

AAKRITI KUMAR

aakriti.kumar@kellogg.northwestern.edu ◊ www.aakritikumar.com ◊ [google scholar](#)

CURRENT APPOINTMENT

| | |
|--|------------------|
| Postdoctoral Researcher Kellogg School of Management, Northwestern University Northwestern Institute on Complex Systems | <i>Aug 2024-</i> |
|--|------------------|

EDUCATION

| | |
|---|--------------------|
| PhD in Cognitive Science University of California, Irvine Thesis: Human Mental Models of Self, Others, and AI Agents | <i>2018 - 2024</i> |
| MS in Statistics University of California, Irvine | <i>2018 - 2022</i> |
| BTech in Engineering Indian Institute of Technology, Madras | <i>2011 - 2015</i> |

RESEARCH EXPERTISE

Human-AI Collaboration, AI-Assisted Decision-Making, Computational Social Science, Behavioral Science, Cognitive Science, Statistical Modeling, Team Dynamics, Digital Experimentation

PEER-REVIEWED PUBLICATIONS

- [1] **A. Kumar**, N. Poungpeth, D. Yang, E. Farrell, B. Lambert, M. Groh. When Large Language Models are Reliable for Judging Empathic Communication. *Nature Machine Intelligence* (*accepted*)
- [2] **A. Kumar**, R. Tham, M. Steyvers (2025). Seeing Things Differently: The Role of Differing Perspectives in Advice-Taking. *Cognitive Science* (*Vol. 47*)
- [3] M. Steyvers, H. Tejeda, **A. Kumar**, S. Karny, X. Hu, L. Mayer, P. Smyth (2025). What large language models know and what people think they know. *Nature Machine Intelligence*
- [4] N. Kamali, K. Nakamura, **A. Kumar**, A. Chatzimpampas, J. Hullman, M. Groh (2025). Characterizing Photorealism and Artifacts in Diffusion Model-Generated Images. *CHI 2025*
- [5] **A. Kumar**, P. Smyth, M. Steyvers (2023). Differentiating mental models of self and others: A hierarchical framework for knowledge assessment. *Psychological Review*
- [6] M. Steyvers & **A. Kumar**, (2023). Three Challenges for AI-Assisted Decision-Making. *Perspectives on Psychological Science*
- [7] M. Kelly, **A. Kumar**, P. Smyth, M. Steyvers (2023). Capturing Humans' Mental Models of AI: An Item Response Theory Approach. *FAccT 2023*
- [8] **A. Kumar**, K. Akash, S. Mehrotra, T. Misu, M. Steyvers (2023). When Do Drivers Intervene In Autonomous Driving?. *Human Robot Interaction (Late-Breaking Report) 2023*
- [9] H. Lemus, **A. Kumar**, M. Steyvers (2022). How Displaying AI Confidence Affects Reliance and Hybrid Human-AI Performance. *Proceedings of the Second International Conference on Hybrid Human-Machine Intelligence*.
- [10] H. Lemus, **A. Kumar**, P. Smyth, M. Steyvers (2022). AI-assisted Decision-Making: A Cognitive Modeling Approach to Infer Latent Reliance Strategies. *Computational Brain & Behavior*

- [11] H. Lemus, **A. Kumar**, M. Steyvers (2022). An Empirical Investigation of Reliance on AI-Assistance in a Noisy-Image Classification Task. *Proceedings of the First International Conference on Hybrid Human-Machine Intelligence*.
- [12] **A. Kumar***, S. Chatterjee¹, P. Shenoy (2022). Meta-Learning of Dynamic Policy Adjustments in Inhibitory Control Tasks. *Cognitive Science* (Vol. 44).
- [13] **A. Kumar**, T. Patel, A. Benjamin, M. Steyvers (2021). Explaining Algorithm Aversion with Metacognitive Bandits. *Cognitive Science* (Vol. 43)
- [14] **A. Kumar**, A. Benjamin, A. Heathcote, M. Steyvers (2021). Comparing models of learning and relearning in large-scale cognitive training data sets. *NPJ Science of Learning*.
- [15] J. Rouder, **A. Kumar**, J. Haaf (2019). Why Most Studies of Individual Differences With Inhibition Tasks Are Bound To Fail. *Psych Bulletin & Review*

PAPERS IN PROGRESS

- [P1] **A. Kumar**, F. Poungpeth, D. Yang, B. Lambert, M. Groh. Practicing with Language Models Cultivates Human Empathic Communication.
- [P2] N. Dundas, **A. Kumar**, Z. Guo, O. Badri, A. Koochek, M. Groh. Mapping Diagnostic Blind spots by Physicians and AI in Dermatology.
- [P3] **A. Kumar**, S. Melumad, M. Groh. Scaffolding Human-AI Collective Intelligence via Metacognitive Forcing Functions.
- [P4] N. Kamali, **A. Kumar**, A. Chatzimparmpas, C. Gerstner, J. Hullman, M. Groh. Guiding Human Visual Attention to Spot AI-Generated Images.
- [P5] **A. Kumar**, P. Petrides, K. Chauncey. Simulating AV System Delays to Study User Stress & Behavioral Intention.

