




Arete Fellowship

NYU Shanghai

Fall 2022 Syllabus



Some efforts to improve the world can be much more impactful than others. Smallpox eradication, the abolitionist movement, germ theory, and the Green Revolution demonstrate what's possible when people come together to solve pressing problems. But these opportunities are few and far between. At the other end of the spectrum, well-intended initiatives have been ineffective or even counterproductive. Doing good better is one of the defining challenges of our time. Recognizing this need, Arete Fellowship NYU Shanghai was established to equip students with the knowledge to do good in the world more effectively and the tools to think critically about their career ambitions. By the end of the semester, Fellows have access to in-depth coaching, career connections, and unique opportunities to help others as part of a happy and meaningful life.

We look forward to learning with you.

nyush.arete@gmail.com
afnyush.org

The Arete Fellowship NYU Shanghai

The Arete Fellowship NYU Shanghai	2
Fellowship Structure	3
Expectations	3
How we hope you'll approach the fellowship	4
Attendance	5
Upon Completion of the Fellowship	5
Feedback and Concerns	5
Curriculum	6
Week One: Can You Trust Your Instincts?	7
Week Two: How Can We Save the World?	9
Week Three: Using Your Career for Good	11
Week Four: Global Health and Development	13
Week Five: Animal Welfare	16
Week Six: Considering the Far Future	19
Week Seven: Avoiding Existential Risks	21
Week Eight: Career Planning	23

Fellowship Structure

The Arete Fellowship NYU Shanghai is a branch of NYU Effective Altruism which is part of the global [Effective Altruism](#) (EA) community. EA emphasizes the compassionate desire to do good in the world, the use of evidence and reason to understand how to do so effectively, and the ambition to act. We will discuss related topics together a great deal this semester.

The fellowship takes place over the course of 8 weeks from the week of October 21st, to the week of December 9th, 2022. Each week consists of a set of study resources and a discussion group. The materials take approximately one hour to complete (taking notes is recommended) and need to be finished in preparation for the weekly discussions.

Expectations

The fellowship is **not** a passive activity. It requires active engagement, curiosity, and discussion. Participants should see the fellowship as a mini-course with one hour of reading homework and an hour of section per week. Weekly materials have been carefully picked for their messages and how engaging they are. They should not be completed simply for the sake of the requirement but instead with the goal of gaining new perspectives. Fellows should use the questions in this syllabus following each reading to reflect and test your reading comprehension.

To complete the Social Impact Career Fellowship, fellows must fulfill the following requirements:

- (1) Keep up with readings and attend weekly discussions (see the attendance policy)
- (2) Attend at least one Arete Fellowship sponsored social event

How we hope you'll approach the fellowship

Taking ideas seriously.

Often, conversations about ideas are like recreational diversions: we enjoy batting around interesting thoughts and saying smart things, and then we go back to doing whatever we were already doing in our lives. This is a fine thing to do — but at least sometimes, we think we should be asking ourselves questions like:

- “How could I tell if this idea was true?”
- “If it is true, what does that imply I should be doing differently in my life? What else does it imply I’m wrong about?”

And, zooming out:

- “Where are my blind spots?”
- “Which important questions should I be thinking about that I’m not?”

Taking ideas seriously means wanting to make our worldviews as full and accurate as possible, since we see that having accurate beliefs allows us to make better decisions about things that we care about.

Disagreements are interesting. When thoughtful people with access to the same information reach very different conclusions from each other, we should be curious about why. Often, we tend to be incurious about this simply because it’s so common that we’re used to it. But if, for example, a medical community is divided on whether Treatment A or B does a better job of curing some disease, they should want to get to the bottom of that disagreement, because the right answer matters — lives may be at stake.

Strong opinions, weakly held. Often people abstain from trying to have opinions about things because they think things like “I’m not an expert” or “It’s hard to know for sure.” Instead, during this Fellowship, we invite you to be bold enough to venture guesses, expressed clearly enough such that it’s easy for someone else, or evidence about the world, to prove you wrong. In the long run, we hope that you'll become stronger and more engaged thinkers; this seems more important than minimizing your error in the short run.

Attendance

Fellows are expected to complete materials and attend meetings weekly (and the concepts workshop). If a fellow cannot make a section, they may attend another section that week. If a fellow cannot attend either, they must make up for the absence with a short, 150-word reflection - [Fellowship session summary](#) - on the week's materials (this will usually take at most 15 minutes to write) and submit it to your section leaders via email.

