EUNICE YIU

Email: ey242@berkeley.edu | Website: https://ey242.github.io

Research Profile

I study how humans and AI systems build, generalize, and update models of the world, and what it means when those processes align or diverge. My work combines rigorous behavioral experiments with controlled model evaluation to probe whether similarities between human and model cognition are superficial or structural.

Employment

Sept 2025 - present — **University of California, Berkeley, Berkeley, CA**
Postdoctoral Scholar, Department of Psychology
Active projects: causal learning and active exploration in humans and LLMs (with Google, Mila); effects of AI on human cognitive and behavioral health (with Brown ARIA)

Education

Aug 2020 – Aug 2025 — **University of California, Berkeley, Berkeley, CA**
Ph.D. in Psychology
Dissertation: Relational Reasoning in Children and Machines: Insights into Causal Generalization and Innovation
Committee: Alison Gopnik, Jitendra Malik, Steve Piantadosi & Shiry Ginosar

Aug 2016 – May 2020 — **Cornell University, Ithaca, NY**
B.A. in Psychology, Biological Sciences, and Economics; Cumulative GPA: 3.996/4.0

Grants & Fellowships

2026 — **Cambridge Boston Alignment Initiative Summer Research Fellowship in AI Safety** ($34,000)

2026 — **Princeton Natural & Artificial Minds Diverse Intelligences Fellowship** ($5,000)

2025-2026 — **Canadian Institute for Advanced Research Next Generation Trainee Fellowship for the Learning in Machines & Brains program** ($10,000 CAD)
Examining whether LLMs replicate the human advantage of active exploration in causal discovery, with implications for how AI augmentation may reshape scientific inquiry.

2025-2026 — **7th Google - Berkeley Artificial Intelligence Research Commons Grant** ($10,000 GCP credits)
Fine-tuning and evaluating VLMs via iterative feedback loops in simulated environments and comparing their learning trajectories to human developmental patterns.

2025 — **Departmental Semester Fellowship at UC Berkeley** ($18,750)

2024-2025 — **6th Google - Berkeley Artificial Intelligence Research Commons Grant** ($41,000)
Designed a comparative study of human and artificial agents on learning and optimizing causal structures through intrinsic objectives

2023-2024 — **5th Meta - Berkeley Artificial Intelligence Research Commons Grant** ($35,000)
Developed a curriculum for foundational AI Models inspired by human cognition

Selected Publications (* equal contribution, † undergraduate mentee)

1. **Yiu, E.**, Kosoy, E., & Gopnik, A. (2024). Transmission Versus Truth, Imitation Versus Innovation: What Children Can Do That Large Language and Language-and-Vision Models Cannot (Yet). *Perspectives on Psychological Science*, 17456916231201401.
2. Samiei, M.*, **Yiu, E.***, Chen, G. C., Lin, D., Richards, B., Gopnik, A. & Precup, D (2026). Human adults and LLMs as Scientists: Who Benefits from Active Exploration? In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 48).
3. **Yiu, E.**, Qraitem, M., Wong, C.†, Majhi, A. N.†, Bai, Y., Ginosar, S., ... & Saenko, K. (2025). KiVA: Kid-inspired visual analogies for testing large multimodal models. In *International Conference on Learning Representations*.
4. **Yiu, E.**, Allen, K., Ginosar, S., & Gopnik, A. (2026). Empowerment gain and causal model construction: children and adults are sensitive to controllability and variability in their causal interventions. *Philosophical Transactions of the Royal Society A*, *384*(2320). Special Issue: World models, A(G)I, and the Hard problem(s) of life–mind continuity.
5. **Yiu, E.***, Sandbrink, K. J.*, & Gopnik, A. (2024). To observe or to bet? Investigating purely exploratory and purely exploitative actions in children, adults, and computational models. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 46).
6. Dahmani, A.*, **Yiu, E.***, & Gopnik, A. (*2025*). Children Spontaneously Design Curricula to Tackle Challenging Tasks. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 47).

Under Review:

1. **Yiu, E.**, Ginosar, S., Gopnik, A., Allen, K.R., & Murphy, K.P. Acting without Knowing in VLMs.
2. Cheng, S., Wu, A., **Yiu, E.**, & Artzi, Y. REMAP: Evaluating Geometric Dual Representations in Spatial Reasoning.

Invited Talks

Jan 2026	Keynote at the Foundations of Agentic Systems Theory Workshop at AAAI, Singapore
March 2025	The Nature of Intelligence Workshop, Santa Fe Institute
June 2024	AI, Psychology & Neuroscience Summer Cluster, Simons Institute for the Theory of Computing
May 2024	Brain Science and Large Language Models Symposium, Leopoldina and Max Planck Institute for Brain Research, Germany (media coverage: *Frankfurter Allgemeine Zeitung*)
Invited Lab Talks	Stanford (Judith Fan, Michael Frank), Harvard (Tomer Ullman), Brown (Daphna Buchsbaum), UC San Diego (Caren Walker)

Professional Service

Reviewer

Cognitive Science	Journal of Experimental Psychology, Topics in Cognitive Science, Cognitive Science Society (CogSci), Society for Philosophy & Psychology (SPP)
Artificial Intelligence	NeurIPS, Association for Computational Linguistics (ACL), Conference on Language Modeling (COLM), Conference on Computer Vision and Pattern Recognition (CVPR)

Co-Organized Workshops and Challenges

2026	Humans of Generative AI, Conference Workshop at Computer Vision Pattern Recognition
2026	Topics in Intelligence: World Models and Social Reasoning
2026	A Unified Account of Motivation in Development, Workshop at Cognitive Development Society
2025	Kid-inspired Visual Analogies (KiVA) Challenge (multi-month competition) at Google DeepMind's Third Perception Test Workshop, International Conference on Computer Vision
2024	AI & Cognitive Development, Preconference Workshop at Cognitive Development Society