Nona Ghazizadeh

(+1) 310-871-2106 | nghaziza@usc.edu, nonaaghazizadeh@gmail.com | nona-ghazizadeh | Nona Ghazizadeh

Education

University of Southern California *August 2024 - present*

Direct Ph.D. Track in Psychology - Brain and Cognitive Science *Los Angeles, CA*

- Advisor: **Morteza Dehghani**
- Current **GPA: 4.0**/4.0

University of Southern California *August 2024 - May 2026*

MS.c. in Psychology - Brain and Cognitive Science *Los Angeles, CA*

- Thesis Title: Alignment Needs 'Cognitive Control': On The Role of Regularization in LLM Alignment
- Current **GPA: 4.0**/4.0

Sharif University of Technology *September 2019 - July 2024*

B.Sc. in Computer Engineering *Tehran, Iran*

- Overall **GPA: 3.85**/4.0 (18.70/20)
- Major **GPA: 3.90**/4.0 (19.00/20)

Research Interests

- **Large Language Model Reasoning**
- **Computational Cognitive Science**
- **Natural Language Processing**

Publications

- **N. Ghazizadeh**, E. Rahmati, M. Dehghani, and P. Piray. Beyond a Penalty: Regularization as Cognitive Control in Alignment. NeurIPS 2026 - Position Paper Track. (Accepted)
- A. S. Ziabari*, **N. Ghazizadeh***, Z. Sourati, F. Karimi-Malekabadi, P. Piray, and M. Dehghani. Reasoning on a Spectrum: Aligning LLMs to System 1 and System 2 Thinking. COLM 2026
- **N. Ghazizadeh**, E. Rahmati, A. S. Ziabari, K. Ahrabian, P. Piray, and M. Dehghani. DIALECT: Continuous-Space Communication Within and Across Language Models. Under review at ICLR 2027, 2026
- E. Rahmati, **N. Ghazizadeh**, Z. Sourati, N. Rouhani, and M. Dehghani. Structural Abstraction as an Inductive Bias for Non-Stationary Language Model Training. Under review at ICLR 2027, 2026
- M. M. Abootorabi, **N. Ghazizadeh**, A. Dalili, A. Ghahramani, M. Dehghani, and E. Asgari. AIMA at SemEval-2024 Task 10: History-Based Emotion Recognition in Hindi-English Code-Mixed Conversations. In *Proceedings of the 18th International Workshop on Semantic Evaluation*, Mexico City, Mexico, June 2024. ACL
- M. Akhi and **N. Ghazizadeh**. HDNA: A graph-based change detection in HTML pages(Deface Attack Detection), 2023
- Z. Sourati*, M. H. Wu*, **N. Ghazizadeh**, J. Kaplan, M. Dehghani, and S. A. Nastase. Cognitive Expert Language Models Better Align with the Corresponding Brain Systems. Under review at ICLR 2027, 2026
- A. Ghahramani Kure, M. Dehghani, M. M. Abootorabi, **N. Ghazizadeh**, S. A. Dalili, and E. Asgari. AIMA at SemEval-2024 Task 3: Simple Yet Powerful Emotion Cause Pair Analysis. In *Proceedings of the 18th International Workshop on Semantic Evaluation*, Mexico City, Mexico, June 2024. ACL
- P. Golazizian, E. Rahmati, J. Trager, Z. Sourati, **N. Ghazizadeh**, …, and M. Dehghani. The Subjectivity of Respect in Police Traffic Stops: Modeling Community Perspectives in Body-Worn Camera Footage. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics*, San Diego, California, July 2026. ACL

Honors & Awards

2026	Received seed funding award from Center for Computational Language Sciences at University of Southern California	*Los Angeles, CA*
2026	Received doctoral research grant award from psychology Department at University of Southern California	*Los Angeles, CA*
2025	Received graduate student research fellowship from University of Southern California	*Los Angeles, CA*
2024	Ranked in the top 25% among 190 students of Computer Engineering students based on GPA at Sharif University of Technology.	*Tehran, Iran*
2019	Ranked in the top 0.5% among nearly 165,000 participants in Nation-Wide University Entrance Exam (Konkour) in Mathematics Branch for fully funded BSc period in Iran.	*Tehran, Iran*

