

# Tim Hua

Website: [timhua.me](http://timhua.me)

## EDUCATION

### Middlebury College

BA, *summa cum laude*, Economics (GPA 3.97/4.00)

Middlebury, VT

May 2023

D.K. Smith Economics Prize for best thesis (*Fox News's Effect on Social and Moral Preferences*).

- Thesis cited by Harvard economist Benjamin Enke in [Moral Boundaries](#)

### Harvard College

Visiting Undergraduate Student (GPA: 4.0/4.0)

Cambridge, MA

2021 - 2022 School Year

- Took coursework in economics, math, and statistics

## AI SAFETY EXPERIENCES

### Neel Nanda and Sam Marks's MATS Scholar

Scholar

Berkeley, CA

Apr 2025 – Current

- Trained model organisms that would act differently in evaluation versus in deployment.
- Elicited the model's deployment behavior in evaluation context through activation steering.
- Paper accepted at Neurips Mechanistic Interpretability workshop and will be on arXiv soon.
- Earlier work studied uncovering "semantic backdoors" in models using interpretability techniques. (Alignment forum post [here](#))

### Tyler Tracy's MARS Mentee in AI Control

Researcher

Berkeley, CA

Jan 2025 – June 2025

- Developed theory and empirical demonstrations on how to optimally combine monitors with different cost and performance profiles (Alignment forum post [here](#)).
- Our paper is accepted as a poster at [Neurips](#).

### Alice Rigg's AI Safety Camp Mentee in Mechanistic Interpretability

Researcher

Remote

Jan 2025 – June 2025

- Studied the functional role of SAE error terms. (Alignment forum post [here](#).)

### Miscellaneous

- Wrote up [AI-Induced Psychosis: A shallow investigation](#). The post received coverage from media outlets ([Business Insider](#), [Futurism](#)) and other bloggers ([Zvi](#), [Tyler Cowen](#))
- BlueDot Economics of Transformative AI Course Facilitator (Feb 2025 Cohort). Attended Stanford DEL's summer course on the [Economics of Transformative AI](#).

## PROFESSIONAL EXPERIENCES

### Walmart Economics Team

Manager, Economist

New York, NY

Intern in summer 2022, full time July 2023 – Aug 2024

Worked on monetary incentives for drivers on Walmart's last mile delivery platform.

- Designed and executed numerous experiments.
- Built causal ML models following [Treatment Allocation Under Uncertain Costs](#).
- Presented work to Walmart chief economist John List, [fifth ranked economist in the world](#).
- Translated technical results into concrete recommendations for business team.
- As an intern, created pricing zones that minimized customers' cross-zone shopping by constructing a weighted graph of Walmart stores and partitioning the graph.

### Brookings Institution

Research Assistant for Professor Carol Graham

Remote

June 2021 – May 2022

- Coauthor on [paper](#) examining mental and physical ill-being during the COVID-19 pandemic.