Yuxuan (Effie) Li

liyuxuan@stanford.edu | https://effie-li.github.io

- 5+ years of research experience in deep learning, machine learning, cognitive science, neuroscience
- Research interests: machine and human intelligence, human-like AI, interpretability

Education

2019 – 2024 **Stanford University, PhD** in Cognitive Psychology (expected Dec 2024). Transcript.
- Weiland Fellow, School of Humanities and Sciences
- Alumna, Center for Mind, Brain, Computation and Technology

2013 – 2017 **Trinity College, BS** in Computer Science and Psychology. *summa cum laude.*

Research Experience

2024 summer **Benchmarking multi-modal LLMs** @ Meta
Yuxuan Li, Vijay Veerabadran, Michael Iuzzolino, Asli Celikyilmaz, Karl Ridgeway. *in prep.*

2023 - 2024 **A learning theory of subgoal choices in transformers and humans** @ Stanford
Yuxuan Li, James McClelland. *CCN24 poster.*
- Developed a data-distributional, learning theory of task decomposition on graphs
- Trained transformers on graph traversal and found human-like subgoal choices

2022 - 2023 **Multi-task learning and systematic generalization in transformers** @ Stanford
Yuxuan Li, James McClelland. *TMLR paper*, *code*.
- Proposed a new encoding method that boosts length generalization in transformers
- Analyzed behavior and representations in transformers trained on algorithmic tasks, leading to insights on emergent multi-task computation in multi-head attention

2023 summer **Representation learning for embodied planning** @ Allen Institute for AI
Yuxuan Li, Luca Weihs. *NeurIPS 2023 Workshop.*
- Evaluated self-supervised goal-directed pretraining objectives for embodied agents

2019 - 2021 **Human goal-directed decision making** @ Stanford
Yuxuan Li, James McClelland. *PLOS paper*, *code*.
- Conducted behavioral studies and built models of human planning processes

2017 – 2019 **Neural decoder of human episodic memory** @ UPenn
Yuxuan Li, Jesse Pazdera, Michael Kahana. *NatComm paper, code.*
- Developed a novel data sampling method and trained neural decoders from large-scale EEG time series data, yielding new insights into human memory

Technical Skills

Programming **Python, R,** some experience with MATLAB, HTML/CSS/JavaScript

Packages **LLM** *(langchain),* **deep learning** (pytorch, pytorch-lightning, allenact, einops), **experiment/server management** (wandb, beaker), **machine learning** (scikit-learn), **data analysis** (scipy, numpy, pandas), **data visualization** (matplotlib), **cognitive** (**neuro**)**science** (mne, ptsa)

Other LaTeX, statistics (linear modeling, generalized linear modeling, mixed-effects models), representation analysis, online behavioral platforms (Amazon MTurk, Prolific)

Publications and Preprints

2024 **Li, Y.**, & McClelland, J.L. Emergent human-like path preferences and implicit subgoal selection in transformers learning graph traversal. *Cognitive Computational Neuroscience.*

2024 **Li, Y.**, Pazdera, J.K., & Kahana, M.J. EEG decoders track memory dynamics. *Nature Communications.*

2023 **Li, Y.**, & Weihs, L. Understanding representations pretrained with auxiliary losses for embodied agent planning. *NeurIPS 2023 Generalization in Planning Workshop.*

2023 **Li, Y.**, & McClelland, J.L. Representations and computations in transformers that support generalization on structured tasks. *Transactions on Machine Learning Research.*

2023 Kahana, M.J., Lohnas, L.J., Healey, K., . . ., **Li, Y.**, . . ., & Weidemann, C.T. The Penn Electrophysiology of Encoding and Retrieval Study. *JEP: LMC*.

2022 **Li, Y.**, & McClelland, J.L. A weighted constraint satisfaction approach to human goal-directed decision making. *PLOS Computational Biology*.

2022 Katerman, B.S., **Li, Y.**, Pazdera, J.K., Keane, C., & Kahana, M.J. EEG biomarkers of free recall. *NeuroImage*.

2018 Grubb, M.A., & **Li, Y.** Assessing the role of accuracy-based feedback in value-driven attentional capture. *Attention, Perception, & Psychophysics*.

Talks and Presentations

Mar 2024 **Li, Y.** Emergent structured computation from learning and its implications for cognitive science and AI. *Microsoft Research Lab – Redmond.*

Nov 2023 **Li, Y.** Systematic generalization and emergent structures in transformers trained on structured tasks. *FriSem seminar, Department of Psychology, Stanford University.*

Apr 2022 **Li, Y.** A weighted constraint satisfaction approach to human goal-directed decision making. *Cognitive Tools Lab, University of California, San Diego.*

Feb 2021 **Li, Y.** Model-based reinforcement learning and the reinforcement learning framework for human behavior. *TA Lecture in PSYCH 209, Stanford University.*

2020, 2021 **Li, Y.** Building online psychology experiments with jsPsych: a tutorial. *TA Lecture in PSYCH 251, Stanford University.*

Honors and Awards

2022 – 2024 Ric Weiland Graduate Fellowship in the Humanities & Sciences. Stanford University.

2013 – 2017 Phi Beta Kappa, Dean's Scholar (top 5%), Faculty Honors, Holland Scholar. Trinity College.

Teaching and Services

Reviewer CogSci 2022 – CCN 2024 – NeurIPS 2024 –

Committee Cognitive Neuroscience Seminar organizing committee, Stanford Psychology, 2021 – 2022

TA Neural network models of cognition, brain decoding, Experimental methods, developmental psychology, introduction to computing, mathematical foundations of computing