

Counterfactual Reasoning, Normative & Descriptive Aspects

Johns Hopkins University, Spring 2019

Course Information

Instructors	Steven Gross Professor Philosophy Department Gilman 272 sgross11@jhu.edu Office hours: F 12:30-1:30 & by appt
	Justin Bledin Assistant Professor Philosophy Department Gilman 206 jbledin@jhu.edu Office hours: TBD & by appt
Class Code	AS.150.459
Class Time	F 10:00-12:30
Class Location	Gilman 288

Course Description

Counterfactual reasoning is reasoning about what would be the case if things had been other than they are. If it had been sunny and so I didn't run into that store for cover from the rain, maybe I would never have met my future partner! How ought one to reason counterfactually? How do people in fact do it? Counterfactual reasoning might seem like a narrow topic, but it is of fundamental importance to both scientific and everyday inquiry, where it is intimately connected to the use of imagination, planning for the future, assessment of and learning from the past, providing explanations, understanding fictions, and constructing experiments. This course will explore both normative and empirical aspects of counterfactual reasoning, drawing upon readings in philosophy, psychology, linguistics, and computer science. An overarching goal of this course is to arrive at a better understanding of counterfactuality that is informed by research across these different disciplines.

Course-level Learning Goals

Given that this seminar spans philosophy, psychology, linguistics, and computer science, the main learning objective of this course is to develop your ability to critically engage in debates both within particular disciplines and in an interdisciplinary context. From linguists and philosophers of language, you will learn about classic "possible worlds" analyses of counterfactual conditionals and more recent causal analyses using structural equation models. From psychologists, cognitive scientists, and AI researchers, you will learn about Bayesian theories of counterfactual reasoning, the role of imagination in counterfactual thought, and various experimental results concerning how we actually reason counterfactually. There will be ample opportunity to reach across disciplinary

lines in trying to solve the puzzles about counterfactuals that most interest you. More generally, philosophy courses—and this one is no exception—are likely to help your intellectual growth in a number of ways: you will be encouraged to develop your reflective abilities, your skills at scrutinizing arguments and presenting arguments for your own positions (both in writing and in discussion), and at anticipating and responding to counter-arguments to these positions.

Prerequisites

No prior knowledge about counterfactuals is required for this course—in fact, you don't even have to know what counterfactuals are! While previous exposure to formal semantics, probability theory, and experimental research isn't required either, it will be helpful.

Topics

All readings for this course will be posted on Blackboard. To get a sense of the topics that we will be covering this semester, you should check out the following recent survey papers:

- William Starr. Counterfactuals. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy* (Spring 2019 Edition). <https://plato.stanford.edu/archives/spr2019/entries/counterfactuals/>.
- Ruth M. J. Byrne. Counterfactual Thought. *Annual Review of Psychology* 67:135–57, 2016.

The Starr SEP entry reviews work by philosophers on semantic, epistemic, and metaphysical issues concerning counterfactual modality. The Byrne article focuses on empirical work by cognitive scientists and psychologists.

We want to keep things fairly flexible in this seminar and let our conversations and your evolving interests guide us. We also expect to have a number of guest presenters throughout the semester, with the dates and topics for their visits yet to be determined. That said, a natural point of entry into the vast body of interdisciplinary research on counterfactuals is the seminal work of David Lewis (a philosopher) and Angelika Kratzer (a linguist) on the truth conditions of counterfactual conditionals, so we open with their theories.

The Counterfactual Bible: The Book of Lewis and Book of Kratzer

Primary readings:

- David Lewis. *Counterfactuals*. Harvard University Press, Cambridge, 1973.
- Angelika Kratzer. Partition and Revision: The Semantics of Counterfactuals. *Journal of Philosophical Logic* 10: 201–216, 1981.

Supplementary readings:

- David Lewis. Ordering Semantics and Premise Semantics for Counterfactuals. *Journal of Philosophical Logic* 10: 217–234, 1981.
- Angelika Kratzer. An Investigation of the Lumps of Thought. *Linguistics and Philosophy* 12: 607–653, 1989.

- Angelika Kratzer. Facts: Particulars or Information Units? *Linguistics and Philosophy* 25: 655–670, 2002.
- Makoto Kanazawa, Stefan Kaufman, and Stanley Peters. On the Lumping Semantics of Counterfactuals. *Journal of Semantics* 22(2):129–151, 2005.
- Angelika Kratzer. Constraining Premise Sets for Counterfactuals. *Journal of Semantics* 22(2):153–158, 2005.

In the next 2-3 weeks of the course, we turn to causal approaches to the semantics of counterfactual conditionals and counterfactual reasoning, which have been very influential on the contemporary scene in philosophy, linguistics, and cognitive science. Much of this research originates in work by Judea Pearl. There has also been interesting experimental work by cognitive scientists testing whether various Bayes nets theories accurately model our understanding of counterfactuals.