Over the course of the fellowship, there will be several virtual and in-person social events, alongside more discussion-based events. Although not mandatory, fellows are strongly encouraged to attend.

Upon Completion of the Fellowship

Arete Fellowship alumni may list the title of “NYU Shanghai Arete Fellow” on their résumé and are invited to join Arete Fellowship NYU Shanghai projects. They will have the opportunity to continue being part of the larger EA community which offers valuable networking opportunities. Additionally, fellowship leaders are dedicated to helping alumni have the opportunities that will make them more effective altruists and can be valuable for education/career planning.

Feedback and Concerns

We are committed to improving the fellowship and making it an environment that is more productive, constructive, welcoming, and safe. Fellowship discussion leaders are open to constructive criticism. **Additionally, this [feedback form](#) (which can be anonymously completed) can be used for concerns or feedback of any kind.**

Curriculum

Week One: *Can you trust your instincts?* Learn about cognitive biases and start thinking more rationally. Understand why ideas discussed in the Arete Fellowship may surprise and challenge your intuitions.

Week Two: *How can we save the world?* Understand the basics of Effective Altruist thinking. What are the biggest problems facing us, and how do we effectively tackle them?

Week Three: *Using Your Career for Good* Start to think about how an Effective Altruist leverages their skills, interests, and 80,000 working hours across a career to make a difference.

Week Four: *Global Health and Development* Understand why many people think global health is a key cause area to focus on. How should we decide which global health interventions are most impactful?

Week Five: *Animal Welfare.* Think critically about the value we give to non-human animals and their suffering. Why might we include animals in our moral circle? As a challenge, try going vegan for a week!

Week Six: *The Far Future* Some Effective Altruists think the most important problems of our time are those whose impacts will primarily be felt in the future. How much should we worry about people many generations after us?

Week Seven: *Existential Risks* Artificial intelligence, pandemic, nuclear war, climate catastrophe...What are the ways humans are most likely to go extinct, and how can we protect our descendants against these threats?

Week Eight: *Career Planning* Apply the philosophical and practical principles you have developed during the fellowship to your own life by evaluating your current career plans.

Week One: Can You Trust Your Instincts?

“There’s a whole...art to finding the truth and accomplishing value from inside a human mind: we have to learn our own flaws, overcome our biases, prevent ourselves from self-deceiving, get ourselves into good emotional shape to confront the truth and do what needs doing, et cetera, et cetera.”

- Eliezer Yudkowsky

Before we begin applying evidence and reason to the way we do good, we need a desire to do good. This piece from Nate Soares is a sober discussion of obligation, emotion, and morality. *This may be one of the most important readings in this syllabus.* What’s the solution to scope insensitivity? How do we achieve this without becoming too cyborg-ish?

Nate Soares: [On Caring](#) (2800 words)

Many ideas in Effective Altruism are challenging or counterintuitive at first. What is happening in your brain when you rely on intuition and when should you trust rational decision making? Please watch Joshua Greene’s talk, “Beyond Point and Shoot Morality,” and think about some of the dilemmas he poses. What is your gut reaction to these questions? Does your answer change when you go into “manual mode”?

Joshua Greene: [Beyond Point-and-Shoot Morality](#) (11 min – 2x speed available)

Human brains were optimized for survival on the savannah, not complex, globalized societies. Read this short introduction to cognitive biases and then consider two biases that are often relevant to Effective Altruism.

Rob Besinger: [Biases: An Introduction](#) (Intro + ‘Noticing Bias’ section) (900 Words)

Eliezer Yudkowsky: [Scope Insensitivity](#) (500 Words)

Eliezer Yudkowsky: [Availability](#) (680 Words)

There are many more cognitive biases. From this list, read short descriptions of at least three or four that interest you.

The Decision Lab: [List of Biases](#) (Skimmable) (Read a few)

(If you don’t know where to start, Hyperbolic Discounting, Identifiable Victim Effect, and In-group Bias are all interesting and relevant to EA.)

Knowing a bias exists does not make you immune to its effects. Unfortunately, you are also biased against realizing you've been biased! One tool we can use to think more rationally is called expected value. This allows us to use the probability of different outcomes, as well as the benefits or costs of each outcome, to predict an action's value. This includes coin flip games, but also career choices, global health interventions, and decisions about the future.

