

Xinyue Hu

xhu26@uci.edu | (845) 546-8258 | huxinyue.com | [Google Scholar](#)

Ph.D. Candidate, Cognitive Science, UC Irvine | Data Science · UX Research · Human-AI Interaction

SUMMARY

Mixed-methods researcher studying how people make decisions with, trust, and coordinate alongside AI systems. Seven years designing and running large-scale online experiments ($N > 2,000$), building web-based interactive study platforms, and modeling behavioral data with Bayesian and mixed-effects methods. Hands-on with LLM-driven agents, RAG-grounded conversational AI, and Wizard-of-Oz prototyping, with published work in Nature Machine Intelligence, ACM THRI, and IEEE RA-L. Prior industry experience at Toyota Research Institute, Honda Research Institute, and Microsoft.

EXPERIENCE

University of California, Irvine

July 2023 – Present

Graduate Student Researcher, Department of Cognitive Sciences (Advisor: Mark Steyvers)

- UCI Beall Applied Innovation collaboration (NarrA.I.tive): lead user research on a RAG-grounded persona AI avatar for engagement and learning.
 - Designed a study comparing the avatar against a generic LLM chatbot and static text on engagement, motivation, and intent to continue use; findings feed directly into RAG tuning and avatar product design.
- Honda Research Institute collaboration: multi-year applied research on how people coordinate and build trust with AI agents, translated into prototypes and design recommendations for deployed human-AI systems.
 - Designed and ran large-scale behavioral experiments ($N > 2,000$) on how people reason, coordinate, and adapt when working with AI agents in real-time multi-agent tasks; combined quantitative modeling with qualitative analysis of participant strategies.
 - Built end-to-end web experiment platforms (JavaScript, HTML/CSS, Firebase) for live human-AI interaction, and analysis pipelines in Python/R using mixed-effects logistic regression, survival analysis, and Bayesian hierarchical modeling.
 - Developing LLM-driven agents with adaptive feedback and tested how agent timing, framing, and language affect group coordination, trust, and perceived usefulness.
- LLM calibration: co-authored a Nature Machine Intelligence study showing gaps between what models know and what users believe they know, with implications for confidence display and trust design in AI products.

Toyota Research Institute, Cambridge, MA

June 2024 – Sept 2024

Research Scientist Intern, Human Interactive Driving

- Designed and ran a between-subjects online experiment ($N = 600$) evaluating 8 trust-repair strategies across 3 shared-control driving scenarios after AI errors.
- Found that neutral AI expression and increased user control drove the largest trust recovery; results informed HMI design recommendations for AI-assisted vehicles.
- First-author paper accepted to an invited session at IEEE ITSC 2025; co-inventor on two resulting patent applications.

Microsoft, Copenhagen, Denmark

March 2022 – May 2023

Student Worker, Software Engineering, Global Applications (Business Central)

- Automated delocalization of product features with Python, improving release consistency across 20+ countries and 40+ localized versions.
- Decoupled 20+ country-specific modules to enable cleaner localization patterns, supporting 3+ global engineering and localization teams.

University of Copenhagen, Copenhagen, Denmark

Jan 2022 – May 2023

Research Assistant, Department of Computer Science

- Ran a VR user study ($N = 40$) with eye-tracking and Unity-based avatars, combining gaze data with micro-phenomenological interviews to understand how users experience embodiment; published at HCII 2025.

SKILLS

Research methods: Experimental design, A/B testing and randomized interventions, large-scale online studies (Prolific/MTurk), Wizard-of-Oz prototyping, user interviews, micro-phenomenology, eye-tracking, survey design

Data & modeling: Python (pandas, scikit-learn, PyTorch, TensorFlow), R (lme4, brms), SQL, Bayesian hierarchical models, mixed-effects regression, survival analysis, agent-based simulation

AI systems: LLM-driven agents, RAG-grounded conversational AI, prompt design, LLM calibration and confidence evaluation, multi-agent coordination

Build & prototype: JavaScript, HTML/CSS, Firebase, Unity, Git, MATLAB, Rhino, Final Cut Pro

Languages: Mandarin (native), English (fluent), German (proficient)

SELECTED PUBLICATIONS

Steyvers, M., Tejada, H., Kumar, A., Belem, C., Karny, S., **Hu, X.**, Mayer, L. W., & Smyth, P. (2025). What large language models know and what people think they know. *Nature Machine Intelligence*.

Zahedi, Z., **Hu, X.**, Mehrotra, S., Steyvers, M., & Kumar, A. (2025). Strategic shaping of human prosociality: A latent-state POMDP framework. *IEEE Robotics and Automation Letters*.

Hu, X., Kumar, A., Mehrotra, S., Misu, T., & Steyvers, M. (2025). Toward promoting prosocial interactions between humans with autonomous agents. *ACM Transactions on Human-Robot Interaction*.

Hu, X., Mehrotra, S., Zahedi, Z., Misu, T., Kumar, A., & Steyvers, M. (2025). Does observing helping robots promote prosociality? Challenges of learning from observation in spatial environments. *Collective Intelligence Conference (CI 2025)*.

Full list of publications and presentations at huxinyue.com and [Google Scholar](https://scholar.google.com/citations?user=HUXINYUE).

SELECTED PRESENTATIONS

Invited session, IEEE ITSC 2025, Gold Coast, Australia. "Enhancing trust repair in driving: The role of AI emotional expression and perceived control."

Invited talk, ACM AutoUI 2024, Stanford, CA. "Toward promoting prosocial interactions between humans with autonomous agents."

Presenter, ACM Collective Intelligence 2025, San Diego, CA. "Does observing helping robots promote prosociality?"

Full list of publications and presentations at huxinyue.com

EDUCATION

University of California, Irvine

Ph.D., Cognitive Science (GPA 3.97/4.0)

Expected June 2027

University of Copenhagen

M.Sc., IT and Cognition

July 2023

Vassar College

B.A., Cognitive Science and German Studies

June 2021