Causal Theories of Counterfactuals

Primary readings:

- Judea Pearl and Dana Mackenzie. *The Book of Why: The New Science of Cause and Effect*. Basic Books, New York, 2018. (Chapter 8)
- Eric Hiddleston. A Causal Theory of Counterfactuals. *Noûs* 39(4):632–657, 2005.
- Lance J. Rips. Two Causal Theories of Counterfactual Conditionals. *Cognitive Science* 34:175–221, 2010.
- Christopher Lucas and Charles Kemp. An Improved Probabilistic Account of Counterfactual Reasoning. *Psychological Review*, 122(4): 700–734, 2015.
- Thomas F. Icard. From Programs to Causal Models. In Alexandre Cremers, Thom van Gessel, and Floris Roelofsen, editors, *Proceedings of the 21st Amsterdam Colloquium*, 2017.

Supplementary readings:

- Boris Kment. Counterfactuals and Explanation. *Mind*, 115(458): 261–310, 2006.
- Rachel Briggs. Interventionist Counterfactuals. *Philosophical Studies* 160(1):139–166, 2012.
- Paolo Santorio. Interventions in Premise Semantics. Forthcoming in *Philosopher's Imprint*.

The rest of the seminar is (more or less) up for grabs. Here are additional topics to draw from:

Background Semantics for Counterfactuals

Primary readings:

- Ivano Ciardelli, Linmin Zhang, and Lucas Champollion. Two Switches in the Theory of Counterfactuals. *Linguistics and Philosophy*, 41(6): 577–621, 2018.

Supplementary readings:

- Daniel Lassiter. Complex Antecedents and Probabilities in Causal Counterfactuals. In Alexandre Cremers, Thom van Gessel, and Floris Roelofsen, editors, *Proceedings of the 21st Amsterdam Colloquium*, pages 115–124, 2017.

Sequences of Counterfactuals and Dynamic Semantic Approaches

Primary readings:

- David Boylan and Ginger Schultheis. Strengthening Principles and Counterfactual Semantics. Unpublished manuscript.

Supplementary readings:

- Kai von Fintel. Counterfactuals in a Dynamic Context. In Michael Kenstowicz, editor, *Ken Hale: A Life in Language*. MIT Press, 2001.
- Frank Veltman. Making Counterfactual Assumptions. *Journal of Semantics* 22(2):159–180, 2005.
- Anthony S. Gillies. Counterfactual Scorekeeping. *Linguistics and Philosophy* 30:329–360, 2007.
- Sarah Moss. On the Pragmatics of Counterfactuals. *Noûs* 46(3):561–586, 2012.

Truthmaker Semantics for Counterfactuals

Primary readings:

- Kit Fine. Counterfactuals Without Possible Worlds. *Journal of Philosophy* 109(3):221–246, 2012.
- Kit Fine. A Difficulty for the Possible Worlds Analysis of Counterfactuals. *Synthese* 189(1):29–57, 2012.

Supplementary readings:

- Kit Fine. A Theory of Truthmaker Content I: Conjunction, Disjunction and Negation. *Journal of Philosophical Logic* 46(6):625–674, 2017.
- Kit Fine. A Theory of Truthmaker Content II: Subject-matter, Common Content, Remainder and Ground. *Journal of Philosophical Logic* 46(6):675–702, 2017.

Backtracking Counterfactuals

- Justin Khoo. Backtracking Counterfactuals Revisited. *Mind*, 126(503):841–910, 2017.

Mental Models and Other Psychological Theories of Counterfactual Thought

Primary readings:

- Philip Johnson-Laird and Ruth M. J. Byrne. Conditionals: a Theory of Meaning, Pragmatics, and Inference. *Psychological Review* 109(4):646–678, 2002.
- Jonathan S. Evans, David E. Over, and Simon J. Handley. Suppositions, Extensionality, and Conditionals: a Critique of the Mental Model Theory of Johnson-Laird and Byrne (2002). *Psychological Review* 112:1040–52, 2005.

- Kai Epstude and Neal J. Roese. The Functional Theory of Counterfactual Thinking. *Personality and Social Psychology Review* 12(2):168–192, 2008.

Supplementary readings:

- Ruth M. J. Byrne. Mental Models and Counterfactual Thoughts about What Might Have Been. *Trends in Cognitive Sciences* 6(10):426–431, 2002.
- Ruth M. J. Byrne. Precis of The Rational Imagination: How People Create Alternatives to Reality. *Behavioral and Brain Sciences* 30(5-6):439–453, 2007.
- Ruth M. J. Byrne. Counterfactual Thinking: From Logic to Morality. *Current Directions in Psychological Science* 26(4):314–322, 2017.
- Kai Epstude and Neal J. Roese. The Functional Theory of Counterfactual Thinking: New Evidence, New Challenges, New Insights. In *Advances in Experimental Social Psychology* (Vol. 56, pp. 1–79). Academic Press, 2017.

Normativity and Modality

- Jonathan Phillips & Joshua Knobe. The psychological representation of modality. *Mind & Language* 33(1):65–94, 2018.
- T. Gerstenberg et al. Lucky or clever? From expectations to responsibility judgments. *Cognition*, 2018.