Work & Research Experience

Research Assistant at University of Southern California

August 2024 - present

Prof. Dehghani

Los Angeles, CA

- Modeling human fast thinking (System 1) and slow thinking (System 2) to understand cognitive biases in large language models, with the goal of manipulating human behavior based on their thinking type, testing on different benchmarks.
- Investigated the mechanistic representations of Gentner's Structure-Mapping Theory in LLMs by constructing a multi-domain analogical reasoning dataset and extracting structural mapping vectors through contrastive activation analysis.
- Connecting cognitive neuroscience and LLM alignment, showing that reference-anchored regularization and cognitive control share the same formal objective and failure modes, and proposing a reframing of alignment research toward designing regularization structure rather than minimizing it.
- Developing a cognitive-inspired reasoning framework using reflection token tagging and augmented language model training by integrating eye-tracking data to identify decision-making points of cognitive reflection in order to enhance LLM reasoning performance and efficiency.
- Designed Abstraction-Augmented Training (AAT), a replay-free online continual learning method that improves stability and generalization via abstraction-aware loss optimization; introduced Relational Cycle and Narrative Abstraction benchmarks and achieved parity or gains over strong experience replay baselines.
- Developed a perspective-aware modeling framework and a multi-perspective dataset of LAPD body-worn camera footage to analyze subjective interpretations of respect. Integrated procedural justice-grounded rubrics and preference data synthesis via SFT and DPO to personalize respect ratings and align model rationales with diverse community lived experiences.

Sharif University of Technology

August 2022 - July 2024

Research Assistant

Dr. Asgari & Prof. Heydarnoori

- Classified emotions in code-mixed conversations by leveraging pre-trained large models and GRU networks to integrate prior and future context as well as the sequential flow of the conversation.
- Extracted emotion-cause pairs in multi-modal conversational contexts by constructing a causality matrix using a transformer-based encoder and applying question answering and prompt engineering to identify causal text segments.
- Implemented a Retrieval-Augmented Generation (RAG) model for a multilingual medical and drug question-answering system.
- Developed a Retrieval-Augmented Generation (RAG) model for a multilingual news question-answering system.
- Ranked GitHub project issues using software engineering methods, creating a call graph with traceability and applying NLP techniques to obtain embeddings and assess their similarities.
- Developed a novel approach for node identification and weighted tree structures for HTML pages, calculating differences in DOM trees to capture changes even in dynamically generated content.

University of New South Wales Sydney

June 2023 - Sep 2023

Research Assistant

Dr. Razzak

- Developed an automated medical report generation system for Fundus Fluorescein Angiography images by using a visual feature extractor and cross-modal memory to align visual and textual features, with reinforcement learning applied to optimize report generation.

Yektanet Technology Company

Apr. 2021 - Sep. 2021

Front-end developer

- Designing and implementing numerous web pages for email marketing and designing and implementing a full SMS marketing and Implementing some UI components for usual use in other teams

Teaching Assistant Experience

Teaching Assistant at University of Southern California

Los Angeles, CA

Statistics

Aug. 2024–May 2025; Aug. 2026–Present

- Hold lab classes, Designing assignments, Correcting assignments, Designing lab handouts, Correcting midterm & final exams

Teaching Assistant at Sharif University of Technology

Sep. 2020 - July. 2024

Modern Information Retrieval, Linear Algebra, Engineering Probability and Statistics, Compiler Design, Electric and Electronic Circuits, Advanced Programming

Tehran, Iran

- Hold problem-solving classes, Designing assignments & projects & quizzes, Correcting assignments & projects & quizzes, Correcting midterm & final exams

Academic Service

- Reviewer — International Conference on Learning Representations (ICLR 2027; ongoing).
- Reviewer — Annual Meeting of the Association for Computational Linguistics (ACL 2026).
- Reviewer — Conference on Language Modeling (COLM 2026).
- Reviewer — Conference on Neural Information Processing Systems (NeurIPS 2026).

Skills

Programming Languages	Python, R, Java, JavaScript
Libraries	NumPy, PyTorch, Scikit-Learn, Tenserflow,Transformers, NLTK, SciPy, Pandas, matplotlib
Frameworks	React, Vue
Databases	PostgreSQL, Redis
Operating Systems	Ubuntu, MacOS, Windows
Other Technologies	Git, Nginx, Jupyter Notebook, Latex, Docker
Soft Skills	Teamwork, Flexibility, Responsibility, Self-Learning, Desire to learn, Problem solving, Creativity
Languages	English, Persian