Humans and Mice Navigate Mazes Alike. Can AI Beat Them?

Rogério Guimarães (rogerio@caltech.edu)
Division of Biology and Biological Engineering, California Institute of Technology
1200 E California Blvd, Pasadena, CA 91125, USA

Alina Zhang (alinazhang@caltech.edu)
Division of Engineering and Applied Science,, California Institute of Technology
1200 E California Blvd, Pasadena, CA 91125, USA

Frank Xiao (frank@caltech.edu)
Division of Engineering and Applied Science, California Institute of Technology
1200 E California Blvd, Pasadena, CA 91125, USA

Evan Wang (ezwang@caltech.edu)
Division of Engineering and Applied Science, California Institute of Technology
1200 E California Blvd, Pasadena, CA 91125, USA

Jieyu Zheng (jzzheng@caltech.edu)
Division of Biology and Biological Engineering, California Institute of Technology
1200 E California Blvd, Pasadena, CA 91125, USA

Pietro Perona (perona@caltech.edu)
Division of Engineering and Applied Science, California Institute of Technology
1200 E California Blvd, Pasadena, CA 91125, USA

Abstract

Mice are natural navigators, capable of few-shot learning in complex environments. Here, we explore how their performance in maze navigation can serve as a benchmark for comparing learning across species and artificial agents. We tested human participants on a virtual binary maze game adapted from a prior mouse study and found not only similar performance, but also striking parallels in learning dynamics. Like mice, humans rapidly optimized reward acquisition, exhibited sudden insights about the maze structure, and showed knowledge of the path home from their very first maze incursion. We then used this embodied navigation task to compare AI agents with both species. We showed that two canonical agents — a Deep Q-Learning (DQN) and a Large Language Model (LLM) — were outperformed by the biological learners. These results highlight the potential of naturalistic learning tasks for cross-species comparisons, and expose challenges and opportunities for advancing AI.

Keywords: navigation; behavior; cross-species; rapid learning; LLM; reinforcement learning; artificial intelligence

Introduction

Moving from less to more desirable locations is among the earliest challenges nervous systems evolved to solve (Sterling & Laughlin, 2017), making navigation a natural task for comparing performance across species. Prior work has compared mice and humans in open environments with global beacons and no walls, using reinforcement learning models to capture their behavior (de Cothi et al., 2022). Yet, real-world environments often feature constrained movement and limited visual cues, such as underground burrows for mice or urban grids for humans. To address this, we tested humans and AI agents in the binary maze proposed by Rosenberg, Zhang, Perona, and Meister (2021). Their study showed that behaviorally naïve mice could quickly learn and memorize reward paths in a symmetric and complex environment complete darkness, a test focused on memory and spatial reasoning.

We adapted this task to a virtual maze (Fig. 1) for humans and AI, applying a restricted field of view to replicate the mouse's tactile-only experience. Humans demonstrated similarly rapid learning within one hour — matching mouse performance and timescale — whereas traditional paradigms like 2AFC show big disparities (days for mice vs seconds for humans). At last, while most current AI benchmarks like ARC-AGI (Chollet, Knoop, Kamradt, & Landers, 2025) focus on human-level reasoning, we show that even achieving mouse-level performance in the binary maze is challenging. We propose complex-maze navigation as a cross-species benchmark for spatial reasoning in biological and artificial agents.

Experiments

We collected data on humans and AI agents in the maze and compared to mice data from Rosenberg et al. (2021).

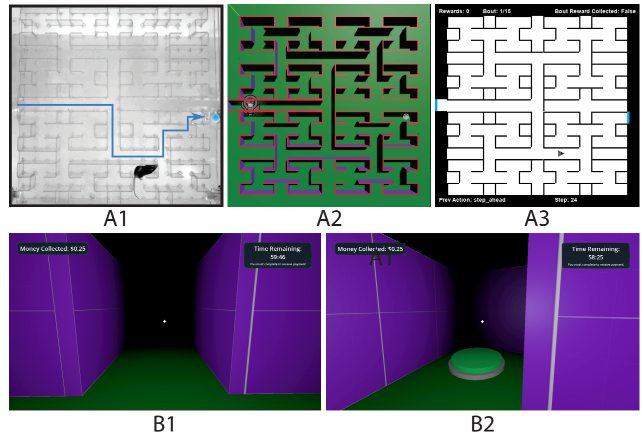

Figure 1: **The Binary Maze.** A. Bird-eye view of the original binary maze for mice from (Rosenberg et al., 2021) (A1), and the virtual binary mazes we developed for humans (A2) and AI models (A3). B. Example of the experience of a human subject in the virtual maze going through a corridor with a restricted field of view (B1) and finding the reward location (B2).