INVITED TALKS & PRESENTATIONS

Cultivating Human Empathy with AI Role-playing Partners. *Parallel Talk, Conference on Digital Experimentation @ MIT, 2025*

Cultivating Human Empathy with AI Partners. *Invited Speaker, DySoc Seminar 2025, University of Tennessee, Knoxville*

Cultivating Human Empathy with AI Partners. *Guest Speaker, Learn Behavioral Leadership Symposium, Chicago*

Large language models can provide expert-aligned judgments of empathic communication. *Lightning Talk, Northwestern Institute on Complex Systems Seminar, 2025*

What Decision-Makers Should Know About Generative AI. *JTIP Symposium 2025: Innovation, AI, and Access, Northwestern School of Law*

AI-assisted Decision Making. *UC Noyce Initiative Symposium, 2022*

Understanding Human-AI Teams. *Noyce Symposium, 2022*

Member, Panel on Human-AI Relationships. *Noyce Symposium, 2022*

Metacognitive Bandits: When Do Humans Seek AI Assistance? *Social Intelligence in Humans and Robots Workshop, ICRA 2021*

Explaining Algorithm Aversion with Metacognitive Bandits. *Workshop on Human-AI Collaboration in Sequential Decision-Making, ICML 2021*

^{1*} indicates equal contribution

AWARDS AND FELLOWSHIPS

| | |
|--|------------------------------|
| UCI Associate Dean's Fellowship 2023, UCI | <i>Total Award: \$7000</i> |
| Noyce Initiative in AI, Law, and Society Fellowship 2021-2023 | <i>Total Award: \$66,000</i> |
| Graduate Dean's Recruitment Fellowship 2018, UCI | <i>Total Award: \$5000</i> |
| Recruitment Fellowship Award, 2018, Department of Cognitive Science, UCI | <i>Total Award: \$5000</i> |

EXPERIENCE

| | |
|---|-------------------------|
| Quantitative Human-Computer Interaction Intern, Motional <i>Human Assistance in Self-Driving Robotaxis:</i> Investigated Motional's remote vehicle assistance capability in assisting their robotaxi. [P5] | <i>Apr'23 - Dec'23</i> |
| Human Behavior Modeling Intern, Honda Research Institute <i>Trust in Autonomous Vehicles:</i> Created a quantitative model to capture the effect of an autonomous vehicle's driving style on users' trust and risk perception. Identified scenarios perceived as risky across two mobility types [8] | <i>Jun'22 - Sept'22</i> |
| Statistical Consultant, Google Research, India <i>AI for Social Good:</i> Helped design data collection and analysis methods to evaluate an ML solution's efficacy in reducing drop-off from an Indian maternal health program. <i>Reliability in People's Inhibitory Control:</i> Proposed and tested RNN models to assess people's cognitive control task traits are reliable. [12] | <i>Jan'20 - Jan'21</i> |

TEACHING EXPERIENCE

| | |
|--|--------------------------------|
| Teaching Assistant, UC Irvine: Probability and Statistics in Psychology I, II, III | <i>Fall 2019 - Spring 2020</i> |
| Lead Teaching Assistant, UC Irvine: Introduction to Human Memory | <i>Winter 2019</i> |
| Teaching Assistant, UC Irvine: Introduction to Psychology | <i>Fall 2018</i> |

SKILLS

| |
|--|
| Programming: Python (pandas, numpy, matplotlib, scikit-learn, pytorch), R, Stan, HTML/CSS, Flask |
| Research Methods: Experiment Design, Computational Cognitive Models, Generalised Linear Models, Bayesian Statistics, Probabilistic Learning, Machine Learning |

MEMBERSHIPS AND SERVICE

| | |
|---|-----------------------|
| Reviewer (TiiS, Trends in Cognitive Sciences, Cognitive Science, Human-Robot Interaction, CHI, Applied Psychology: Health and Well-Being) | <i>2020 - Present</i> |
| Department Colloquium Coordination Committee Member | <i>2021 - 2022</i> |
| Cognitive Science Society Student Member | |