Devin Soni: [What is Expected Value](#) (750 Words)

If you have any questions about EA jargon and lingo, make sure to refer to the [terms page](#) on our website. And hit us up at our email (effectivealtruism.nyu@gmail.com) if you have any further questions!

Book recommendation: [The Scout Mindset](#), by Julia Galef: if you are interested in cognitive biases and how to accurately form beliefs about the world, this is a great read. Galef explains in detail how to identify when you are using motivated reasoning, instead of evidence, to reach a conclusion. Ask Cian or Sage for a reimbursement!

Week Two: How Can We Save the World?

“The challenge for us is this: How can we ensure that, when we try to help others, we do so as effectively as possible?”

- William MacAskill

This week, we are exploring the basic concepts of Effective Altruism. Firstly, Sam Deere explains that, perhaps unsurprisingly, it’s good to do good – but this concept may have surprising implications.

Sam Deere: [Four Ideas You Already Agree With](#) (1500 Words)

Do you disagree with any of these ideas? Do you disagree with any of the arguments Deere makes about how these ideas should affect people’s life decisions? Why do you think many people who hold these values don’t act on these ideas?

Next, an exploration of the “effective” part of EA. Owen Cotton-Barratt gives a talk (with a full transcript below) explaining how we can figure out if what we’re doing actually is effective, with the illustrative metaphor of searching for gold.

Owen Cotton-Barratt: [Prospecting for Gold](#) (7800 words--skimmable)

Take note of the three key components for effectiveness, and of some key principles outlined for maximum cooperation and truth-seeking. Think about some altruistic causes you have heard about, such as researching breast cancer, aiding animal shelters, preventing climate change, or NYUAD’s COVID-19 fundraising. How do these causes fit into the framework presented here? We will discuss this at our meeting, so come prepared with your own examples.

After thinking about this on your own, read the Centre for Effective Altruism’s overview of EA, which includes information on EA’s mission, community, and goals.

Center for Effective Altruism: [Introduction to Effective Altruism](#) (3200 Words)

Were any of the main cause areas listed here surprising to you?

Remember that EA is cause-neutral. This means it is a movement about doing good, not about the cause areas presented in this introduction and in the fellowship. While these causes are currently seen as effective ways to help the world, EA as a community aims to be responsive and update if new information finds another area to be even more important.

Finally, watch at least **one** of the two talks below.

Though Michael Page from the Centre for Effective Altruism recognizes the biggest challenges to EA, he argues that we shouldn't shy away from the movement. EA is diverse and there is disagreement. How set in stone are your moral beliefs?

Michael Page: [Embracing the Intellectual Challenge of Effective Altruism](#) (option 1) (19 min)

Will MacAskill takes a step back to look at our progress, our problems, and our potential as a civilization. Effective Altruism thinks big. Have human moral understandings kept up with scientific advancements? What should be our key priorities this century?

Will MacAskill: [What are the Most Important Moral Problems of our Time?](#) (option 2) (12 min)

If you need an explanation of a term, check the [EA Terms List](#) created by HCEA. This is a good resource across the Fellowship.

Week Three: Using Your Career for Good

“You have 80,000 hours in your career. Make the right career choices, and you can help solve the world’s most pressing problems, as well as have a more rewarding, interesting life.”

- 80,000 Hours

This week, we are talking about a big way you can use a resource you each have to make the world a better place — your career. How can you use your time and talents to focus your career on doing good? 80,000 Hours is an organization which has a number of resources published on making career decisions through an Effective Altruism framework. First, look over their career guide.

80,000 Hours: [A Guide to Using your Career to Help Solve the World’s Most Pressing Problems](#) (6000 Words)

After reading the article, scroll back up to the “[Five Categories of Career to Consider](#)” section. Read at least **one**, or however many are intriguing to you.

Then, look at the “[Priority Paths](#)” list. Read at least **two**, or however many interest you, and for at least one, explore a few of the links at the bottom. Feel free to look at some of the other recommendations on the list below the Priority Paths as well. These priority paths may not be what you currently predict for your career, but they may be worth seriously investigating. Why are these careers high impact? What are the steps you could take towards the career path you explored?

It may be easy to forget the potential impact you could have if you choose a highly effective career. This talk, also from 80,000 Hours, is pretty inspiring in reminding us how much good it is possible to achieve over the course of our lifetimes.

