HS2921: Effective Altruism in Theory and Practice

Semester 1, Academic Year 2024-2025

Syllabus

Instructors
Dr. Joel Chow, joelchow@nus.edu.sg
Dr. Martin Mattsson, martin.mattsson@nus.edu.sg

Course Description
How can we best use our resources to help others? And how do we decide whom (or what) we should devote resources to? How do we ensure that our efforts are effective in improving the wellbeing of others?

Economics and philosophical approaches converge as we discuss these questions. The former allows us to construct empirically meaningful measures of wellbeing and evaluate amongst a wide array of possible policy or charitable options. The latter provides a methodology for answering non-empirical moral questions, such as our obligations towards the global poor, non-human animals, and future populations.

Prerequisites
There are no pre-requisites, but this course is reserved for students from Year 2 onwards.
Students are encouraged to have completed at least two of the Level-1000 CHS Common Curriculum courses, which may include: HSI1000, HSH1000, CHS Data Literacy course, and CHS Digital Literacy course.

Disclosure
Mattsson has received funding from Open Philanthropy for one of this research projects. Open Philanthropy's approach to philanthropy is discussed in detail throughout this course.

Assessment Components

1. Participation (20%)

Participation will be graded based on the quality of contributions to class discussion. Quality here refers to good-faith attempts to further the discussion of the class (asking thoughtful questions, asking for clarification that helps to further discussion, etc.). Class discussion here is understood quite broadly, it includes discussion in the classroom, polls during lectures, and a forum post submitted by the students on Friday of week 5.

To help students who might not be so familiar with group discussions of this nature, a video (both visual format and podcast format) modelling a discussion by two different groups (faculty members and students from different faculties including Science, Law, FASS, CDE, Computing) will be made available to students before the semester begins.[LTW2] This will give them an idea of what to expect and how interdisciplinary groups interact and communicate. (Note: this podcast is being developed by one of the primary instructors of this course [Joel Chow] using his classroom experience from teaching at NUS College).

2. Quizzes (25%)

Students will be quizzed on the starred readings using Canvas. Quizzes will take place at the start of the class.

3. Group project (25%)

Throughout this semester, we have seen that there are a few key areas of moral concern: i) the global poor; ii) animal welfare; and iii) long-term existential risks. We have also discussed and examined tools that social scientists use to evaluate the efficacy of potential interventions and ways in which individuals might seek to shift behaviour.

For this group assignment, in groups of 5 - 6, we would like you to discuss and focus on a sub-area of concern within the three broad categories mentioned above. We would like you to do some research using the concepts and ideas we have discussed, and identify a problem area you believe will be ripe for an EA intervention. The group will then detail the potential intervention, including explaining why you think it would be effective, and how you would track its efficacy. You should also note the limitations of your approach as well as highlight key assumptions you have made in your attempt to evaluate the efficacy of the project. Throughout the project, the group will be required to keep track of their discussions through a journal or using a Google form that will be sent to each group. Each group will also fill in a peer evaluation and self-evaluation form in the middle and at the end of the semester, keeping track of your progress and teamwork as a group.

The project will be graded according to the following metrics:

- Good teamwork and collaboration through the ideation, conception, and evaluation phases. This score will be assigned at an individual level. (20%);

- Choice of problem area, including providing a good justification using reason and/or empirical evidence for why this is a potentially very impactful project (20%);

- Thoughtful use of concepts, theories, and tools discussed in the course to track and evaluate the efficacy of the proposed intervention or a good discussion of why empirical evidence will be hard to use and a sound argument for why the project needs to be evaluated on a theoretical basis (25%);

- Thoughtful and clear expression of the limitations of the approach, and assumptions made in attempts to evaluate the efficacy of the intervention (20%);

- Creative means of expressing and presenting ideas (15%).

4. Essay (30%)

For the essay, students can choose between the following options:

a) Case Study Analysis

For this option, using what you have learned, we would like you to analyse and evaluate an intervention from an effective altruist perspective. You should select a project that has been carried out and which you have reliable data to analyse and evaluate. Using what we have discussed in class, examine the intervention and explain the strengths and weaknesses of the intervention.

b) A Proposal

For this option, using what you have learned, please identify a key area where you believe an effective altruist intervention will be effective in improving the wellbeing of your target group, explaining why you have chosen to focus on this group, how the intervention will be efficacious, how you intend to track the efficacy of the said intervention, and the possible limitations and challenges of implementing this intervention.

Communication

If you have a question, please email both of us or ask us during or after class. All assignments should be uploaded to the respective Canvas assignment folders and should not be sent to us by email.