Human Subjects

Human participants were recruited via Prolific and completed the experiment remotely using standard first-person game controls. Maze speed was set to 1.5 tiles/s, and field of view was limited to 1.5 tiles ahead (Fig. 1) to match mice navigating in the dark. Subjects received monetary rewards for collecting virtual rewards placed at the same location as the water port in the original experiment and at the home chamber. The extra home reward was needed to motivate the homing behavior, which mice do intrinsically for food and safety. Subjects could only collect the rewards if they alternated between them, which required navigating the maze. Like the mice, they were not trained in any maze navigation tasks prior to the behavioral test, and were not informed of the reward contingency.

Humans achieved similar in-maze reward collection in 1h to mice (humans: 16.4 ± 10.0, mice: 18.6 ± 5.4). Moreover, 30% of both species showed the previously reported sudden insight (5/17 humans, 3/10 mice), a discontinuity in the learning process after which navigation becomes more efficient (Fig. 2). Human reward curves were steeper, driven by more exploitative behavior. Unlike mice, humans refrained from exploration after learning the path to reward, and thus did not show increase in long direct paths from within the maze to reward. We also found in humans the asymmetry in learning the paths to reward and back home reported in mice (Fig. 3), suggesting humans retained home-location information while exploring. On their first maze bout, both species reached the reward significantly slower than they returned home (mice: 8.1 min vs. 4.5 min; humans: 6.4 min vs. 1.8 min). A key behavioral difference was that some humans adopted wall-following strategies (Fig. 2), which mice did not. Post-experiment questionnaires revealed this came from prior knowledge of maze-solving heuristics, which slowed reward collection.

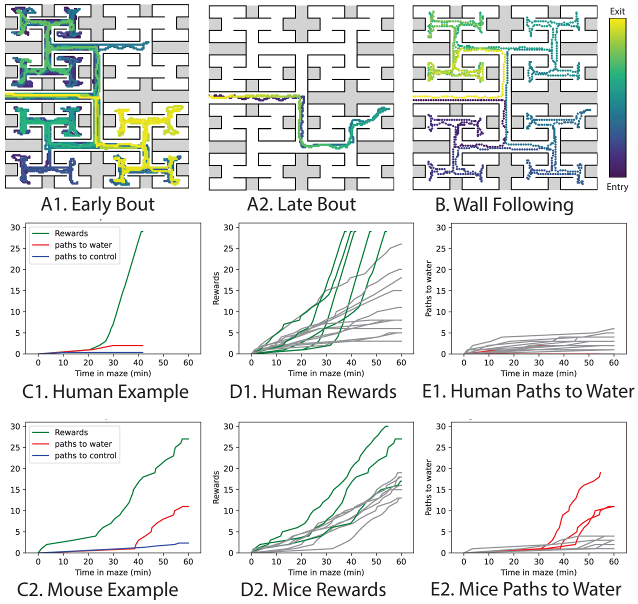

Figure 2: **Humans show Rapid Learning and Sudden Insight.** A. Example of an early (A1) vs late (A2) bout of a human that had the sudden insight. B. A late bout from another subject that stuck to wall following. C. Example results from a single human (C1) and mouse (C2), with all cumulative rewards, and direct long paths (> 6 junctions) from within the maze to reward and to control end nodes. D. Cumulative Rewards for all all humans (D1) and mice (D2). E. Cumulative long paths to water for all humans (E1) and mice (E2).

AI Agents

To test blind AI agents, we created an environment with only local observations: the objects in front, behind, left, and right, classified as wall, reward, or empty. Actions included stepping forward/back or turning 90°. Inputs were encoded as one-hot vectors for the DQN (Mnih et al., 2013) and as text prompts for the LLM (OpenAI o1-mini) (Jaech et al., 2024). When given only local observations, neither agent collected any reward, performing much worse than humans and mice under the same conditions. After augmenting input with positional coordinates and orientation (emulating path integration), the DQN learned the optimal path after 100,000 steps. The LLM collected a few rewards within 600 steps, still below average biological performance (Fig. 4) within a similar time frame. Previous work has explored the navigation abilities of LLMs (Martorell, 2025) and even blind RL agents (Wijmans et al., 2023), but never in direct comparison to nonhuman animals.

This task also poses computational challenges for LLMs. They require full input-output history to support in-context learning, hitting input token limits at 600 steps that force a hard limit in experiment duration. Runtime and price are also high for an LLM with long prompts: running 600 steps on OpenAI o1-mini took $\sim$7h and cost \$15. By contrast, a small DQN can consistently solve the maze after 100,000 steps, which takes one minute on a single GPU.

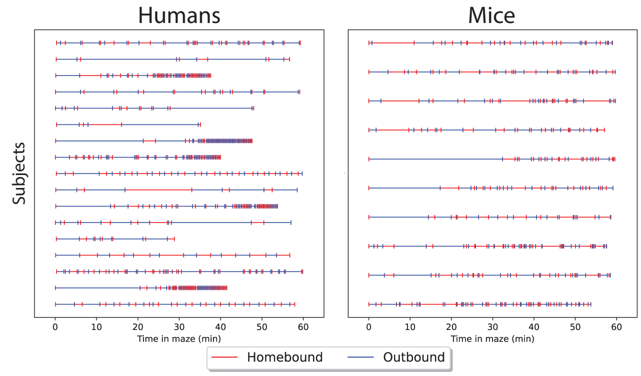

Figure 3: **Faster homing in humans vs. mice** Time in maze during outbound vs. homebound bouts. Both species returned home faster than they reached the reward, even on the first attempt.