80,000 Hours: [How big of a difference can you make?](#) (7 min)

On a bigger scale, it is interesting to see how much progress the world has made on many metrics we care about. Looking at the charts collected by Our World in Data (a project at Oxford University) suggests the combined effort of many people’s work has made a significant difference.

Our World in Data: [12 Key Metrics to Understand the State of the World](#)

Finally, we are ending with a few **optional** further resources.

First, another video from the 80,000 hours series (the entire series is worth watching, and available on Youtube [here](#)). This video is only four minutes long, and outlines how to make a Plan A, Plan B, and Plan Z for your career. (This advice is useful whether or not you currently are hoping for a career in an EA cause area!) If you have time, try planning your own A, B, and Z career plans, even knowing they are likely to change.

80,000 Hours: [How to make a career plan](#) (optional) (4 min)

This next reading is recommended for those who are interested in entrepreneurship and social enterprise. If you are considering the for-profit or startup realms, how do you think you could maximize your impact?

Robert Wiblin: [Doing Good Through For-Profits: Wave and Financial Tech](#) (optional) (1100 Words)

And finally, a reading recommended for those who are considering going into policy or governance. What are the key ways to affect change in policy discussed here?

Hiski Haukkala: [Policy Makers Love Their Children Too](#) (optional) (20 min -- 2x speed available)

Week Four: Global Health and Development

"If we can save one thousand lives with one intervention and ten thousand with another at an equal price, then merely moving our funding from the first to the second saves nine thousand lives."

- Toby Ord

This week, we are exploring the first of three often-prioritized cause areas in Effective Altruism: global health and development.

One key question we need to consider is the question of local versus global. Here, read an excerpt from philosopher Peter Singer's book *One World Now*, which discusses the balance between community and impartiality. How should we pragmatically balance these often opposing values?

Peter Singer, [One World Now excerpt](#) (1800 Words)

Next, read an overview of this cause area, with arguments for and against focusing on global health as an effective altruist. Which arguments are you compelled by? Think back to the scale/tractability/neglectedness framework, as well as our discussion of cognitive biases. Why do you think people and governments spend so little on foreign aid?

Jess Whittlestone: [Global Health and Development](#) (3700 Words)

One thing discussed in Whittlestone's piece is the importance of cost effectiveness in this cause area. Toby Ord, a philosopher at Oxford, wrote this piece arguing that cost effectiveness is not only important, but morally obligatory. Reading this article is not required but is highly recommended.

Toby Ord: [The Moral Imperative toward Cost Effectiveness in Global Health](#) (optional) (2800 Words)

This raises an obvious question: How do we know which interventions are in fact cost effective? One group trying to answer this question is called GiveWell. Look through their list of recommended charities, picking at least two to read about more. (Keep your career plans in mind as well!) What do you notice are commonalities between them?

GiveWell: [Top Charities](#) (8 Charities)

Another group which works on figuring out which interventions are most effective is a research group of economists at MIT and Harvard, who recently won the Nobel Prize for their work. The group, called J-PAL, has pioneered “randomista development,” where specific interventions are tested using randomized controlled trials or RCTs. Here, watch a TED Talk from a founder of J-PAL.

Esther Duflo: [Social Experiments to Fight Poverty](#) (16 Minutes)

This method of evaluating interventions is common in Effective Altruism. People appreciate it for being less susceptible to bias than other methods of evaluating interventions.

However, randomista development has been criticized as well. One critique is that it can only apply to quantifiable problems, inherently ignoring other values such as equality or social justice. Further, it is impossible to run an RCT on certain forms of change, in particular social or policy changes which address systemic barriers to prosperity and health. Lant Pritchett describes randomista development as “kinky” development, suggesting it is only able to make small kinks or adjustments in a trajectory towards our goals. This reading is optional, but highly recommended if you are interested in going into global health or development.

Hauke Hillebrandt: [Growth and the Case Against Randomista Development](#) (until 21:07) (optional)

If you choose not to watch this talk, read this quote from the [related EA Forum post](#):

“GDP per capita is strongly correlated with many objective and subjective measures of welfare. Thus, empirical evidence shows that only so much can be achieved for a country at a low level of income per head. If a country has an income per head below \$5,000, it is very likely to do poorly on most objective and subjective measures of welfare. If a country’s income per head is above \$20,000, it is very likely to do well on most objective and subjective measures of welfare.

