018.

[30] M. Botvinick, S. Ritter, J. X. Wang, Z. Kurth-Nelson, C. Blundell, D. Hassabis, Reinforcement learning, fast and slow, Trends in cognitive sciences 23 (5) (2019) 408–422.

[31] E. Nisioti, M. Mahaut, P.-Y. Oudeyer, I. Momennejad, C. Moulin-Frier, Social Network Structure Shapes Innovation: Experience-sharing in RL with SAPIENS, arXiv preprint arXiv:2206.05060 (2022).
URL `http://arxiv.org/abs/2206.05060`

[32] S. J. Gershman, N. D. Daw, Reinforcement learning and episodic memory in humans and animals: an integrative framework, Annual review of psychology 68 (2017) 101–128.

[33] M. Lengyel, P. Dayan, Hippocampal contributions to control: the third way, in: Advances in neural information processing systems, 2008, pp. 889–896.

[34] C. Blundell, B. Uria, A. Pritzel, Y. Li, A. Ruderman, J. Z. Leibo, J. Rae, D. Wierstra, D. Hassabis, Model-free episodic control, arXiv preprint arXiv:1606.04460 (6 2016).
URL `http://arxiv.org/abs/1606.04460`

[35] I. T. Freire, A. F. Amil, V. Vouloutsi, P. F. Verschure, Towards sample-efficient policy learning with dac-ml, Procedia Computer Science 190 (2021) 256–262.

[36] I. T. Freire, A. F. Amil, P. F. Verschure, Sequential episodic control, arXiv preprint arXiv:2112.14734 (2021).

[37] A. Pritzel, B. Uria, S. Srinivasan, A. P. Badia, O. Vinyals, D. Hassabis, D. Wierstra, C. Blundell, Neural episodic control, in: International conference on machine learning, PMLR, 2017, pp. 2827–2836.

[38] Z. Lin, T. Zhao, G. Yang, L. Zhang, Episodic memory deep q-networks, arXiv preprint arXiv:1805.07603 (2018).

[39] G. Buzsáki, D. Tingley, Space and time: the hippocampus as a sequence generator, Trends in cognitive sciences 22 (10) (2018) 853–869.

[40] D. Santos-Pata, A. F. Amil, I. G. Raikov, C. Rennó-Costa, A. Mura, I. Soltesz, P. F. Verschure, Epistemic autonomy: self-supervised learning in the mammalian hippocampus, Trends in Cognitive Sciences (2021).

[41] T. Schaul, J. Quan, I. Antonoglou, D. Silver, Prioritized experience replay, arXiv preprint arXiv:1511.05952 (2015).

[42] D. J. Foster, M. A. Wilson, Reverse replay of behavioural sequences in hippocampal place cells during the awake state, Nature 440 (7084) (2006) 680–683.

[43] H. F. Ólafsdóttir, C. Barry, A. B. Saleem, D. Hassabis, H. J. Spiers, Hippocampal place cells construct reward related sequences through unexplored space, Elife 4 (2015) e06063.

[44] M. G. Mattar, N. D. Daw, Prioritized memory access explains planning and hippocampal replay, Nature neuroscience 21 (11) (2018) 1609–1617.

[45] A. Boroomand, P. E. Smaldino, Superiority bias and communication noise can enhance collective problem solving, Journal of Artificial Societies and Social Simulation 26 (3) (2023).

[46] J. Perolat, J. Z. Leibo, V. Zambaldi, C. Beattie, K. Tuyls, T. Graepel, A multi-agent reinforcement learning model of common-pool resource appropriation, Advances in neural information processing systems 30 (2017).