Grading Policy

Regular attendance and class participation are expected. Late submissions will be graded down for every 24 hours they are late.

Course Schedule

Framework: What is Welfare and Why Does It Matter?

Week 1 - Introduction

Week 2 - Why Welfare Matters (Quiz 1, group formation due Friday 11:59pm)

Week 3 - What is Welfare?

Week 4 - Measuring Welfare and Cause Prioritization

Week 5 - Markets, Market Failures, Government Failures, and Cause Prioritization (Quiz 2, forum post due Friday 11:59pm)

Theme 1 - Why and How Can We Promote the Welfare of the Global Poor?

Week 6 - Global Health and Development - The "evidence-based" altruism approach

Week 7 - Global Health and Development - Other approaches to improving global health and development (Quiz 3)

Theme 2 - If Welfare Matters, What about Animal Welfare?

Week 8 - Should We be Concerned about Animal Welfare?

Week 9 - The Economics of Social Movements (Quiz 4)

Theme 3 - What about Future Generations?

Week 10 - What (if anything), Do We Owe to our Distant Descendants?

Week 11 - Existential Risks and Discounting (Quiz 5)

Criticisms and conclusion

Week 12 – Guest Lecture: Peter Singer

Week 13 - Criticisms of Effective Altruism (Essay due Friday at 11:59pm)

Framework:

Week 1: Introduction (Chow + Mattsson)

This class has three guiding questions:

- How can we best use our resources to help others?

- How do we decide whom (or what) we should devote our resources to?

- How do we ensure that our efforts are effective in improving the wellbeing of others?

In week 1, we will examine why these questions are important. We will also see that in order to answer all three questions comprehensively, we need a conceptual and practical toolkit that brings in skills, methods, and frameworks from (at least) two disciplines: philosophy and economics. Briefly, philosophy seeks to answer normative questions about how individuals and societies should behave. Economics, on the other hand, is concerned about the efficient allocation of scarce resources. Put together, these disciplines can help us identify key areas of moral and practical concern, and attempt to identify the most efficient means of improving the lives of individuals and societies around us.

In order to answer these questions fully, we will need to have a basic grounding in some disciplinary tools. In this introductory lecture, you will be introduced to basic philosophical concepts and methodology (normative vs. descriptive approaches; thought experiments); economic concepts (cost-benefit analysis, opportunity costs and trade-offs, utility functions, and game theoretic concepts); and some social science methodology (Randomised Controlled Trials, data analysis).

Finally, we begin to examine one potential answer to the above questions: Effective Altruism (EA). In a nutshell, we will define EA as "the use of evidence and reason in search of how to benefit others as much as possible, and taking action on that basis". This definition, while seemingly simple, as we will see, might have radical implications for our thinking about the way we approach our personal and professional lives, as well as how we conceive of helping others. Throughout the semester, we will unpack what this means carefully using the tools from philosophy and economics, and bring in other disciplines such as social psychology and politics.

Week 1 required Readings:

- Introduction to Effective Altruism (Online Article: www.effectivealtruism.org/articles/introduction-to-effective-altruism) (20 min)

Supplementary Material:

- "The How and Why of Effective Altruism", Peter Singer (TED Talk: www.ted.com/talks/peter_singer_the_why_and_how_of_effective_altruism)
- "The Definition of Effective Altruism", Will McAskill (Book Chapter: https://academic.oup.com/book/32430/chapter/268751648)

Week 2 - Why Welfare Matters (Chow)

In week 1, we began examining the three guiding questions of this course. Recall the first question: how can we best use our resources to help others? One might argue that this question makes an assumption, that we *should* help others. But why should we care for others in the first place? Many philosophers argue that we should care for others because we should care about the welfare of one another. This week, we unpack the idea of why we should care about the welfare of others, and how that might have significant implications for thinking about global poverty.

Week 2 required readings:

- "Famine, Affluence, and Morality", Peter Singer

Supplementary Readings:

  - "Beneficence, Duty, and Distance", David Miller
  - "Sometimes there is Nothing Wrong with Letting a Child Drown", Travis Timmerman

Week 3 - What is Welfare? (Chow)

Week 3 - What is Welfare?

In week 2, we examined the idea that we should care about the wellbeing of others, and that if we take this idea seriously, it seems that we are obligated to help the global poor. But this raises the question–what precisely is welfare? How should we best understand it? This week, we explore three major philosophical theories of wellbeing: i) hedonism, ii) preference satisfaction, and iii) objective-list theories. We also end with a question: how might we seek to operationalise and measure welfare?