As discussed above, there are also good reasons to believe that increased GDP per capita causes many of these increases in welfare. This suggests that when we are working out how to increase human welfare to the greatest extent possible, then we should start by figuring out how best to increase GDP per capita... Instead, randomized development (RD) is overwhelmingly focused on evaluating the success of programmatic attempts to solve a problem in a

specific target population, such as depression, educational attainment, intestinal worms or malaria.”

(Some interventions which are prioritized by GiveWell and EA do take social change and political advocacy into account -- for an example of this, look at [Considering policy advocacy organizations: Why GiveWell made a grant to the Centre for Pesticide Suicide Prevention](#) (optional))

Think about your stance on the two big questions explored this week: Local versus global focus, and randomista development versus upstream change.

Finally, an optional reading on a growing EA cause area – mental health. If you are interested, read this overview by Michael Plant.

Michael Plant: [Cause Profile: Mental health](#) (optional) (5300 Words)

Week Five: Animal Welfare

“Animal Liberation will require greater altruism on the part of human beings than any other liberation movement [because] human beings have the power to continue to oppress other species forever... Will we rise to the challenge and prove our capacity for genuine altruism?”

- Peter Singer

Animal welfare is a cause which may be surprising to those who haven't thought about it much before. Among philosophers, there have been interesting discussions about human treatment of non-human animals. Watch this video exploring some discrepancies in our relationships with these beings. Do you find yourself convinced that animals should be included in moral consideration? Why or why not?

Crash Course Philosophy: [Non-Human Animals](#) (9 min)

Effective Altruism concerns itself with preventing suffering as much as possible. This mission has led many to look at the suffering of non-human animals as a particularly large scale, tractable, and neglected cause area. Read this overview of animal welfare from the Center for Reducing Suffering, which focuses in particular on factory farmed animals.

Center for Reducing Suffering: [Altruism, Numbers, and Factory Farms](#) (2100 Words)

The overview touches on the scale of the problem, but reading these numbers can be hard to comprehend. (Remember scope insensitivity!) This gives estimates for the total number of animals in captivity. As you look at the tables at the top, pick a few species you are interested in and really focus on digesting the numbers as much as possible. If you scroll down a bit, the same tables appear (in yellow), but with a duration of time it would take to commemorate the animals' lives with a moment of silence. (No need to read the entire piece)

Saulius Šimčikas: [Estimates of Global Captive Vertebrate Numbers](#) (14000 Words -- skimmable)

With a sense of scale in mind, we now have to consider whether these large numbers of animals actually are worthy of moral consideration. Brian Tomasik makes the case that, from our best scientific understanding, animals do suffer.

Brian Tomasik: [Suffering in Animals Versus Humans](#) (1900 Words)

Is moral status dependent on suffering in your view? Or, what features would something need to have moral status? This is a tough question, and we will discuss this more at our meeting.

One other way to think about this question is through the framework of a moral circle. Some Effective Altruists make the argument that, while animals may not be in everyone's circle of moral concern now, these circles are generally expanding. Over the history of ethics (specifically in Europe and the USA), people have become more open to caring about communities beyond their own, such as people of marginalized nationalities, genders, and races. Does this mean our circles should continue to expand and include animals? Sigal Samuel explores this question.

Sigal Samuel: [Should Animals, Plants, and Robots have the Same Rights as You?](#) (2900 Words)

Similar to GiveWell which we read about last week, an organization called Animal Charity Evaluators (ACE) examines the efficacy of various interventions in this cause area. Explore their recommendations, reading more about at least one or two. (Keep in mind career options if you are interested!)

Animal Charity Evaluators: [Recommended Charities](#) (600 Words)

You read about the scale of this problem already this week. Finally, here are two optional readings, with a bit more information on tractability and a particular area of neglectedness.

OpenPhil, which is a philanthropy aligned with EA, published this post with more information on what they see as the most important strategies for effective animal advocacy. TLDR? Campaign factory farm corporations to make conditions better, create good alternatives to animal meat, and promote animal welfare in Brazil, China, and India. Read the whole piece here.

OpenPhil: [How can you do the most good for animals?](#) (optional) (1700 Words)

A particularly counter-intuitive idea within animal welfare is that of wild animal welfare, which argues that wild animals endure a lot of suffering which is incredibly neglected. If you are curious about this idea, read Brian Tomasik's overview of it here.

