

## **Ethics and the Future -- Assigned Readings**

Note: This is a list of the readings that I assigned (or recommended) to the students (a mix of undergraduates and graduate students) for my seminar, “Ethics and the Future,” taught at Yale in Spring 2021. As the reading list was prepared only for the students in the course, I’m afraid I didn’t include links for any papers that I was including in the course pack, though I believe all of these readings can be easily found online, with the exception, of course, of Ord’s The Precipice, and Parfit’s Reasons and Persons.

I am making the reading list publicly available at the request of Pablo Stafforini.

In choosing suitable readings I was aided tremendously by the list of readings selected by William MacAskill and Christian Tarsney for their own seminar taught at Oxford in 2019, and especially by the potential syllabus that Joshua Monrad wrote up for my benefit. (Since it was Joshua who rightly insisted to me that this would make an interesting and important topic for a seminar, I am doubly in his debt.)

The readings are organized by topic. We spent a week on each.

--Shelly Kagan

### **Background on Existential Risks:**

1. Toby Ord, The Precipice, Chapters 3-6 and Appendices C and D (about 124 pages)

### **The Basic Case for Longtermism:**

1. Perhaps start with this very brief overview: Todd, “Future Generations and Their Moral Significance” (about 7 pages), which can be found online at: <https://80000hours.org/articles/future-generations/>
2. Then look at the somewhat longer (but still breezy) exposition in Ord, The Precipice, Intro and Chapters 1-2, and Appendix E (65 pgs.)
3. Then read Chapters 1 and 3 from Nick Beckstead’s dissertation, On the Overwhelming Importance of Shaping the Far Future (about 44 pages)
4. Finally, Greaves and MacAskill, “The Case for Strong Longtermism” (about 25).

That will come to about 140 pages, most of which reads fairly quickly. If you want even more (consider what is listed next as recommended but not required)--look at:

5. Bostrom, “Existential Risk Prevention as Global Priority” (17 pages), and also

6. Bostrom, “Astronomical Waste” (10 pages)

7. Finally, there is a passage from Derek Parfit, Reasons and Persons, that is quoted regularly in the longtermist literature (for example, by Beckstead). If you would like to see it in its original context, it is on pp. 453-4 (2 pages).

All of these things can be found in the Files folder for the class, other than the Todd, the Ord, and the Parfit.

### **The Social Discount Rate:**

1. Start with Cowen, Discount Rates Table, a short passage from his Stubborn Attachments, which gives a quick sense of how even a “modest” discount rate effectively wipes out the significance of the long term future (1 page).

2. Then Parfit, Reasons and Persons, Appendix F (7 pages), for arguments against the social discount rate.

3. Cowen and Parfit, “Against the Social Discount Rate,” (from Peter Laslett & James S. Fishkin (eds.) Justice between age groups and generations, Yale University Press: New Haven, 1992, pp. 144–161) repeats much of the Parfit but gives some additional arguments. To (mostly) avoid the repetition, only read the two introductory pages (pp. 144-145) and the section on “economic arguments” (pp. 150-158). Though the first such economic argument (on opportunity costs) very closely follows the earlier Parfit, it does add some extra details. (11 pages.)

4. Then Ord, The Precipice, Appendix A (6 pages) for further discussion.

5. Next, read Greaves, “Discounting for Public Policy,” section 7, which is pages 404-409 (5 pages). That’s the bit on the “pure” discount rate. (The rest of the paper isn’t required, but is recommended for anyone who would like a thorough (though a bit technical) survey of some of the economics debates on the discount rate.)

6. Finally, Mogensen, “The Only Ethical Argument for Positive Delta” (33 pages).

That’s about 62 pages.

7. If you are interested in further discussion of the discount rate from an economist’s perspective, you could take a look at Broome, “Discounting the Future,” (29 pages) though this is primarily on discounting with regard to future resources, not pure discounting of future welfare, so it is only recommended.

## **Population Ethics I:**

Parfit, Reasons and Persons, Chapters 16-18, and Appendix G (70 pages).

## **Population Ethics II:**

1. Start with Boonin, “How to Solve the Non-Identity Problem” (30 pages)
2. Next, Harman, “Can We Harm and Benefit in Creating?” (25 pages)
3. Then McMahan, “Climate Change, War, and the Non-Identity Problem” (27)
4. Beckstead, Overwhelming Importance, Chapter 4 (23 pages)
5. Ord, The Precipice, Appendix B (6 pages)
6. Finally, a few pages from Kagan, “Singer on Killing Animals” from the middle of p. 141 (“Some jargon...”) to the lower middle of p. 145 (“the various alternatives”) (4 pages)

That’s about 115 pages.

## **Aggregation Principles:**

1. Start with Hurka, “Value and Population Size” (12 pages).
2. Beckstead, Overwhelming Importance, Chapter 5 (32 pages)
3. Huemer, “In Defence of Repugnance” (32 pages)
4. Parfit, Reasons and Persons, Chapter 19 (24 pages). This can be slow going, because there’s a lot of detailed examination of possible objections to the basic argument. But the basic idea of the argument shouldn’t be too hard to get; indeed, the Huemer presents a simplified version of it. So don’t worry about following all the details.

That’s about 100 pages.

