Guy Davidson

guy.davidson@nyu.edu | +1 415-510-9167 | https://github.com/guydav | https://guydavidson.me

I am a cognitive scientist excited about understanding the human mind and leveraging ideas from human cognition to develop more human-like artificial intelligence. My research combines human experiments, data analysis, and computational modeling to study cognition. I am particularly interested in cognitive goals and the use of structured, program-like representations to capture them. I have also led work on cognitively-inspired machine learning and contributed to more theoretical psychology efforts.

EDUCATION

NYU CENTER FOR DATA SCIENCE New York, NY
Ph.D. candidate 2019–Present
Advised by Professors Brenden Lake and Todd Gureckis, my thesis research attempts to advance the understanding of goals in human cognition. My work proposes a framework to represent goals as reward-producing programs and supports this idea by combining evidence from rich human experiments and computational models. My research offers a novel treatment of goals in cognitive science and aims to contribute to the development of richer goal representations for artificial agents.

BRAINS, MINDS, AND MACHINES SUMMER COURSE Woods Hole, MA
Hosted by the MIT and Harvard Center for Brains, Minds, and Machines 08/2021

MACHINE LEARNING SUMMER SCHOOL London, England
Hosted by University College London and Imperial College London 07/2019

MINERVA UNIVERSITY San Francisco, CA
BSc in Computational Sciences 2015–2019
I graduated summa cum laude with a concentration in Machine Learning **GPA: 3.98/4.0**. In my **capstone project**, I investigated the scaling behavior of different meta-learning algorithms: how quickly new tasks are learned as a function of previous training, using a novel benchmark paradigm inspired by visual question answering (externally advised by Mike Mozer, published at CVPR 2020).

PUBLICATIONS AND PRESENTATIONS

Davidson, G., Todd, G., Togelius, J., Gureckis, T. M., Lake, B. M. (2024). Goals as Reward-Producing Programs. *Preprint. Soon to be submitted to Nature Machine Intelligence*: `https://guydavidson.me/files/goals_as_programs.pdf`

Davidson, G., Orhan, A. E., Lake, B. M. (2024). Spatial Relation Categorization in Infants and Deep Neural Networks. *Cognition*.

Sharma, S. **Davidson, G.**, Khetarpal, K., Kanervisto, A., Arora, U., Hofmann, K., Momennejad, I. (2024). Toward Human-AI Alignment in Large-Scale Multi-Player Games. *ArXiv*.

Davidson, G., Todd, G., Gureckis, T. M., Togelius, J., Lake, B. M. (2023). Generating Human-Like Goals by Synthesizing Reward-Producing Programs. *Intrinsically Motivated Open-ended Learning Workshop @ NeurIPS 2023*.

Davidson, G., Gureckis, T. M., & Lake, B. M. (2022). Creativity, Compositionality, and Common Sense in Human Goal Generation. *Proceedings of the 44th Annual Meeting of the Cognitive Science Society, CogSci 2022*.

Davidson, G., Lake, B. M. (2021). Examining Infant Relation Categorization Through Deep Neural Networks. *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society, CogSci 2021*.

Bennett, D., **Davidson, G.**, & Niv, Y. (2021). A model of mood as integrated advantage. *Psychological Review*.

Davidson, G., Lake, B. M. (2020). Investigating Simple Object Representations in Model-Free Deep Reinforcement Learning. *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society, CogSci 2020*.

Davidson, G., Lake, B. M. (2020). Systematically Comparing Neural Network Architectures in Relation Leaning. *Object-Oriented Learning (OOL): Perception, Representation, and Reasoning Workshop @ ICML 2020*.

Davidson, G., Mozer, M. C. (2020). Sequential mastery of multiple visual tasks: Networks naturally learn to learn and forget to forget. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Davidson, G., Radulescu, A., & Niv, Y. (2019). Contrasting the effects of prospective attention and retrospective decay in representation learning. *The 4th Multidisciplinary Conference on Reinforcement Learning and Decision Making*.

Bennett, D., **Davidson, G.**, & Niv, Y. (2019). Momentum and mood in policy-gradient reinforcement learning. *The 4th Multidisciplinary Conference on Reinforcement Learning and Decision Making*.

SELECTED INVITED TALKS

CoLab, Princeton University 05/2024

Invited talk, Cognitive Control of Action Workshop, Princeton University 03/2024

CoCoDev Lab, Harvard University 02/2024

CoCoSci Lab, Massachusetts Institute of Technology 02/2024

TEACHING

DS-GA 1016: COMPUTATIONAL COGNITIVE SCIENCE NYU

Section leader, grader 01/2021-05/2021, 01/2022-05/2022

Served as a section leader and grader for graduate-level course focused on computational approaches to modeling cognition.

DS-UA 112: INTRODUCTION TO DATA SCIENCE NYU

Section leader 09/2019-12/2019

Served as a section leader for new undergraduate course introducing students to NYU's newly approved Data Science major.

RESEARCH & PROFESSIONAL EXPERIENCE

MICROSOFT RESEARCH New York, NY

Research Intern 05/2022–08/2022

Developed methods inspired by the cognitive psychology concept of task-sets (abstract task representations) to analyze and predict behavior in a large-scale gameplay dataset in a multiplayer game. Initial results highlighted consistent differences between players by their propensity to flee or attack in fight-or-flight scenarios. Mentored by Ida Momennejad and Harm van Seijen.

PRINCETON NEUROSCIENCE INSTITUTE Princeton, NJ

Research Intern 05/2018–08/2018

Interned with Professor Yael Niv's lab, to investigate human reinforcement learning (RL) in multidimensional environments:

- Modeled data from previous experiments, making discoveries regarding the dissociable roles of attention and decay in human RL, and the efficacy of eye-tracking and fMRI-based attention measures. Project presented at RLDM 2019.
- Implemented a reinforcement learning experiment in a customizable web platform, enabling data collection using Amazon Mechanical Turk and building a framework used by several current lab members to develop new experiments.
- Developed a simulation environment for bandit problems to motivate work framing mood as a momentum variable.

AIDOC MEDICAL Tel Aviv, Israel

Research Engineer 05/2017–11/2017

- Implemented research-supporting tools in Azure cloud environment to facilitate and expedite deep learning experimentation, reducing idea-to-experiment turnaround 10x from 1-2 hours to 5-10 minutes.

AMAZON PRIME AIR Seattle, WA

Software Development Engineer Intern 05/2016–08/2016

MINERVA PROJECT San Francisco, CA

Software Development Engineer Intern 01/2016–05/2016

SIMILARWEB Tel Aviv, Israel

Software Engineer 06/2015–09/2015

ISRAEL DEFENSE FORCES INTELLIGENCE BRANCH Israel

Team leader, training instructor, software engineer 08/2008–10/2014

- Team leader: grew team of four software developers to ten, responsible for developing tools to enhance analysis capabilities and solve production issues. Reduced turnaround time more than 2x, contributed to Israel Defense Prize-winning project.
- Instructor: managed two other instructors, training ten pupils in software engineering and computer networking.
- Software engineer: built in-house data ingestion pipelines and analysis tools. Developed primarily in Python and Java and cultivated debugging and fault analysis expertise.

SKILLS AND TECHNICAL COMPETENCIES

- Neural network research and development (mostly PyTorch)
- Data science and analysis (Numpy, Pandas, Sklearn, Matplotlib)
- Software engineering, architecture design, and testing
- Human participant research experiment design
- Web experiment development (React, vue, tailwind, Firebase)
- Research environment and stimulus design (Unity, Blender)