

Sam Lau

PhD Candidate
Department of Cognitive Science
University of California, San Diego (UCSD)

Updated: October 12, 2022
Email: lau@ucsd.edu
<https://samlau.me>

RESEARCH INTERESTS Human-computer interaction, end-user programming, data science education

EDUCATION **University of California, San Diego**

Ph.D. Cognitive Science 2018 – 2023 (expected)
Advisor: Philip Guo

University of California, Berkeley

M.S. Computer Science 2017 – 2018
Advisor: Joshua Hug
GPA: 3.93

B.S. Electrical Engineering and Computer Science 2013 – 2017
GPA: 3.96

AWARDS AND HONORS UCSD Cognitive Science Teaching Excellence Award 2019

UC Berkeley EECS Distinguished Graduate Student Instructor Award 2018

TEACHING EXPERIENCE **Instructor**

UCSD DSC 10: Principles of Data Science Summer 2022
100% student recommendation rate, 44% response rate

Berkeley Data 100: Principles and Techniques of Data Science Summer 2019
Teaching rated 6.2 / 7.0 (dept avg 5.8), 92% response rate
First Berkeley summer offering of Data 100

Berkeley Data 8: Foundations of Data Science Summer 2017
Teaching rated 6.3 / 7.0 (dept avg 5.8), 84% response rate
First Berkeley summer offering of Data 8

Graduate Teaching Assistant

UCSD COGS 18: Introduction to Python Winter 2022

UCSD COGS 124: HCI Technical Systems Research Fall 2020

UCSD COGS 108: Data Science in Practice Fall 2019, Winter 2019, Winter 2020

UCSD COGS 10: Cognitive Consequences of Technology Spring 2019

Berkeley Data 100: Principles and Techniques of Data Science Spring 2017, Fall 2017

Berkeley Data 8: Foundations of Data Science Fall 2016, Spring 2016, Fall 2015

Berkeley CS 169: Software Engineering

Spring 2015

Berkeley CS 61AS: Structure and Interpretation
of Computer Programs

Spring 2014, Fall 2014

BOOKS AND
MONOGRAPHS

Textbooks

- [B.1] Samuel Lau, Joseph Gonzalez, Deborah Nolan. Principles and Techniques of Data Science. O'Reilly Media, 2023 (forthcoming). www.textbook.ds100.org.
One of the first textbooks to cover the entire data science lifecycle. Used by 2,000 Berkeley students annually with an additional 30,000 readers from 145 countries.

PEER-
REVIEWED
PUBLICATIONS

Conference Papers

- [C.4] Sam Lau, Justin Eldridge, Shannon Ellis, Aaron Fraenkel, Marina Langlois, Suraj Rampure, Janine Tiefenbruck, Philip Guo. The Challenges of Evolving Technical Courses at Scale: Four Case Studies of Updating Large Data Science Courses. ACM Conference on Learning @ Scale (L@S), 2022.
Instructors of large technical courses run into major logistical challenges that take time away from teaching.
- [C.3] Sam Lau, Deborah Nolan, Joseph Gonzalez, Philip Guo. How Computer Science and Statistics Instructors Approach Data Science Pedagogy Differently: Three Case Studies. ACM Technical Symposium on Computer Science Education (SIGCSE), 2022.
Real-world case studies show tradeoffs in balancing computer science and statistics in data science teaching.
- [C.2] Sam Lau, Sruti Srinivasa Ragavan, Ken Milne, Titus Barik, Advait Sarkar. TweakIt: Supporting End-User Programmers Who Transmogrify Code. ACM Conference on Human Factors in Computing Systems (CHI), 2021.
Placing live previews of code outputs directly in spreadsheets enables data analysts to tweak and reuse Python examples without needing Python expertise.
- [C.1] Sam Lau, Ian Drosos, Julia M. Markel, Philip J. Guo. The Design Space of Computational Notebooks: An Analysis of 60 Systems in Academia and Industry. IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC), 2020.
Computational notebooks vary widely in data import, code editing, code execution, and output format.

Workshop and Poster Papers

- [P.3] Sam Lau, Philip Guo. Data Theater: A Live Programming Environment for Prototyping Data-Driven Explorable Explanations. Workshop on Live Programming (LIVE), 2020

Separating logic from presentation simplifies the process of creating explorable explanations.

- [P.2] Samuel Lau, Tricia J. Ngoon, Vineet Pandey, Scott Klemmer. Experiment Reconstruction Reduces Fixation on Surface Details of Explanations. Poster in Proceedings of C&C 2019: ACM SIGCHI Conference on Creativity and Cognition, 2019

Asking people to mentally replicate an experiment briefly reduces the allure of scientific terminology.

- [P.1] Vinitra Swamy, Allen Guo, Samuel Lau, Wilton Wu, Madeline Wu, Zachary Pardos, David Culler. Deep Knowledge Tracing for Free-Form Student Code Progression. Poster in Proceedings of AIED 2018: *International Conference on Artificial Intelligence in Education*, June 2018

Deep learning models trained on free-form student code predict learning pace.

Journal Articles

- [J.1] Shou-Tian Zheng, Xiang Zhao, Samuel Lau, Addis Fuhr, Pingyun Feng, Xianhui Bu. Entrapment of metal clusters in metal-organic framework channels by extended hooks anchored at open metal sites. In JACS: *Journal of the American Chemical Society*, 2013.

OTHER EMPLOYMENT AND PROJECTS

Berkeley Institute of Data Science, Berkeley, CA 01/2017 - 06/2017

Student Research Engineer – designed distributed infrastructure for hosted computational notebooks. Architecture now used at multiple universities, including UC Berkeley and UC San Diego.

Counsyl, San Francisco, CA 05/2016 - 08/2016

Software Engineering Intern – designed and built appointment scheduling web application.

Khan Academy, Mountain View, CA 05/2015 - 08/2015

Software Engineering Intern – worked on article authoring system for instructors to make interactive content. As of 2020, used for over 95% of articles on Khan Academy.

Berkeley Public Schools Fund, Berkeley, CA 08/2013 - 06/2014

Software Engineering Intern – built crowdfunding system used to raise over \$66,000 for 20 Berkeley public schools.