Developmental Aspects

- Beck, S. R., Robinson, E. J., Carroll, D. J., & Apperly, I. A. (2006). Children’s thinking about counterfactuals and future hypothetical as possibilities. *Child Development* 77:413–426.
- Shtulman, A., & Carey, S. Improbable or Impossible? How Children Reason About the Possibility of Extraordinary Events. *Child Development* 78(3): 1015–1032, 2007
- Gopnik, A., & Walker, C. M. Considering counterfactuals: The relationship between causal learning and pretend play. *American Journal Of Play* 6(1):15–28, 2013.

Episodic Counterfactual Thinking

- De Brigard, F., & Parikh, N. (In press). Episodic Counterfactual Thinking. *Current Directions in Psychological Science*.
- De Brigard, F., Szpunar, K. K., & Schacter, D. L. Coming to grips with the past: Effect of repeated simulation on the perceived plausibility of episodic counterfactual thoughts. *Psychological Science* 24(7):1329–1334, 2013.

Imagery

- Kosslyn, S. M., and S. T. Moulton. Imagining predictions: mental imagery as mental emulation. *Philosophical Transactions of the Royal Society B* 364:1273–1280, 2009.
- N Chater, M Oaksford. Programs as causal models: Speculations on mental programs and mental representation. *Cognitive Science*, 2013

- Christopher P. Barlett and Laura A. Brannon. “If Only...”: The Role of Visual Imagery in Counterfactual Thinking. *Imagination, Cognition and Personality* 26(1-2):87–100, 2006-7.

Knowledge of and through Counterfactuals

- Timothy Williamson. Philosophical Knowledge and Knowledge of Counterfactuals. *Grazer Philosophische Studien* 74:89–123, 2007.
- Timothy Williamson. *The Philosophy of Philosophy*. Blackwell Publishing, Oxford, 2007.
- Timothy Williamson. Knowing By Imagining. In Amy Kind (ed.), *Knowledge through Imagination*, OUP.
- L. A. Paul, J. McCoy, and T. Ullman. Modal Prospection, forthcoming (2019) in *Metaphysics and Cognitive Science*, eds. Alvin Goldman and Brian McLaughlin. Oxford University Press (US).
- Dominic Gregory. Counterfactual Reasoning and Knowledge of Possibilities. *Philosophical Studies* 174:821–835, 2017.

The Value of Counterfactual Thought

- Thomas Icard, Fiery Cushman, and Joshua Knobe. On the Instrumental Value of Hypothetical and Counterfactual Thought. *Proceedings of the 40th Annual Meeting of the Cognitive Science Society*, 2018.

Requirements

There are three requirements for taking this course.

- Participation (worth 30% of your final grade). We expect you to attend each class having carefully read through at least the primary assigned readings. Come to seminar and office hours prepared to ask questions and try out ideas. Moreover, all students are required, starting in week 2, to post to the Blackboard discussion board once a week by 7pm on the Thursday before seminar. All students are required to read what others have posted before seminar. These brief posts should raise a question, or try out a point about that week’s reading, or perhaps draw a connection to something else we read (or something you have learned about elsewhere), etc. If you’re at a complete loss about what to write, provide a clear abstract of the reading. Posts needn’t be more than a paragraph. Individual posts aren’t graded, but your overall posting contributes to your course grade.
- Take-home midterm exam (worth 30% of your final grade). This will be assigned on March 8 and due a week later on March 15.
- Final term paper (worth 40% of your final grade). A 10 page double-spaced paper on a topic of your choice is due at the end of the semester. You must submit a one paragraph proposal to Steven or Justin on your proposed topic, either by email or in person during office hours, by April 5. After that, there will be a peer-review process where each student must submit a draft or skeleton of their paper to another randomly-selected student on April 19 and must return the draft they are assigned the following week on April 26. The term paper itself is due in the final session of the seminar on May 3.

Graduate students taking this seminar for credit can opt not to take the midterm and instead submit a 20-page double-spaced final paper (worth 70% of their final grade). Graduate students are still expected to participate in seminar and post weekly to Blackboard.

Academic Integrity

The strength of the university depends on academic and personal integrity. In this course, you must be honest and truthful. Do not plagiarize—if you are unsure what counts as *plagiarism*, ask! Students caught cheating may receive an F in the course and can face more serious consequences in extreme cases.

For more information, see the Homewood Student Affairs site on academic ethics:

<https://studentaffairs.jhu.edu/student-life/student-conduct/academic-ethics-undergraduates>.

See also the e-catalog entry on the undergraduate academic ethics board:

<http://e-catalog.jhu.edu/undergrad-students/student-life-policies/#UAEB>

Disabilities

Any student with a disability who may need accommodations in this class must obtain a letter from Student Disability Services, 385 Garland, 410.516.4720, studentdisabilityservices@jhu.edu. Accommodations recommended by SDS will be implemented.

Mental Health

If you are struggling with anxiety, stress, depression or other mental health related concerns, please consider visiting the JHU Counseling Center. If you are concerned about a friend, please encourage them to seek out these services. The Counseling Center is located at 3003 North Charles Street in Suite S-200 and can be reached at 410.516.8278 and online at <http://studentaffairs.jhu.edu/counselingcenter/>

Enjoy the course!