Violet Xiang

ziyxiang@stanford.edu | violetxi.github.io | github.com/violetxi | Google Scholar

Education

Stanford University Expected Sep 2026
Ph.D. in Psychology
Research focus: LLM post-training, reinforcement learning, reward modeling, and computational cognitive science. Stanford Autonomous Agents Lab, advised by Prof. Nick Haber.

Indiana University 2018–2020
M.S. in Computer Science

Indiana University 2011–2016
B.S. in Informatics, B.A. in Mathematics

Industry Research

Radical Numerics Oct 2025–Jan 2026
Member of Technical Staff Intern
- Built an end-to-end post-training and evaluation stack for DPO-tuning genomic foundation models, including preference-data processing, training orchestration, and evaluation harnesses for biomedical DNA-sequence design.

Toyota Research Institute Jun-Sep 2025
Research Scientist Intern, Human-AI Interactive Learning
- Built a teacher-training objective for corrective feedback in which teacher outputs were scored by how much they raised a frozen student's likelihood of the reference solution.
- Established a reliable final-answer reward signal and characterized why it failed to transfer to long reasoning traces.

Publications

*Equal contribution.

LLM Post-Training, Reward Modeling, and Evaluation

V Xiang, A Setlur, C Blagden, N Haber, A Kumar. ExpRL: Exploratory RL for LLM Mid-Training. *COLM 2026* | *code*.
- Designed and open-sourced ExpRL: an RL-based mid-training method using reference solutions as reward scaffolds for an LLM judge that scores partial progress with dense outcome- and process-level rewards for online RL.
- Beat the strongest baseline at each stage on held-out AIME 2026 pass@1 (Qwen3-4B) by 3.6 points after ExpRL priming alone and by 4.7 points after identical sparse-GRPO post-training.

D Fein*, M Lamparth*, **V Xiang**, M J Kochenderfer, N Haber. One Bias After Another: Mechanistic Reward Shaping and Persistent Biases in Language Reward Models. *ICML 2026* | *code*.
- Identified five reward-model biases (length, uncertainty, position, sycophancy, model-style) across leading Bradley-Terry (BT) RMs.
- All five biases persist: 22.6-point accuracy drops on hedged correct answers, up to 60% agreement with wrong user suggestions, and positional swings up to 28 points.
- Reduced the length, uncertainty, and position biases via null-space probe projections without retraining or RewardBench-2 loss.

D Fein*, S Russo*, **V Xiang***, K Jolly, R Rafailov, N Haber. LitBench: A Benchmark and Dataset for Reliable Evaluation of Creative Writing. *EACL 2026, Long Paper* | *dataset*.
- Curated LitBench, the first standardized benchmark for creative-writing evaluation: a 43k-pair training corpus and a 2.5k-pair debiased, human-labeled test set.
- Trained BT and generative reward models that beat the strongest zero-shot judge by 5 points in human-preference agreement.
- Validated reward-model rankings via an online human study on newly LLM-generated stories.

V Xiang, C Blagden, R Rafailov, N Lile, S Truong, C Finn, N Haber. Just Enough Thinking: Efficient Reasoning with Adaptive Length Penalties Reinforcement Learning. *NeurIPS 2025 Efficient Reasoning Workshop Spotlight.*
- Developed a GRPO-compatible adaptive length penalty for DeepScaleR-1.5B that reduced generation length by more than 50% at matched accuracy across MATH-500, AIME, and OlympiadBench while allocating 5.35× more tokens to hard than easy prompts.

T Hua, H Hua, **V Xiang**, B Klieger, S Truong, W Liang, F-Y Sun, N Haber. ResearchCodeBench: Benchmarking LLMs on Implementing Novel Machine Learning Research Code. *NeurIPS 2025 Datasets & Benchmarks Track, Spotlight* | *website*.
- Co-developed an open-source execution harness for 212 challenges from 20 recent ML papers; across 32 evaluated models, the strongest reached 33.0% on the HARD split, and 43 tasks remained unsolved by every model.

V Xiang, C Snell, K Gandhi, A Albalak, A Singh, C Blagden, D Phung, R Rafailov, N Lile, D Mahan, L Castricato, J-P Franken, N Haber, C Finn. Towards System 2 Reasoning in LLMs: Learning How to Think With Meta Chain-of-Thought. *Preprint, 2025.*
- Proposed a training roadmap combining linearized search traces, process supervision, synthetic data, and reinforcement learning, supported by evidence of in-context search behavior in frontier reasoning models.

Multi-Agent Learning

L Cross, **V Xiang**, A Bhatia, D L K Yamins, N Haber. Hypothetical Minds: Scaffolding Theory of Mind for Multi-Agent Tasks with Large Language Models. *ICLR 2025, Poster.*

- Contributed to the design and implementation of an LLM Theory-of-Mind agent that outperformed MARL baselines on 3 of 4 environments and 6 of 8 eight-agent scenarios across 30 Melting Pot tasks.

V Xiang, L Cross, J-P Fränken, N Haber. From Centralized to Self-Supervised: Pursuing Realistic Multi-Agent Reinforcement Learning. *NeurIPS 2023 ALOE Workshop, Poster.*

- Built and compared centralized, decentralized, asynchronous, and self-supervised MARL pipelines, showing how centralized-information assumptions can overstate performance in realistic mixed-motive settings.

Human and Machine Learning

B Long*, R Sparks*, **V Xiang***, S Stojanov*, et al.. The BabyView Dataset: High-Resolution Egocentric Videos of Infants' and Young Children's Everyday Experiences. *Computational Cognitive Neuroscience 2025 Proceedings.*

- Co-developed an 868-hour longitudinal egocentric-video dataset spanning ages 5 months to 3 years, with gold-standard annotations and evaluation pipelines for speech transcription, diarization, pose estimation, and out-of-distribution transfer.

C Zhuang*, **V Xiang***, Y Bai, X Jia, N Turk-Browne, K Norman, J J DiCarlo, D L K Yamins. How Well Do Unsupervised Learning Algorithms Model Human Real-Time and Lifelong Learning?. *NeurIPS 2022 Datasets & Benchmarks Track.*

- Developed naturalistic real-time and lifelong visual-learning benchmarks and identified memory mechanisms as critical for learning from sparse, low-diversity input streams.

B Long, S Goodin, G Kachergis, V A Marchman, S F Radwan, R Z Sparks, **V Xiang**, et al.. The BabyView Camera: Designing a New Head-Mounted Camera to Capture Children's Early Social and Visual Environments. *Behavior Research Methods 56, 2024.*

L Yuan, **V Xiang**, D Crandall, L Smith. Learning the Generative Principles of a Symbol System from Limited Examples. *Cognition, 2020.*

Privacy and Security

Y Liu, **V Xiang**, E J Seong, A Kapadia, D S Williamson. Defending Against Microphone-Based Attacks with Personalized Noise. *Proceedings on Privacy Enhancing Technologies, 2021.*

Technical Skills

Library & Tools: FSDP, Docker, vLLM, SGLang, Ray/RLlib, SLURM

Deep Learning Frameworks: PyTorch, Tensorflow, Keras

Programming Languages: Python, C++, C, JavaScript

Teaching

2023-25 Teaching Assistant, Neural Network Models of Cognition, Stanford University
2023 Teaching Assistant, Introduction to Statistical Methods, Stanford University
2021 Teaching Assistant, Longitudinal Design and Data Analysis, Stanford University
2019 Mentor, Proactive Health Informatics REU, Indiana University
2019 Teaching Assistant, Computer Vision, Indiana University
2018 Teaching Assistant, Introduction to Data Structures and Algorithms, Indiana University