5. If you’d like a systematic review of most of the main proposals that have been made about aggregation functions, you might have a look at Greaves, “Population Axiology.” That can be a bit of a hard slog, and I’ve already given you one with the Parfit, so let’s just consider that recommended. (18)

### **The Survival of Humanity:**

1. Lenman, "On Becoming Extinct" (16 pages)
2. Frick, "On the Survival of Humanity" (22 pages) – all of it is relevant to the course, but sections 9 and 10 are the ones most relevant for this week's particular topic.

That's a mere 38 pages!

### **Deontology (Options and Constraints):**

1. Reread Greaves and MacAskill, "The Case for Strong Longtermism," section 6 (pp. 25-29) (4 pages)
2. Mogensen, "Moral Demands and the Far Future," only read section 4 on Agent-Centred Prerogatives (pp. 9-11) (2 pages).
3. Mogensen, "Staking Our Future" (27 pages)
4. MacAskill and Mogensen, "The Paralysis Argument" (36 pages)
5. Finally, Unruh, "The Constraint against Doing Harm and Long-Term Consequences" (18)

That's about 87 pages.

6. Adams, "Must God Create the Best?" (16 pages) argues against the requirement to maximize the good, from the perspective of religious ethics. That's only recommended.

### **Cluelessness:**

1. Lenman, "Consequentialism and Cluelessness" (29 pages)
2. Then reread MacAskill and Mogensen, "The Paralysis Argument," section 3.1 (2 pages).
3. Greaves, "Cluelessness" (18 pages). (This was published in 2016, in the Proceedings of the Aristotelian Society, 116, pp. 311-339.)
4. Finally, Mogensen, "Maximal Cluelessness" (22 pages)

That's about 71 pages

### **Very Small Probabilities:**

1. Start with Bostrom, “Pascal’s Mugging” (3 pages!)
2. Then Balfour, “Pascal’s Mugger Strikes Again” (6)
3. Beckstead, Overwhelming Importance, Chapter 6 (25)
4. Tarsney, “The Epistemic Challenge to Longtermism,” only read section 6.2 (Fanaticism, pp. 30-31), which suggests a rough way to measure just how dependent longtermism is on very small probabilities--given some calculations he has done earlier in the paper. (2 pages)
5. Wilkinson, “In Defence of Fanaticism.” This gets rather technical in section 6 (“Egyptology and Indology,” pp. 18-28), so consider that part merely recommended. (The rest comes to about 19 pages.) I also think I caught a typo or two, which caused me some initial confusion.

That comes to 55 pages.

6. Beckstead and Thomas, “A Paradox for Tiny Probabilities and Enormous Values,” (36 pages) repeats the main arguments from Beckstead, Chapter 6 (as well as some material from his Chapter 7), so it is only recommended. However, it elaborates on a number of points--exploring them in greater detail--so if these issues about tiny probabilities especially interest you, you might want to read it as well.
7. If you are especially intrigued by the technical issues of how to handle the possibility of infinite value, you might want to look at Beckstead, Overwhelming Importance, Chapter 7 (the first two sections, 7.1 and 7.2, are particularly relevant for our class). (The entire chapter is 23 pages.)
8. And if you are really interested, even more generally, about how to extend ethics to cover infinity, one overview can be found in Bostrom, “Infinite Ethics.” (55 pages)

### **Moral Uncertainty:**

1. Gracely, “On the Incomparability of Judgments Made by Different Ethical Theories” (5 pages)
2. Ord and MacAskill, “Why Maximize Expected Choice-Worthiness?” (25)
3. Greaves and Ord, “Moral Uncertainty about Population Axiology” (30)

That comes to 60 pages.

4. Harman, “The Irrelevance of Moral Uncertainty” (26 pages) argues that moral uncertainty actually doesn’t affect what one ought to do at all. That’s only recommended.

### **Intuitions:**

1. Beckstead, Overwhelming Importance, Chapter 2, and sections 7.3 and 7.4 (36 pages in all)
2. Reread Huemer, “In Defence of Repugnance,” sections 1.5 and 2 (5 pages)
3. Reread Boonin, “How to Solve the Non-Identity Problem,” section 3 (9 pages)
4. Reread Wilkinson, “In Defence of Fanaticism,” section 2 (1 page)

That’s 51 pages.

### **What to Do:**

If you are persuaded by longtermism, what should you do? Here are some things worth looking at:

1. Reread Greaves and MacAskill, “The Case for Strong Longtermism,” section 3 (9 pages)
2. Then reread Beckstead, Overwhelming Importance, section 3.3 (pp. 67-72) (6 pages)
3. Ord, The Precipice, Chapter 7 and Appendix F (35 pages)
4. For a concrete example of something that could be done, read Matheny, “Reducing the Risk of Human Extinction,” section 4 and then section 7 through to the end (4 pages).

That comes to about 54 pages required. Here are a few more recommendations:

5. If you wanted to pick a career working on one of these problems, what might that be? The organization 80,000 Hours has a web page suggesting areas one might specialize in, “Our Current View of the World’s Most Pressing Problems.” (That’s roughly 9 pages, but much of the text is only visible when you click on the relevant “show” buttons. Clicking on all of them brings it up to about 20 pages, I think.) You can find it at: <https://80000hours.org/problem-profiles/#Other-problems-we've-looked-at-which-seem-less-impactful-to-work-on-than-global-health>
6. For some thoughts on changing institutions so as to take long term considerations into account, John and MacAskill, “Longtermist Institutional Reform” (18 pages)
7. Finally, for one vision of what humanity’s future could hold, end with Ord, The Precipice, Chapter 8 (25 pages)