## Vivienne Bihe Chi

(+1) 734-604-9056 chi.vivienne@gmail.com linkedin.com/in/vivienne-chi/

**EDUCATION Brown University** Providence, RI Ph.D. in Cognitive Science 2019 - 2025 (expected) University of Michigan Ann Arbor, MI 2015 - 2019 B.S. in Computer Science, Cognitive Science Industry Google | Student Researcher | Mountain View, CA 2024 EXPERIENCE Gemini Safety Honda Research Institute | Research Scientist Intern | San Jose, CA 2023

rionda Research institute | Research Scientist Intern | San Jose, CA 202

Cognitive Modeling

Ageless Innovation | Researcher | Pawtucket, RI 2020-2023

• Robotics Design and Engineering

Wundertax (fintech startup) | Product Management Intern | Berlin, Germany 2017

INTERESTS LLM Safety, Responsible AI, Human-AI Interaction

Publications

Manuscripts

Chi, V. B.\* & Malle, B. F. Interactive Human-Robot Teaching Recovers and Builds Trust, Even With Imperfect Learners. *ACM/IEEE International Conference on Human-Robot Interaction (HRI)* [24.9% acceptance rate, oral presentation], 2024.

Chi, V. B.\*, Mehrotra, S., Teruhisa, M., Kumar, A. Should I Help a Delivery Robot? Cultivating Prosocial Norms through Observations. *ACM Conference on Human Factors in Computing Systems (CHI)*, 2024.

Chi, V. B.\*, Ulwelling, E., Salubre, K., Mehrotra, S., Teruhisa, M., Kumar, A. More than just a Tool: People's Perception and Acceptance of Prosocial Delivery Robots as Fellow Road Users. *In Proceedings of Workshop on Robo-Identity: Designing for Identity in the Shared World (HRI)*, 2024.

Malle, B. F., Chi, V. B.\* Norm Learning, Teaching, and Change. Current Opinion in Psychology, 2024.

Chi, V. B.\*, Hsu, S. H., Chou, Y. C., Lou, W. C. Machine Learning for Early-Onset Dementia Prediction Based on Financial Capabilities: A Comparative Analysis. *Vulnerable groups: inclusion, exclusion and challenges (book chapter)*, 2023.

Chi, V. B.\* & Malle, B. F. People Dynamically Update Trust When Interactively Teaching Robots *ACM/IEEE International Conference on Human-Robot Interaction (HRI)* [25.2% acceptance rate, oral presentation], 2023.

Chi, V. B.\* & Malle, B. F. Calibrated Human-Robot Teaching: What People Do When Teaching Norms to Robots. 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) [oral presentation], 2023.

Malle, B. F., Rosen, E., Chi, V. B.\*, Ramesh, D. What Properties of Norms can we Implement in Robots? 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2023.

Malle, B. F., Rosen, E., Chi, V. B.\*, Ramesh, D. Robots Capable of Norm-Appropriate Action (not Moral Agency). ACM/IEEE International Conference on Human-Robot Interaction (HRI) Perspectives on Moral Agency in Human-Robot Interaction Workshop, 2023.

Malle, B. F., Chi, V. B\*., Rebola, C. B. Older Adults' Understanding of Signals Communicated by Robot Companions for Caregiving *Innovation in Aging*, *Volume 7*, *Issue 1*, 2023.

Chi, V. B.\* & Malle, B. F. When People Teach a Robot, They Systematically Choose Their Strategies Based on Observed Learning Evidence and Emerging Trust Feelings. *International Convention of Psychological Science (ICPS)*, 2023.

**PUBLICATIONS** 

Manuscripts

Hsiung, E., Rosen, E., Chi, V. B.\*, Malle, B. F. Learning Reward Functions from a Combination of Demonstration and Evaluative Feedback ACM/IEEE International Conference on Human-Robot Interaction(HRI), 2022.

Chi, V. B.\* & Malle, B. F. Instruct or Evaluate: How People Choose to Teach Norms to Social Robots. ACM/IEEE International Conference on Human-Robot Interaction(HRI), 2022.

Rosen, E., Hsiung, E., Chi, V. B.\*, Malle, B. F. Norm Learning with Reward Models from Instructive and Evaluative Feedback. IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2022.

Malle, B. F, Austerweil, J. L, Chi, V. B.\*, Kenett, Y., Beck, E. D, Thapa, S., Allaham, M. Cognitive Properties of Norm Representations Proceedings of the Annual Meeting of the Cognitive Science Society, 2021.

Malle, B. F., Rosen, E., Chi, V. B.\*, Berg, M., Haas, P. A General Methodology for Teaching Norms to Social Robots 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2020.

Chi, V. B.\* & Malle, B. F. Towards a Model of Social Norm Activation(In Prep)

Chi, V. B. Teaching Norms to Humans and Robots: a Survey (In Prep)

Chi, V. B.\*, Rébola, C. B., Aydin, Z., Malle, B. F. Intuitively Understandable Communication Signals for a Nonverbal Companion Robot (Under Review)

Chi, V. B.\*, Dennis, E., Salubre, K., Mehrotra, S., Teruhisa, M., Kumar, A. Doing the Right Thing: Fostering Prosocial Behavior through Observational Learning of Social Norms in Future Hybrid Societies (Under Review)

Phillips E., Malle, B.F., Rosero A., Kim M.J., Kim. B., Melles L., Chi, V. B.\* Systematic methods for Moral HRI: A framework and testbeds for studying responses to robot norm conflict (Under Review)

Reviewer

AAAI Fall Symposium on Artificial Intelligence for Human-Robot Interaction

ACM Transactions on Human-Robot Interaction,

IEEE International Conference on Robot and Human Interactive Communication,

ACM/IEEE International Conference on Human-Robot Interaction,

Invited Talks

Morality Lab at Boston College, Chestnut Hill, MA, 2024

Social-Cognitive Seminar Series at Brown University, Providence, RI, 2024

3rd Workshop on Human-Interactive Robot Learning (HIRL), Boulder, CO, 2024

The Gerontological Society of America Annual Scientific Meeting, Tampa, Florida, 2023

First Multidisciplinary Seminar on Social Sciences with Emphasis on Vulnerability and Vulnerable Groups, Universidad de las Américas Puebla, online, 2023

International Convention of Psychological Science, Brussels, Belgium, 2023

Measuring Belief Systems Workshop, Princeton University, Princeton, New Jersey, 2022 22nd meeting of the European Society for Cognitive Psychology, Lille, France, 2022

Association for Psychological Science Virtual Convention, Virtual, 2021

Teaching

CLPS 0950 Intro to Programming Independent Instructor of Record Spring 2023 CLPS 2980 Foundations of Advanced Statistics Spring 2022 CLPS 0220 Making Decisions Spring 2021 CLPS 0900 Statistical Methods

Fall 2020

SKILLS

Programming: Python, R, C++, MATLAB, HTML/CSS, JavaScript

Languages: English, Mandarin, German