Brian Tomasik: [Should We Intervene in Nature?](#) (optional) (1000 Words)

Effective Altruism involves constantly challenging our pre-existing biases and assumptions. Sometimes, that includes disagreeing with the consensus in the effective altruism community. This article has a well-thought out criticism of the way Animal Charity Evaluators, a respected EA organization, ranks animal charities.

Harisson Nathan: [The Actual Number is Almost Surely Higher](#) (optional) (10000 Words)

Week Six: Considering the Far Future

“Insofar as you care about making the world a better place, your key concern should be to increase the chance that the future goes well rather than badly.”

- Benjamin Todd

This week, we are going to think on the biggest scale yet: that of the far future. The long term value thesis argues that much of the potential suffering we could alleviate, and much of the potential value we could achieve, does not exist in the present day but instead in the future. Read Benjamin Todd’s overview of arguments for and against this idea.

Ben Todd: [Presenting the Long-Term Value Thesis](#) (3600 Words)

By this point, we have talked about the idea that, if we value each life regardless of where in the world someone lives, it may be worth our money to help those who are farther from our community in location. Thoughts about the far future suggest that this may be true across time as well – maybe it is important to take future generations into consideration, and try to make their lives better. This question matters because the future will likely be incredibly big, as this short post explains.

EA Concepts: [The Number of Future People](#) (200 Words)

In some ways, we already account for future generations. One example of this is the urgency of addressing climate change, as Martin Rees writes about.

Martin Rees: [The Crime Against Future Generations](#) (900 Words)

Rees points out a number of examples of times we already have a long term-ist’s approach to thinking about the world. We want to avoid asteroid collisions, climate change, and harms from nuclear waste, even though these are likely to affect our descendents rather than just us. Are you convinced by this idea? Do you think you have any moral obligation to future generations?

One concept which is important in this conversation but psychologically hard to grasp is the difference between extinction and near-extinction. This talk gives evidence on how people understand the future and extinction.

Stefan Schubert: [Psychology of Existential Risk and Long-Termism](#) (3500 Words)

Next, please watch/read at least one of the following:

Nick Bostrom discusses what is possible for humanity: never dying of old age or even aging at all, reaching a state of peak happiness. If we can get the far future right, the outcome could be amazing. Do you think “transhumanism,” or modifying human nature to solve our intrinsic problems, is a good idea? Do you predict it will happen in the future? Are we obligated to ensure maximum happiness (and become transhumans), or just to prevent suffering?

Nick Bostrom: [A Philosophical Quest for our Biggest Problems](#) (option 1) (16 min)

Martin Rees discusses the opportunities, and potential pitfalls, available to people alive today. It may be the case that we have the potential to make dramatic changes to the far future, but our window of opportunity could close soon. Do you agree these problems are urgent? Or do you predict humans will solve them in the future?

Martin Rees: [Is This Our Final Century?](#) (option 2) (17 min)

Evan Williams writes about the state of the world today. Every other era in history has suffered from significant moral failures, and there are so many things which could be going wrong right now. This may suggest that in a few years or centuries, humanity will look back at our time period very critically, and we should be applying the same scrutiny to ourselves today. What are the biggest moral catastrophes you think are currently ongoing?

Evan Williams: [The Possibility of an Ongoing Moral Catastrophe](#) (option 3) (7000 Words -- skimmable)