Week 3 required readings:

- "Well-Being", Roger Crisp (sections 1, 4, and 5)

Week 4 - Cause Prioritization, Marginal Thinking, Expected Utility and Measuring Welfare (Mattsson)

Week 4 - How Can We Measure Welfare and Prioritize Between Different Causes/Approaches/Projects?

In week 3, we will reintroduce the Importance-Neglectedness-Tractability framework for cause prioritization. We will talk about how one can go about measuring welfare using Disability-Adjusted Life Years (DALYs) as well as more subjective measures of well-being. We will introduce the concept of thinking about marginal (instead of average) impact. We will explain how expected utility maximization works and how it is related to the "Veil of Ignorance" thought experiment.

Week 4 required readings:

The Moral Imperative toward Cost-Effectiveness in Global Health, Toby Ord, 2013, https://www.cgdev.org/sites/default/files/1427016_file_moral_imperative_cost_effectiveness.pdf

Supplementary materials:

What is social impact? A definition, Benjamin Todd, https://80000hours.org/articles/what-is-social-impact-definition/

The WELLBY: a new measure of social value and progress, Frijters et. al. (2024), Nature, https://www.nature.com/articles/s41599-024-03229-5

Key Ideas, Happier Lives Institute (About using aggregated subjective wellbeing as a measure of welfare) www.happierlivesinstitute.org/key-ideas/

Podcast: Will MacAskill on Effective Altruism, Moral Progress, and Cultural Innovation, Ep. 156 of "Conversations with Tyler", https://conversationswithtyler.com/episodes/william-macaskill/

Podcast: Toby Ord on the perils of maximising the good that you do (https://80000hours.org/podcast/episodes/will-macaskill-moral-philosophy/)

Podcast: Will MacAskill on moral uncertainty, utilitarianism, and how to avoid being a moral monster (https://80000hours.org/podcast/episodes/will-macaskill-moral-philosophy/)

Tutorial 1:

- Formation of groups

- Group discussions on how the philosophical normative questions on welfare might be answered by economic analysis

- Group discussions on what kinds of duties (if at all) we owe to the global poor

Week 5 - Markets, Market Failures, Government Failures, and Finding Opportunities to have Large Impact (Mattsson)

In week 5, we will briefly introduce the first and second fundamental theorems of welfare economics. We will discuss under what conditions these theorems holds and when we can expect markets to deliver efficient allocations. We will then talk about different classes of market failures and government failures and how we can use these frameworks in our search for opportunities to do have a large positive marginal impact.

Week 5 required readings:

Philanthropic Cause Prioritization, Emily Oehlsen, 2024, Journal of Economic Perspectives (https://www.aeaweb.org/articles?id=10.1257/jep.38.2.63)

Supplementary materials:

Blogpost: Radical Empathy, Holden Karnofsky, https://www.openphilanthropy.org/research/radical-empathy/

Podcast: 80,000 hours with Ajeya Cotra on worldview diversification and how big the future could be. https://80000hours.org/podcast/episodes/ajeya-cotra-worldview-diversification/

Theme 1 - Global Health and Development

Week 6 - Maximizing Marginal Improvements and Generating Evidence of Impact (Mattsson)

This week we will discuss the differences in incomes and health outcomes between high- and low-income countries and how these have changed over time. We will talk about the basic case for why focusing on interventions in low-income countries allows us to have a large positive impact. We will talk about the evidence-based approach to doing good and why randomized controlled trials have become an important tool to generate evidence on the impact of different interventions.

Week 6 required reading:

Podcast: Elie Hassenfeld on two big-picture critiques of GiveWell's approach, and six lessons from their recent work with Robert Wiblin, https://80000hours.org/podcast/episodes/elie-hassenfeld-givewell-critiques-and-lessons/#transcript

The whole podcast is great but you are only required to listen from Minute 0 to Minute 30 and then again from 2h20min to 2h40min (chapter "Whether economic policy is what really matters overwhelmingly"). You can also just read the transcript, if you prefer reading over listening.

Website: Read through GiveWell's research report on the Malaria Consortium: https://www.givewell.org/charities/malaria-consortium

(This is to give you a sense of what type of research and evaluations that GiveWell does)

Watch short video about deworming: https://www.youtube.com/watch?v=byhCzOIWw0o

Supplementary materials:

The rest of the podcast: https://80000hours.org/podcast/episodes/elie-hassenfeld-givewell-critiques-and-lessons/#transcript

Original Research Paper: Miguel, Edward, and Michael Kremer. "Worms: identifying impacts on education and health in the presence of treatment externalities." Econometrica 72.1 (2004): 159-217.

