PSYC35880 COGS24533

Computational Social Cognition

Spring 2026

Instructor:	Prof. Xuechunzi Bai	**Time:**	Tue & Thu 9:30 – 10:50
Email:	baix@uchicago.edu	**Location:**	Green Hall 117
Office:	Pick Hall 116	**Office Hour:**	Friday 13:00 – 13:50

Objective

How do people make sense of the social world? Why do first impressions form so quickly, attitudes change slowly, and groups sometimes become wiser, or more biased, than the individuals within them? The goal of this course is to introduce students to the science of how people think, judge, and decide in social situations, and to show how these processes can be studied with simple computational ideas. We will explore topics such as social perception, categorization, attitude change, decision making, and collective behavior, asking not only what people do, but also how these patterns can be explained and predicted. The course combines classic theories in social psychology with computational tools from cognitive science. No prior programming experience is required; the emphasis is on learning how to break complex social behavior into clear, testable components. By the end of the course, students will gain a foundation for reverse-engineering social decisions and begin building simple formal models of social cognition themselves.

Structure

This course is organized as a seminar. We will usually spend two class sessions on each broad topic. The first session introduces the conceptual background, major questions, and historical development of the topic through a short lecture. The second session focuses on recent empirical work, with students presenting selected papers and leading discussion. Each class will also include activities such as small-group discussion, debate, and collaborative reflection, with the goal of treating ideas as something to examine, question, and build together. Because the course depends on active exchange, all discussions take place *in person*, and attendance is *mandatory*.

Grading

- Submit comments to Canvas (10%)
- Participate in class discussions (20%)
- Paper presentation (40%)
- Final paper (30%)

Submit comments to Canvas (10%). Before each class, please upload your comments for the papers through Canvas. Each paper should receive a comment of 100-200 words, including aspects you liked, disliked, found confusing, or were inspired by after reading the paper. The comments serve as a reminder for you to ask questions and engage in discussions. Due the day before class, *every Monday, Wednesday at 11:59 pm.*

Participate in class discussions (20%). Group discussion is a central part of the course and an opportunity to *practice* asking questions, listening carefully to others, and developing ideas through conversation. Thoughtful participation matters more than having polished answers; some of the most productive discussions begin with uncertainty, curiosity, or disagreement. Throughout the quarter, we will have many opportunities for interactive activities, including small-group discussion, role-playing as authors and reviewers, scientific debate, and raising broader questions that connect individual papers to larger themes.

Research presentation (40%). Each student will be asked to present two papers throughout the quarter. Each presentation will last 15 minutes, followed by a 15-minute group discussion. Your presentation should cover the motivation (why the topic is important), background (what has already been done), novel contribution (such as a new method, population, mechanism, or application), methods and results (what was found), and questions for group discussion. The presentation should be clear, concise, well-practiced, and accompanied by slides.

Research paper (30%). You will write a research paper as your final project. It consists of two components: an in-class presentation (15%) and a written paper (15%). More detailed instructions will be provided later. The written component is due on *May 29th, noon*.

Policies

Class start on time. Important administrative announcements will sometimes be made at the start of class. Please come with enough time in advance so you don't miss anything.

Late submission. Assignments are expected to be submitted on time so they can contribute to class discussion. Late submissions will receive a 5% deduction for each day late; missing submissions will receive a 10% deduction. Work submitted after class discussion may still receive credit, but some of its value is lost because it no longer contributes to collective learning. Please communicate early if unexpected circumstances arise.

Academic Integrity. Course assignments are meant to be completed on your own. If you are stressed to the point of cheating – don't. It's just not worth it. Contact me and we can work something out. The University of Chicago's academic integrity statement can be found here.

Use of AI assistants. The policy for use of AI assistants parallels that for human collaboration. It's fine to discuss class materials with AI assistants or use them to generate ideas for research projects or suggestions about papers to read (although be aware they can hallucinate references). If you have used AI assistants, you must specifically describe whatever help you received. It's *not* okay to do anything that results in the generation of questions or proposal contents. If you find yourself putting a paper or abstract from class into an AI assistant, or copying an answer from an AI assistant, you are in violation of the course policies.

Classroom climate. I'm committed to making sure that everyone is treated with dignity and respect in our class. If you have concerns about harassment, discrimination, or unequal treatment in this course, I invite you to contact me. Doing so will never impact your grade. You may also reach out to the University's Office for Equal Opportunity Programs, which includes the Office for Sexual Misconduct Prevention and Support. In handling reports, I will protect your privacy as much as possible. Please know, however, that in some cases (e.g., sexual harassment) I am required to report information. You can find more information about the University's policy on harassment, discrimination, and sexual misconduct here.

Disability services. Students with disabilities are offered accommodations through Student Disability Services. For information about course accommodations, contact SDS at (773) 702-6000 or disabilities@uchicago.edu.

Email. So that I don't miss an email from you, please include "psyc35880-cogs24533" in the subject line of every email you send me. I'll get back to you during normal working hours.

Schedule

Subject to change.

Part I: FOUNDATIONS

March 26, Introduction

- Fiske, S. T., & Taylor, S. E. (2021). *Social Cognition*. (4th ed.). London: Sage Publications Ltd. Ch. 1, pp. 1–21.
- Griffiths, T. L., Chater, N., & Tenenbaum, J. B. (2024). *Bayesian Models of Cognition*. (1st ed.). Cambridge, MA: MIT Press. Ch. 2, pp. 37–58.

March 31, Computational social cognition and levels of analysis 1

- Cushman, F., & Gershman, S. (2019). Editors' introduction: Computational approaches to social cognition. *Topics in Cognitive Science, 11(2)*, 281–298.
- Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman. Ch. 1, pp. 8–38.