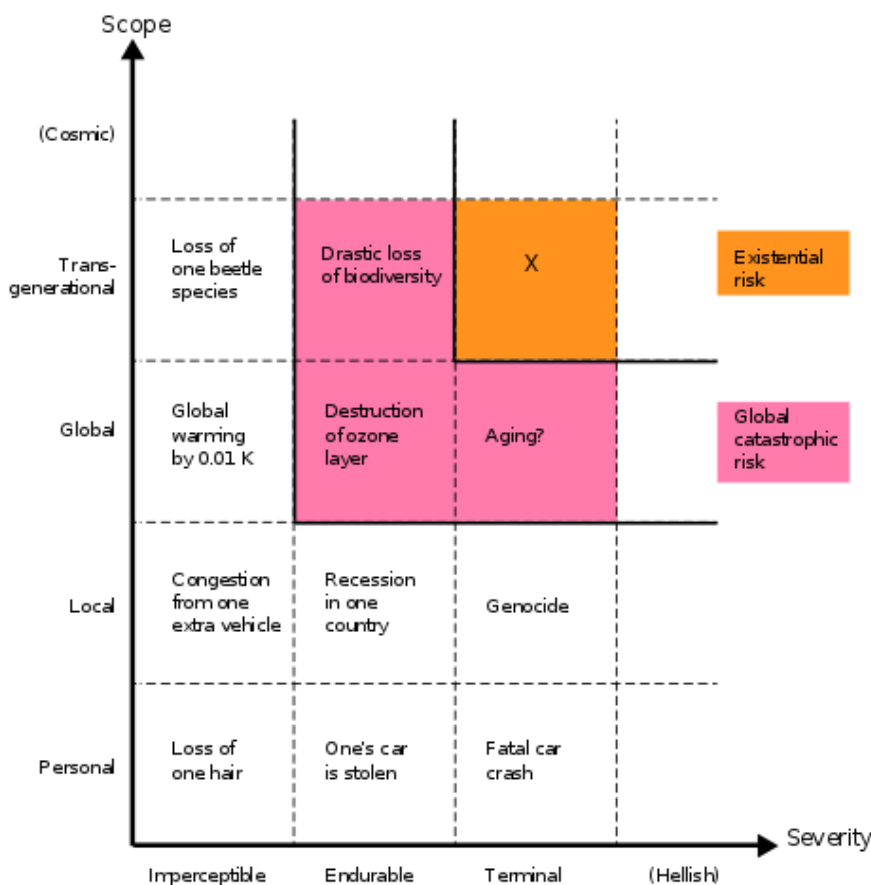
Week Seven: Avoiding Existential Risks

“If we can preserve civilization long enough, we may be able to eradicate poverty, achieve peaceful globalization, cure most diseases and mental health issues, genetically modify ourselves to be happier and healthier, engineer our biosphere to reduce suffering of all kinds, spread out into the cosmos... It’s a long way up from here if we can manage to keep surviving and making progress.”

- Anonymous

Last week, we talked about why the far future might matter a lot. This week, we will discuss what to do about it. Which interventions and causes are most important in the far future? How can we avoid a global catastrophic risk? Read this overview from Open Philanthropy Project on this question.

Open Philanthropy Project: [Global Catastrophic Risks](#) (1200 Words)



What do we mean when we talk about a “global catastrophic risk”?

If we care about the long term future, this chart is important. We want to avoid things in the pink boxes, and in particular we want to avoid anything in the orange box. But what might the orange box contain?

Artificial intelligence is one new technology which many people in Effective Altruism think has the potential to dramatically change the future. This technology is often misunderstood, but important to think seriously about. This blog post, and its sequel, may be two of the most interesting things you will read. They are a bit long (but not too dense) so skim if you need to, but try to absorb the main ideas.

Tim Urban: [The AI Revolution](#) (7500 Words) and [The AI Revolution Part 2](#) (16000 Words - skimmable)

Next, read at least **one** of the following profiles on far future cause areas. Keep your career plans in mind if any of these cause areas are interesting to you.

Roman Duda: [Biorisk Reduction](#) (10000 Words)

Peter McIntyre: [Nuclear Security](#) (2500 Words)

Roman Duda: [Climate Change \(Extreme Risks\)](#) (5000 Words)

Robin Wilbin: [Positively Shaping Artificial Intelligence](#) (5500 Words)

These are not the only things which may be major influences on the far future. Here, 80,000 Hours describes some other causes which may be important from a long term perspective. Read at least two of the short descriptions of these cause areas.

80,000 Hours: [Potential Longtermist Issues](#) (2500 Words)

From all of these options, be prepared to share a bit about what you have read with other fellows in your discussion section. And, think a bit on your own about any other issues you think may have important downstream effects in the future.

Caring about the far future does not mean we want to ignore the present – for example, it may be the case that increasing access to education in our generation leads to positive long term effects, such as more knowledge and research in the future.

However, these choices can be risky, because unlike the RCT evidence we use in global health interventions, there is less definitive evidence about the far future. Most of why many EAs prioritize the far future comes down to the expected value: there may only be a 1% chance your action makes an impact, but if it were to make an impact, this impact would be so large that the expected value is large. If we are risk-neutral, and go off expected value, the long termist view may be smart. But, if we are risk-averse, this may not be a path worth taking.

