

# Messi H.J. Lee

[hojunlee@wustl.edu](mailto:hojunlee@wustl.edu) | [lee-messi.github.io](https://lee-messi.github.io) | [OSF](#) | [Github](#) | [Google Scholar](#) | [LinkedIn](#)

## Research Interests

---

**AI Bias** (e.g., **Stereotyping in Large Language and Vision Language Models**),

Behavioral Evaluation of LLMs, Computational Social Science, Text-as-data, Stereotyping, Prejudice

## Education

---

**Washington University in St. Louis**

Exp. 2025

- PhD Candidate in the Division of Computational and Data Sciences
- McDonnell International Scholar
- Graduate Student Affiliate, Center for the Study of Race, Ethnicity & Equity

**Hong Kong University of Science and Technology**

2017-2020

- BSc in Quantitative Social Analysis
- Capstone Project | “Part-time Electoral Campaigns on Social Media”
- Academic Achievement Medal & First Class Honors

## Full Conference & Journal Publications (Peer Reviewed)

---

**Lee, M. H. J.**, Montgomery, J. M., & Lai, C. K. (June 2024). Large Language Models Portray Socially Subordinate Groups as More Homogeneous, Consistent with a Bias Observed in Humans. FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency. [\[ACM Digital Library\]](#) [\[Replication Materials\]](#)

**Lee, M. H. J.**, Montgomery, J. M., & Lai, C. K. (2024). America's racial framework of superiority and Americanness embedded in natural language. *PNAS Nexus*, pgad485. <https://doi.org/10.1093/pnasnexus/pgad485> [\[Replication Materials\]](#) [\[OSF\]](#)

## Work in Progress

---

**Lee, M. H. J.**, Montgomery, J. M., & Lai, C. K. (Submitted). More Distinctively Black and Feminine Faces Lead to Increased Stereotyping in Vision-Language Models. [\[Pre-print\]](#) [\[Replication Materials\]](#)

**Lee, M. H. J.** & Lai, C. K. (Submitted). Probability of Differentiation Reveals Brittleness of Homogeneity Bias in Large Language Models. [\[Pre-print\]](#) [\[Replication Materials\]](#)

**Lee, M. H. J.,** Montgomery, J. M., & Lai, C. K. (In Prep.). Do Large Language Models Amplify Homogeneity Bias in Human Participants?.

**Lee, M. H. J.,** Montgomery, J. M., & Lai, C. K. (In Prep.). How Does Post-Training of Large Language Models Affect Homogeneity Bias?.

## Workshop Publications (Lightly Peer-Reviewed)

---

**Lee, M. H. J.,** Montgomery, J. M., & Lai, C. K. (2023). The Effect of Group Status on the Variability of Group Representations in LLM-generated Text. NeurIPS 2023 Socially Responsible Language Modelling Research (SoLaR) Workshop. [[OpenReview](#)]

## Talks & Presentations

---

**Lee, M. H. J.,** Montgomery, J. M., & Lai, C. K. (June 2024). *Large Language Models Portray Socially Subordinate Groups as More Homogeneous, Consistent with a Bias Observed in Humans*. ACM Conference on Fairness, Accountability, and Transparency 2024 in Rio de Janeiro, Brazil.

**Lee, M. H. J.,** Montgomery, J. M., & Lai, C. K. (May 2024). *Large Language Models Portray Socially Subordinate Groups as More Homogeneous, Consistent with a Bias Observed in Humans*. Poster Presentation @ STL DataFest 2024 in St. Louis, MO.

**Lee, M. H. J.** (Mar 2024). *Large Language Models Portray Socially Subordinate Groups as More Homogeneous, Consistent with a Bias Observed in Humans*. TRIADS “Messy” Data Meeting @ Washington University in St.Louis.

**Lee, M. H. J.,** Lai, C. K., & Montgomery, J. M. (Feb 2024). *Outgroup Homogeneity-like Effect in Large Language Models*. Poster presentation @ The Society of Personality and Social Psychology Conference in San Diego, CA. [[Poster](#)]

**Lee, M. H. J.,** Lai, C. K., & Montgomery, J. M. (Dec 2023). *The Effect of Group Status on the Variability of Group Representations in LLM-generated Text*. Poster presentation @ NeurIPS 2023 Socially Responsible Language Modelling Research (SoLaR) Workshop in New Orleans, LA. [[Poster](#)]

**Lee, M. H. J.,** Montgomery, J. M., & Lai, C. K. (Feb 2023). *A Test of the racial position model in natural language*. Poster presentation @ The Society of Personality and Social Psychology Conference in Atlanta, GA. [[Poster](#)]

**Lee, M. H. J.** (April 2022). *Racial/ethnic stereotypes in natural language*. Social and Personality Psychology Brown Bag @ Washington University in St.Louis.

**Lee, M. H. J., & Lai, C. K.** (Feb 2022). *The spread of stereotypes on social media*. Virtual poster presentation @ The Society of Personality and Social Psychology Conference.

