

PHIL 2460: The Epistemology of Bias

Fall 2021

Class: Mondays 1–3:30pm

Classroom(!): 1008B Cathedral of Learning / <https://pitt.zoom.us/j/95332310151>

Office: 1029G CL / <https://pitt.zoom.us/j/4971426138> OH: Wed 3–5pm; by appointment

Kevin Dorst

kevindorst@pitt.edu

Course Description: Brian is biased. So are the algorithms he uses, the social networks he belongs to, and the search committees he sits on. This much is obvious. But what does it mean? In what sense(s) can people, algorithms, networks, and selection processes each be biased? And what should we make of the fact that they are? This course will use readings from philosophy, cognitive science, economics, and law to offer a tour through the contemporary epistemological study of bias in these various domains. We'll try to get clear on what these various biases amount to, what causes them, and what we can(not) and should (not) do about them.

Readings: All readings will be made available on the course website:

https://www.kevindorst.com/fa_21_bias.html

Grading: If you're taking the class for credit, your grade will be determined by a paper (70%), a presentation (15%), and participation (15%).

Paper (70%): Standard 8000-ish word term paper. A *draft* must be turned in by **December 10**, to be revised with the final version turned in by **January 9**.

Presentation (15%): I'll ask you to lead discussion of one of the papers we're going to cover in the course. Construct a handout to lead us through the main points of the paper, raising questions/concerns/objections for class discussion along the way. Your presentation should cover half of one session—approximately 1–1.25 hours. The point is to practice; afterwards we'll have a brief, private "post-op" on things that went well and ways to improve.

Participation (15%). Please do all the readings beforehand and come to class prepared to discuss and raise questions. If you find yourself hesitant to speak up in class, an alternative is to shoot me questions about the reading beforehand, and/or schedule meetings or come to office hours to discuss.

Course Policies:

Academic Integrity. You know the drill: <https://www.provost.pitt.edu/faculty/academic-integrity-freedom/academic-integrity-guidelines>

Disability Accommodations. If there's any issue to touch base about, let me know! Disability Resources and Services: <https://www.diversity.pitt.edu/disability-access/disability-resources-and-services>; also 412-648-7890.

Zoom. Please come in person, if you can (and are vaccinated)! But I will also set up zoom in the classroom as an option for people who (a) are not able to be in Pittsburgh, or (b) are quarantining, at risk, or otherwise unable to come to class.

Schedule

I. Biased People

8/30: Intro

- Kelly MS, *Bias: A Philosophical Study*, Ch. 4, “Biased People”
- Johnson 2019, “The Structure of Bias”

9/6:

- **No Class (Labor day)**

9/13: Perception

- Siegel 2017, *The Rationality of Perception*, Chs. 1–3.

9/20: Culpability

- Begby 2013, “The Epistemology of Prejudice”
- Begby 2018, “Straight Thinking in Warped Environments”
- Siegel 2018, “Replies to Begby, Ghijsen and Samoiloa” [focus on Begby section]

II. Biased Algorithms

9/27: What are they?

- Goodman 2016, “Economic Models of (Algorithmic) Discrimination”
- Johnson 2020, “Algorithmic bias: on the implicit biases of social technology”

10/4: Why are they bad?

- Creel and Hellman 2021, “The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision Making Systems”
- Johnson 2020, “Are Algorithms Value-Free?”

10/11: When are they bad?

- Hedden 2021, “On statistical criteria of algorithmic fairness”
- Hellman 2021, “Big Data and Compounding Injustice”

III. Biased networks

10/18: Networks and credibility

- Zollman MS, *Network Epistemology*, Chs. 1, 4, and 8
(Kevin Zollman will visit the class.)

10/25: Bad science?

- Smaldino and McElreath 2016, “The natural selection of bad science”
- Weatherall, O’Connor, and Bruner 2020, “How to beat science and influence people: Policymakers and propaganda in epistemic networks”

11/1: Self-correction?

- Romero 2016, “Can the behavioral sciences self-correct? A social epistemic study”
- Bruner and Holman 2019, “Self-correction in science: Meta-analysis, bias and social structure”

IV. Biased Selection

11/8: Commensuration Bias

- Lee 2015, “Commensuration Bias in Peer Review”
- Heesen 2021, “The Necessity of Commensuration Bias in Grant Peer Review”

11/15: Publishing

- Heesen 2018, “When Journal Editors Play Favorites”
- Heesen and Remeijn 2019, “Epistemic Diversity and Editor Decisions: A Statistical Matthew Effect”

11/22:

- **No class (Thanksgiving)**

11/29: Shortlisting

- Bovens 2016, “Selection under uncertainty: Affirmative action at shortlisting stage”
- Mulligan 2017, “Uncertainty in Hiring Does Not Justify Affirmative Action”

12/6: Statistical Discrimination

- Aigner and Cain 1977, “Statistical Theories of Discrimination in Labor Markets”
- TBD. *Maybe* Dorst and Salow MS