

24.223 Paper 2 Prompts

Kevin Dorst

Instructions: Please write a roughly 2000-word paper (7–8 pages double-spaced) that responds to one of the following prompts.

ChatGPT and AI: You can use whatever methods you like to brainstorm and outline, but once you start writing your draft, all your writing must be your own. **You must write your draft, from start to finish, in Google Docs.** When you submit your paper, you will **share editing privileges with me for your paper**, so that—if needed—I can use Draftback to verify that you wrote it yourself, rather than using an AI to write it.

The aim is to make one small contribution to the conversation started by the paper and in class. You should begin with a thesis paragraph, explaining what you will argue and how you will proceed. You should then reconstruct the argument you will be engaging with in a way that is understandable to someone who has not read the paper or come to class. (You may presuppose the basics of Bayesian epistemology and decision theory, though it is good to remind your readers what any terms and equations mean.) After doing so, you should make your contribution, which should come in the form of raising or responding to an objection to the argument you have reconstructed. I will grade your paper by evaluating it for (1) sentence-level and structural *clarity*, (2) accuracy and adequacy of its *exposition* of the target argument, and (3) the novelty and substance of its *engagement* with that argument.

Papers should be submitted over email to kmdorst@mit.edu by sharing editing privileges with your Google Doc. The due date is **11:59pm on December 13**. Late papers will be marked down 1/3 of a letter grade for each late day. Due to the grade deadline, **any paper turned in after December 18 will automatically receive a 0**.

Prompts:

1. O'Connor and Weatherall use a bandit model to argue that selective trust can lead to rational polarization. Is this a plausible model of trust? If you think it is, how would you respond to the strongest objection to it? If you think it's not, how should O'Connor and Weatherall respond to this claim?
2. Kelly (2008) argues that selective scrutiny is rational, and that as a result those who aren't aware of this fact are rational to polarize when presented with mixed evidence. Explain his argument, and either object to it or defend it against an objection.
3. Salow and Kevin both argue that confirmation bias (or "intentionally biased inquiry") should be understood in terms of Reflection failures. Explain why they think this. A natural objection is that someone could be biased without having any idea that they are. What's the best way to formulate this objection? Is there a plausible way for them to respond?
4. Singer et al. argue that coherence-minded memory management is more rational than weight-minded management. This is because weight-minded management can sometimes lead you to predict that you're going to change your belief on the basis of worse reasons than you have now. Explain what this means and reconstruct their argument for it, and then raise and evaluate an objection.
5. Dallmann introduces a model of limited-memory management, and uses it to argue in favor of the rationality of an obstinate (or selective) response to new information. Explain his model and argument. What's the best way to object to it?