**Lee, M. H. J., Keller, F., & Yang, J.** (Aug 2020). *Central coordination patterns distinguish AstroTurfing from grassroots campaigns*. Virtual poster presentation @ Politics and Computational Social Science & Political Networks.

## Awards

---

<b>Travel Grant</b> ACM Conference on Fairness, Accountability, and Transparency (\$1,000)	2024
<b>OpenAI API Researcher Access Program</b> OpenAI (\$5,000)	2024
<b>Outstanding Research Award</b> The Society for Personality and Social Psychology	2024
<b>Graduate Student Conference Travel Award</b> The Center for the Study of Race, Ethnicity & Equity at WashU (\$750)	2023
<b>Winner and People's Choice of the Digital Transformation Track</b> McDonnell International Scholars Academy (\$3,000)	2023
<b>Small Grant</b> The Center for the Study of Race, Ethnicity & Equity at WashU (\$900)	2022
<b>Diversity Graduate Travel/Registration Award</b> The Society for Personality and Social Psychology (\$500)	2022
<b>Academic Achievement Medal</b> The Hong Kong University of Science and Technology	2020
<b>Undergraduate Research Award Second Runner Up</b> The Hong Kong University of Science and Technology	2020
<b>Scholarship for Continuing Undergraduate Students</b> The Hong Kong University of Science and Technology	2019
<b>Scholarship for Continuing Undergraduate Students</b> Ng Hong Mun Educational Foundation	2018
<b>Outstanding Student Scholarship</b> Ng Hong Mun Educational Foundation	2017

## Teaching

---

### Washington University in St.Louis, Assistant in Instruction

2023

- Text as Data (Instructor: Dr. Brandon Stewart)

### Washington University in St.Louis, Assistant in Instruction

2022-2023

- Quantitative Methods I & II (Instructor: Dr. Shelly Cooper)

## Skills

---

### Natural Language Processing & Text-as-Data

- Behavioral evaluations of Large Language Models
- Fine-tuning (e.g., LoRA) and Prompting Large Language Models for downstream tasks
- Contextualized Embeddings and word embeddings
- Deep learning and natural language processing
- Text-as-data approaches (e.g., Structural Topic Models)

### Computational Social Science & Psychology

- Survey experiments using Qualtrics, CloudResearch, MTurk, Prolific

### Programming Languages and Packages Frequently Used

- python: PyTorch, TensorFlow, HuggingFace, bitsandbytes, peft, sentence-transformers, transformers, openai, scikit-learn, gensim, numpy, pandas
- R: text, tidyverse, dplyr, ggplot2, psych, lme4, lmerTest, afex
- Other Languages: Java, C++, SQL

## Work Experience

---

### Researcher

2022-

- Coordinated data collection efforts for a large-scale study on stereotype threat among Black college students across the United States, aiming to recruit a large sample size (anticipated N = 2,700) from 30+ institutions to enhance the robustness of findings in stereotype threat research.
- Pre-print for the project: <https://osf.io/preprints/psyarxiv/6hju9>

### Research Assistant in Computational Social Science

2019-2020

- Worked on a pivotal project examining coordinated activity patterns of Twitter accounts involved in political astroturfing during the 2012 South Korean Presidential Election.
- Conducted Social Network Analysis of Twitter account activities using Gephi, GraphViz, and R programming language identifying anomalous retweet patterns unique to astroturfing campaigns.
- Uncovered critical insights into the tactics of digital political influence, enhancing the understanding of social media's role in election dynamics.

## **Student Researcher in Research Methods and Political Science**

2019

- Led a research project analyzing sentiment of censored song lyrics during the Third and Fourth Republic of Korea, shedding light on the underlying socio-political sentiments of the time.
- Utilized state-of-the-art text-as-data approaches in Python and R to uncover distinctive sentiment patterns within censored lyrics and to draw correlations between the sentiments and the prevailing socio-political climate, contributing to a deeper understanding of historical censorship impacts.

## **Professional Service**

---

### **Reviewer**

- Association for Computational Linguistics (ACL) Student Research Workshop 2024
- Trustworthy Multi-modal Foundation Models and AI Agents (TiFA) Workshop at ICML 2024
- Personality and Social Psychology Bulletin (March 2024)
- Socially Responsible Language Modelling Research (SoLaR) Workshop at NeurIPS 2023

### **Program Committee**

- Socially Responsible Language Modelling Research (SoLaR) 2.0 Workshop at NeurIPS 2024

### **Conference Volunteer**

- Volunteer at the STL DataFest 2024
- Volunteer at the Political Methodology (PolMeth) Annual Meeting 2022

### **University Service**

- President of the Korean Graduate Student Association (KGSA) 2021-2022
- McDonnell International Scholars Academy Information Session @HKUST 2023

## **Research Coverage & Recognition**

---

Department of Political Science, Washington University in St.Louis ([Link](#))

2024

Center for the Study of Race, Ethnicity & Equity ([Link](#))

2024