These decisions about risk-aversion versus risk-neutrality are important when considering the far future, but are not unique to this cause area. This blog post from GiveWell describes an intervention of deworming young people, which has a low percentage chance of being effective but a high impact if it is effective. If you are interested in these decisions, feel free to read the optional post.

GiveWell: [Deworming might have a huge impact, but might have close to zero impact](#) (optional) **(2000 Words)**

Where do you fall in your level of comfort with risk in decisions about where to donate or focus your career? How do you personally want to balance well evidenced interventions versus riskier interventions which may have a far greater impact?

Week Eight: Career Planning

“Most of us want to make a difference. We see suffering, injustice and death, and are moved to do something about them. But working out what that ‘something’ is, let alone actually doing it, can be a difficult and disheartening challenge. Effective altruism is a response to this challenge.”

- Center for Effective Altruism

This week, as you are almost done with the Arete Fellowship, we want to discuss how to keep thinking about and acting on the ideas we’ve talked about as aspiring effective altruists.

There are three steps to prepare for this session:

Step 1: Sketch an outline for your [A/B/Z Career Plan](#) using [this](#) template.

This guide will help you to develop a model of your ideal career plan based on thinking through what problem areas you feel are the most important to work on, the types of roles you would like to have and the sorts of career capital you hope to develop while young. It will have you come up with a Plan A, your ideal scenario, Plan Bs, nearby alternatives, and a Plan Z, your temporary fallback if everything else fails.

Step 2: Make profile on [EA Hub](#)

EA Hub is a networking tool that will help you to easily connect with professionals working directly on EA prioritized cause areas. After setting your profile you should look up other profiles filtering by expertise, career interest, and cause areas. If these groups apply to you, you can do this here too: [EA Consulting](#), [EA Finance](#), [EA Medicine](#), [EA Supply Chain](#). After doing this, **email/message three people** (at least 1 similar to your A, and at least one similar to one of your Bs) to set up conversations. Alternatively, ask your facilitator to make an email introduction for you.

Step 3: Find at least one job or internship adjacent to your A (if you have somewhere else you would rather look, that is fine, too):

There are many resources with opportunities open to undergraduates including the [EA Internships Board](#), the [80k Jobs Board](#), [EA Job Postings](#), [EA NYC Jobs](#), and other [social impact job boards](#). Do you have your own project idea? Check out some of the [available grants from EA](#) for individuals and organizations working on projects!

During the session we will work on drafting applications so it may be helpful to come with an up-to-date resume or past cover letters to adapt.

Here are some additional ways to integrate EA into your life:

Firstly, keep in touch with the EA movement! Here are a few newsletters which are related to Effective Altruism. It's very highly recommended you sign up for them.

Center for Effective Altruism: [Newsletter](#)

Updates on research, job opportunities, and the EA community.

80,000 Hours: [Newsletter](#)

New postings to their job board, and updates on podcasts and interviews with career advice.

Vox's Future Perfect: [Newsletter](#)

Vox's column on Effective Altruism, with interesting updates and explainers of EA perspectives on current events.

If you are on Facebook, feel free to join EA communities there, including [Effective Altruism Group](#), and [Effective Altruism Job Postings](#) (plus maybe [Dank EA Memes](#) if you like)

Another amazing resource for reading and writing in the marketplace of EA ideas is the [Effective Altruism Forum](#). The EA forum is one of the best successes of the EA movement at building community and finding solutions to challenges, and as Arete alumni [you are certainly qualified](#) to post as well. If you want regular forum updates, you can subscribe to the [Forum digest](#) as well. Please find **at least one post** on the Forum which is interesting to you and read it.

EA Community: [Effective Altruism Forum](#)

One thing to think about in your own life is whether, and how much, you hope to donate. This is an interesting calculator to see, on a global scale, [How Rich Am I?](#) Try your family income, minimum wage income (\$15,080), and whatever other numbers you're interested in. If you choose to make a more serious commitment to giving, you may be interested in [One For the World](#) (1% giving pledges) or [Giving What We Can](#) (flexible, usually 10% giving pledges)!

If you're still interested in learning more, check out the [Resources page](#) on our website.

Finally, read the last post in 80,000 Hours' career guide.

Benjamin Todd: The End: [A Cheery Final Note – Imagining Your Deathbed](#) (500 words)