Conclusions and Further Directions

Our key findings are the surprisingly similar learning patterns in mice and humans, including rapid reward acquisition and sudden insight, and their contrasting exploratory tendencies. Future directions include more immersive environments for humans, like virtual reality or real mazes. We also show that naturalistic navigation tasks can be a unifying learning benchmark across species and AI agents. We propose that navigation with quantitative results from humans and other animal species can be used as a spatial reasoning benchmark for AI agents, and we provide initial results with two models. Our results are limited in the diversity of agents so future work will expand the testing to different architectures. Different frameworks to adapt the 3D maze task to text-based LLMs that are more computationally efficient must also be tested. Furthermore, the current version of the maze can be easily solved by algorithms that exploit prior knowledge about its grid-world discrete nature, like Dyna (Sutton, 1991), which limits their utility for public benchmarks. Future work will also introduce continuous, noisy environments that better reflect real-world navigation challenges, and should only be solvable by artificial agents capable of performing equally well if deployed in novel real world environments.

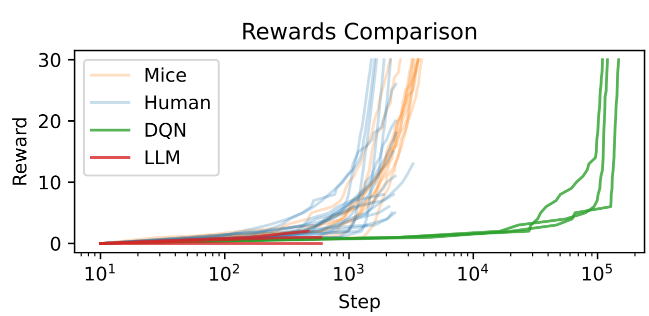

Figure 4: **AI vs Biological Agents.** Cumulative rewards of DQN and LLM with path integration vs humans and mice.

Acknowledgements

This project is funded by Simons Collaboration on the Global Brain.

References

Chollet, F., Knoop, M., Kamradt, G., & Landers, B. (2025, January). *ARC Prize 2024: Technical Report.* arXiv. Retrieved 2025-04-11, from http://arxiv.org/abs/2412.04604 (arXiv:2412.04604 [cs]) doi: 10.48550/arXiv.2412.04604

de Cothi, W., Nyberg, N., Griesbauer, E.-M., Ghanamé, C., Zisch, F., Lefort, J. M., ... Spiers, H. J. (2022, September). Predictive maps in rats and humans for spatial navigation. *Current Biology*, *32*(17), 3676–3689.e5. Retrieved 2023-01-02, from https://www.sciencedirect.com/science/article/pii/S0960982222010958 doi: 10.1016/j.cub.2022.06.090

Jaech, A., Kalai, A., Lerer, A., Richardson, A., El-Kishky, A., Low, A., ... others (2024). Openai o1 system card. *arXiv preprint arXiv:2412.16720*.

Martorell, N. (2025, February). *From Text to Space: Mapping Abstract Spatial Models in LLMs during a Grid-World Navigation Task.* arXiv. Retrieved 2025-06-12, from http://arxiv.org/abs/2502.16690 (arXiv:2502.16690 [cs]) doi: 10.48550/arXiv.2502.16690

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013, December). Playing Atari with Deep Reinforcement Learning. *arXiv:1312.5602 [cs]*. Retrieved 2021-12-12, from http://arxiv.org/abs/1312.5602 (arXiv: 1312.5602)

Rosenberg, M., Zhang, T., Perona, P., & Meister, M. (2021, July). Mice in a labyrinth show rapid learning, sudden insight, and efficient exploration. *eLife*, *10*, e66175. Retrieved 2024-04-10, from https://elifesciences.org/articles/66175 doi: 10.7554/eLife.66175

Sterling, P., & Laughlin, S. B. (2017). *Principles of neural design*. Cambridge: MIT press.

Sutton, R. S. (1991). Dyna, an integrated architecture for learning, planning, and reacting. *ACM Sigart Bulletin*, *2*(4), 160–163. (Publisher: ACM New York, NY, USA)

Wijmans, E., Savva, M., Essa, I., Lee, S., Morcos, A. S., & Batra, D. (2023, June). Emergence of Maps in the Memories of Blind Navigation Agents. *AI Matters*, *9*(2), 8–14. Retrieved 2024-04-18, from https://dl.acm.org/doi/10.1145/3609468.3609471 doi: 10.1145/3609468.3609471