

Sarah A. Wu

sarahawu@stanford.edu · <https://sarahawu.github.io/>

Education

Stanford University	2020 –
Ph.D. in Psychology	
Advisor: Tobias Gerstenberg	
Diverse Intelligences Summer Institute	2021
Massachusetts Institute of Technology	2016 – 2020
B.S. in Mathematics with Computer Science; Brain & Cognitive Sciences	

Honors & Awards

Stanford Interdisciplinary Graduate Fellowship	2023 – 2026
Stanford Institute for Research in the Social Sciences Grant	2022, 2021
Best Paper Award, NeurIPS Cooperative AI Workshop	2020
Computational Modeling Prize in Higher Cognition, Cognitive Science Society	2020
NSF Graduate Research Fellowship	2020 – 2023
Phi Beta Kappa	2020
MIT Hans Lukas Teuber Award for Outstanding Academics	2020, 2019
Amgen National Scholar	2018

Publications

Journal Articles

Sarah A. Wu and Tobias Gerstenberg (under review). If not me, then who? Responsibility and replacement.

Sarah A. Wu*, Rose E. Wang*, James A. Evans, Joshua B. Tenenbaum, David C. Parkes, and Max Kleiman-Weiner (2021). Too many cooks: Bayesian inference for coordinating multi-agent collaboration. *Topics in Cognitive Science*, 13(2).

Sarah A. Wu and Edward Gibson (2021). Word order predicts cross-linguistic differences in the production of redundant color and number modifiers. *Cognitive Science*, 45(1).

Book Chapters

Rose E. Wang*, **Sarah A. Wu***, James A. Evans, Joshua B. Tenenbaum, David C. Parkes, and Max Kleiman-Weiner (2021). Too many cooks: Bayesian inference for coordinating multi-agent collaboration. In S. Muggleton and N. Charter (Ed.), *Human-like Machine Intelligence*. Oxford University Press.

Peer-reviewed Conference Proceedings

Sarah A. Wu, Shruti Sridhar, and Tobias Gerstenberg (2023). A computational model of responsibility from counterfactual simulations and intention inferences. *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

Sarah A. Wu, Shruti Sridhar, and Tobias Gerstenberg (2022). That was close! A counterfactual simulation model of causal judgments about decisions. *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

Sarah A. Wu*, Rose E. Wang*, James A. Evans, Joshua B. Tenenbaum, David C. Parkes, and Max Kleiman-Weiner (2020). Too many cooks: Coordinating multi-agent collaboration through inverse planning. *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*.

Conference Presentations

“A computational model of responsibility from counterfactual simulations and intention inferences”
Poster at the 45th Annual Meeting of the Cognitive Science Society (CogSci), 2023

“That was close! A counterfactual simulation model of causal judgments about social agents”

Talk at the 48th Annual Meeting of the Society for Psychology and Philosophy ((E)SPP), 2022
Poster at the 44th Annual Meeting of the Cognitive Science Society (CogSci), 2022
Poster at the Robotics: Science and Systems (RSS) Social Intelligence in Humans and Robots Workshop, 2022

“The role of counterfactual reasoning in responsibility judgments”

Talk at the 47th Annual Meeting of the Society for Psychology and Philosophy (SPP), 2021

Poster at the 43rd Annual Meeting of the Cognitive Science Society (CogSci), 2021

“Too many cooks: Bayesian inference for coordinating multi-agent collaboration”

Spotlight talk & poster at the NeurIPS Cooperative AI Workshop, 2020

“Word order predicts cross-linguistic differences in the production of redundant color and number modifiers”

Talk at the 26th Architectures and Mechanisms for Language Processing (AMLaP), 2020

“Too many cooks: Coordinating multi-agent collaboration through inverse planning”

Talk at the 42nd Annual Meeting of the Cognitive Science Society (CogSci), 2020

Talk at the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), 2020

Invited Presentations

Experimental Jurisprudence Workshop, Michigan Law School

Oct. 2022

Teaching

Teaching Assistant

Stanford PSYCH 205 Foundations of Cognition

Spring 2023

Stanford PSYCH 251 Experimental Methods

Fall 2023, Fall 2022, Fall 2021

Stanford PSYCH 252 Statistical Methods for Social & Behavioral Sciences

Winter 2022

Stanford SYMSYS 1 Minds and Machines

Summer 2021

MIT 6.046 Design and Analysis of Algorithms

Spring 2020, Fall 2019, Spring 2019

MIT 6.036 Introduction to Machine Learning

Fall 2018

Instructor

IIS Curie-Sraffa High School, Milan, Italy (through MIT Global Teaching Labs)

2019

Mentoring

Shruti Sridhar (undergraduate, 2021 –), Gabe Gaw (undergraduate, 2022 –), Siying Zhang (research assistant, 2022 –)

Professional Service & Activities

Conference Reviewing

CogSci 2023, CogSci 2022, CogSci 2021

Workshop Organizing & Program Committee

NeurIPS 2023 AIPsyPhil Workshop, ICML 2023 Counterfactuals Workshop, NeurIPS 2022 nCSI Workshop

Department & University Service

Stanford Psychology Advising Coach

2023 –

Stanford Psychology FriSem Seminar Organizer

2022 – 2023

Stanford Psychology Faculty Meeting Representative

2020 – 2022

MIT Educational Counselor

2020 –

DEI & Outreach

Stanford Future Advancers of Science & Technology (<https://fast.stanford.edu/>)

2020 –

President

2023 –

Chief Program Officer

2022 – 2023