April 2, Computational social cognition and levels of analysis 2

- Krafft, P. M., & Griffiths, T. L. (2018). Levels of analysis in computational social science. *Proceedings of the Annual Meeting of the Cognitive Science Society*.
- Bai, X., Fiske, S. T., & Griffiths, T. L. (2022). Globally inaccurate stereotypes can result from locally adaptive exploration. *Psychological Science, 33(5)*, 671-684.

Part II: CORE THEMES IN COGNITIVE SCIENCE AND AI

April 7, Rules and symbols 1

- Haugeland, J. (1997). What is mind design? In J. Haugeland (Ed.), *Mind Design II: Philosophy, Psychology, Artificial Intelligence*. Cambridge, MA: MIT Press. pp. 8–21.
- Newell, A., Rosenbloom, P. S., & Laird, J. E. (1989). Symbolic architectures for cognition. In M. I. Posner (Ed.), *Foundations of Cognitive Science*. Cambridge, MA: MIT Press. pp. 93–131.
- *Optional:* Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton. pp. 11–48.

April 9, Rules and symbols 2

- Charlesworth, T. E., Caliskan, A., & Banaji, M. R. (2022). Historical representations of social groups across 200 years of word embeddings from Google Books. *Proceedings of the National Academy of Sciences, 119(28)*, 1–12
- Peterson, J. C., Uddenberg, S., Griffiths, T. L., Todorov, A., & Suchow, J. W. (2022). Deep models of superficial face judgments. *Proceedings of the National Academy of Sciences, 119(17)*, 1–9.
- *Optional:* Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, 41(3), 258–290.

April 14, Neural network 1

- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior, 8(2)*, 240–247.
- Elman, J., Bates, E., Karmiloff-Smith, A., Johnson, M., Parisi, D., & Plunkett, K. (1996). *Rethinking Innateness: Development in a Connectionist Perspective*. Cambridge, MA: MIT Press. Ch. 2. pp. 47-106
- *Optional:* LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436–444.

April 16, No class

April 21, Neural network 2

- Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science, 7(3)*, 136–141.
- Freeman, J. B., & Lin, C. (2025). A high-dimensional model of social impressions. *Trends in Cognitive Sciences*, 29(9), 790–801.
- *Optional:* Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103(2)*, 284–308.

April 23, Large language models

- Ku, A., Campbell, D., Bai, X., Geng, J., Liu, R., Marjieh, R., McCoy, R. T., et al. (2026). Levels of analysis for large language models. *Philosophical Transactions of the Royal Society*.
- Shiffrin, R., & Mitchell, M. (2023). Probing the psychology of AI models. *Proceedings of the National Academy of Sciences, 120(10)*.
- *Optional:* Bai, X., Wang, A., Sucholutsky, I., & Griffiths, T. L. (2025). Explicitly unbiased large language models still form biased associations. *Proceedings of the National Academy of Sciences, 122(8)*.

April 28, Probabilistic inference 1

- Griffiths, T. L., Chater, N., & Tenenbaum, J. B. (2024). *Bayesian Models of Cognition*. (1st ed.). Cambridge, MA: MIT Press. Ch. 1, pp. 3–35.
- Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review, 98(3)*, 409.

April 30, Probabilistic inference 2

- Gershman, S. J., & Cikara, M. (2020). Social-structure learning. *Current Directions in Psychological Science, 29(5)*, 460–466.
- Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences, 107(20)*, 9066–9071.

Part III: MULTI-AGENT AND SOCIAL PROCESSES

May 5, Attribution and communication

- Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences, 20(8)*, 589–604.
- Sumers, T. R., Hawkins, R. D., Ho, M. K., & Griffiths, T. L. (2021). Extending rational models of communication from beliefs to actions. *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society.*
- *Optional:* Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science, 336(6084)*, 998.

May 7, Making decisions

- Callaway, F., Hardy, M., & Griffiths, T. L. (2023). Optimal nudging for cognitively bounded agents: A framework for modeling, predicting, and controlling the effects of choice architectures. *Psychological Review, 130(6)*, 1457.
- Bai, X., Griffiths, T. L., & Fiske, S. T. (2025). Costly exploration produces stereotypes with dimensions of warmth and competence. *Journal of Experimental Psychology: General, 154(2)*, 347–357.
- *Optional:* Shafir, E., & Tversky, A. (1995). Decision making. In E. E. Smith & D. N. Osherson (Eds.), *Invitation to Cognitive Science, Vol. 3: Thinking* (2nd ed.). Cambridge, MA: MIT Press. pp. 377–425.

May 12, Collective behavior

- Hardy, M. D., Thompson, B. D., Krafft, P. M., & Griffiths, T. L. (2023). Resampling reduces bias amplification in experimental social networks. *Nature Human Behaviour, 7(12)*, 2084–2098.
- Garcia, D., Galesic, M., & Olsson, H. (2024). The psychology of collectives. *Perspectives on Psychological Science, 19(2)*, 316–319.
- *Optional:* Rabb, N., & Sloman, S. A. (2024). Radical collective intelligence and the reimagining of cognitive science. *Topics in Cognitive Science, 16(2)*, 164–174.

Part IV: CRITIQUES AND WRAP-UP

May 14, Limitations of computational approaches to social cognition

- Fiske, S. T. (2004). Mind the gap: In praise of informal sources of formal theory. *Personality and Social Psychology Review, 8(2)*, 132–137.
- Jones, M., & Love, B. C. (2011). Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behavioral and Brain Sciences, 34(4)*, 169–188.
- *Optional:* Wang, A., Kapoor, S., Barocas, S., & Narayanan, A. (2024). Against predictive optimization: On the legitimacy of decision-making algorithms that optimize predictive accuracy. *ACM Journal on Responsible Computing*, 1(1), 1–45.

May 19, In class presentation

May 21, In class presentation