Original Research Paper: Baird, Sarah, Craig McIntosh, and Berk Özler. "When the money runs out: Do cash transfers have sustained effects on human capital accumulation?." Journal of Development Economics 140 (2019): 169-185.

Review Paper: Olken, Benjamin A. "Banerjee, Duflo, Kremer, and the rise of modern development economics." The Scandinavian Journal of Economics 122.3 (2020): 853-878.

Website: Givewell, https://www.givewell.org/charities/top-charities

Tutorial 2:

- Group discussions integrating philosophical and economic theories on wellbeing

- Group discussions on why it is important to have a *measurement* of wellbeing

- Group discussions on how measurements of wellbeing can help us guide our decision-making (this relates to the central concern of EA and the three guiding questions of the course)

Week 7 - Can We Have an Even Bigger Impact? Growth, Institutions, Migration, and Technological Breakthroughs (Mattsson)

In the previous week, we spoke about how evidence is generated and how we can decide between different tried and tested approaches to having a large positive impact. In this week we will be discussing ideas and projects with even bigger potential but where there is much less evidence in favor of these appraoches. First we will talk about theories about the underlying reasons for why some countries have much higher average income levels while others have much lower average income. We will then talk about what could be done to boost economic growth in low-income countries. Finally, I will make a case for why promoting migration and developing certain new technologies might be the most impactful things you can do.

Week 7 required reading:

Forum post: "Growth and the case against randomista development" by Hauke Hillebrandt and John Halstead, https://forum.effectivealtruism.org/posts/bsE5t6qhGC65fEpzN/growth-and-the-case-against-randomista-development

Supplementary materials:

Book: Why nations fail: The origins of power, prosperity and poverty. By James A. Robinson and Daron Acemoglu, 2012.

Clemens, Michael A. *"Economics and emigration: Trillion-dollar bills on the sidewalk?."* Journal of Economic perspectives 25.3 (2011): 83-106.

Economic growth is enough and only economic growth is enough, Lant Pritchett and Addison Lewis, 2022, url: https://lantpritchett.org/wp-content/uploads/2022/05/Basics-legatum-paper_short.pdf

Podcast: Rachel Glennerster on how "market shaping" could help solve climate change, pandemics, and other global problems, by Luisa Rodriguez and Keiran Harris URL:https://80000hours.org/podcast/episodes/rachel-glennerster-market-shaping-incentives/

Theme 2 - If Welfare Matters, What about Animal Welfare?

Week 8 - Should We be Concerned about Animal Welfare? (Chow)

So far, we have discussed welfare with respect to humans. We have seen how if we are concerned with the welfare of humans, we might have good reason to believe that we are obligated to help the global poor. And we have also seen that in one popular view of welfare, welfare is understood as seeking pleasure and avoiding pain. This raises a natural question: what about non-human animals? This week, we examine the idea that if welfare matters, it seems that we have good reason to be concerned about the welfare of non-human animals as well. As we saw with the global poor, it seems that this idea has far-reaching consequences for our everyday lives as well as our social institutions and organisations.

Finally, we also examine an economic explanation for a potential paradox: why do many individuals report having empathetic feelings towards animals but continue consuming meat?

Week 8 required readings:

- "All Animals are Equal", Peter Singer

Supplementary Readings:

- "An Economic Model for the Meat Paradox", Hestermann et al.

Tutorial 3

- Group discussion on the argument that we should treat animal welfare

Week 9 - The Economics of Social Movements (Mattsson)

Building upon our discussion in week 8, we have seen that in the past few decades, the animal rights movement has emerged from a generally fringe movement to a significant social force, changing the way we think about animal welfare in many areas. In this lecture we will discuss different strategies for improving the lives of farmed animals.

Week 9 required readings:

- Podcast: Lewis Bollard on the 7 most promising ways to end factory farming, and whether AI is going to be good or bad for animals, by Luisa Rodriguez and Keiran Harris. From 00:30 (Potential solutions) to 01:31 (Farm animal welfare in Asia), URL:https://80000hours.org/podcast/episodes/lewis-bollard-factory-farm-advocacy-gains/#potential-solutions-003055

Supplementary Readings:

- Podcasts: The rest of the podcast above with Lewis Bollard.
  - See also the following podcasts:
  - Bob Fischer on comparing the welfare of humans, chickens, pigs, octopuses, bees, and more By Luisa Rodriguez and Keiran Harris, URL: https://80000hours.org/podcast/episodes/bob-fischer-comparing-animal-welfare-moral-weight/
  - Lewis Bollard on big wins against factory farming and how they happened, By Robert Wiblin and Keiran Harris, URL: https://80000hours.org/podcast/episodes/lewis-bollard-big-wins-against-factory-farming/

- News article: Can a college course in moral philosophy convince people to eat less meat? by Dylan Matthews, https://www.vox.com/future-perfect/21354037/moral-philosophy-class-meat-vegetarianism-peter-singer
- Website: https://welfarefootprint.org/
- Website: https://thehumaneleague.org/
- Website: https://80000hours.org/problem-profiles/factory-farming

Theme 3 - What about Future Generations?

Week 10 - What (if anything), Do We Owe to our Distant Descendants? (Chow)

So far this semester, we have seen that if we care about the wellbeing of others, we have reason to hold that we owe duties of moral concern to the global poor and animals. We have also seen through the lens of economics, how we might evaluate and choose amongst possible courses of action in advancing the wellbeing of the global poor and non-human animals.

But what about the welfare of future generations? On one level, it would seem as though this is intuitively necessary: a state's monetary and fiscal policy requires to some extent that economic policy takes into account the welfare of future generations. This has become known in philosophy and economics as the debate over how, if at all, we should discount the future.

This week, we explore the idea that while it would appear intuitive to discount future generations, if we take seriously the idea that welfare matters, it might appear as though discounting future generations is not only inconsistent, but potentially ethically indefensible, as Frank Ramsey (1928, 1931), a philosopher, mathematician, and economist, put it.

Week 10 required readings

- "The Future is Vast", Max Roser
- "The Moral Case for Long-Term Thinking", Hilary Greaves, William MacAskill, and Elliott Thornley

Tutorial 4

- Discussion of long-termism

Week 11 - Existential Risks and Discounting (Mattsson)

Last week, we gave an introduction to the moral issues of how to think about future generations. This week I will show how economists think about discounting future generations and when I think this is valid vs. when I do not think this type of discounting is correct. I then apply this framework to climate change. I will then discuss other ways in which we may improve the long-run future such as reducing existential risks.

Week 11 required readings

- Discounting 101, A review of discounting and how it applies to climate change by Brian Prest. URL:https://media.rff.org/documents/Discounting_Explainer_-_Final.pdf
- Website: 80,000 hours "The case for reducing existential risks" URL: https://80000hours.org/articles/existential-risks/

Supplementary Readings:

- **Book:** MacAskill, William. What We Owe The Future. Simon and Schuster, 2022.
- **Book:** Ord, Toby (2020). The Precipice: Existential Risk and the Future of Humanity. Also see the website: https://theprecipice.com/
- **Video lecture:** "Shaping humanity's longterm trajectory" by Toby Ord, EA Global London 2023, url: https://www.youtube.com/watch?v=jb7BoXYTWYI
- **Video lecture:** "The Precipice Revisited" by Toby Ord, EA Global Bay Area 2024, url:https://www.youtube.com/watch?v=vQ3ml6wcsn4&t=783s

Criticisms and conclusion

Week 12 – Guest Lecture by Peter Singer

- Submit questions for Peter via Canvas

Week 13 - Criticisms of Effective Altruism (Chow)

This week, we turn to criticisms of Effective Altruism in both theory and practice. What are the problems with the underlying theories that effective altruism rely on? Furthermore, what are the problem with how Effective Altruism has been practised?

Week 13 required readings

- **Book review:** Stop the Robot Apocalypse, by Amia Srinivasan, https://www.lrb.co.uk/the-paper/v37/n18/amia-srinivasan/stop-the-robot-apocalypse

- **Article:** How effective altruism let Sam Bankman-Fried happen, by Dylan Matthews, https://www.vox.com/future-perfect/23500014/effective-altruism-sam-bankman-fried-ftx-crypto

Supplementary Readings:

- **Article:** Effective altruism's most controversial idea, by Sigal Samuel, https://www.vox.com/future-perfect/23298870/effective-altruism-longtermism-will-macaskill-future

- **Opinion piece:** The Deaths of Effective Altruism, by Leif Wenar https://www.wired.com/story/deaths-of-effective-altruism/

- **Webpage with rebuttals to some criticisms:** Misconceptions about effective altruism, by Benjamin Todd https://80000hours.org/2020/08/misconceptions-effective-altruism/

Tutorial 5

- Discussion about criticisms of EA