

TRANSCRIPT

Three Sentences That Could Change the World — and Your Life

The philosopher William MacAskill lays out the case for longtermism: “Future people count. There could be a lot of them. And we can make their lives better.”

Tuesday, August 9th, 2022

[MUSIC PLAYING]

Ezra Klein

I’m Ezra Klein. This is “The Ezra Klein Show.”

Today’s show is built around three sentences that could — I think that should — change the future. Here they are: Future people count. There could be a lot of them. We can make their lives better. That is it. They sound obvious, even banal, but two things are true about them. First, we do not act as if they are true. We do not act day-to-day as if future people count. We do not take seriously our responsibility to them.

And second, buried in those sentences are startling possibilities, particularly the second one — the number of people there could yet be is mind-boggling. You don’t need particularly strange assumptions to think it plausible that future people could outnumber current people by a ratio of 10,000 to one, 100,000 to one, a million to one, or even more. Those three sentences are the foundation of a worldview called longtermism.

And I don’t agree with everything that its most hardcore adherents take from it. You’ll hear that in this episode. But I even more profoundly disagree with how much most of us dismiss or ignore in it. And day-to-day, I include myself in that number. This is a worldview worth grappling with, maybe not the only one we should use to evaluate our actions and our policies, but certainly one of the ones we should be using.

And so I asked Will MacAskill, a philosopher at Oxford University, a founder of the effective altruism movement, and the author of the forthcoming book “What We Owe The Future,” which is probably the best single book distillation of longtermism you’ll find, to come on the show and talk through this worldview, its ideas, and its challenges.

As always, my email, if you have guest suggestions, feedback, recommendations for things we should read or watch or listen to, ezrakleinshow@nytimes.com.

[MUSIC]

Will MacAskill, welcome to the show.

William Macaskill

Thanks so much for having me on.

Ezra Klein

So something I love about longtermism is that it’s really just built on three sentences. Future people count. There could be a lot of them. We can make their lives better. Walk me through them.

William Macaskill

OK, I’ll take each one in turn. So the idea that future people count — that is, their interests matter morally, and we should take their interests seriously in just the same way we take the interests of people alive today seriously. I think that idea is just common sense.

So when we consider the issue of climate change, sure, there are lots of harms that will impact us in the present generation, but it will also — you know, the carbon dioxide that we emit today, a significant fraction of that will stay in the atmosphere. And it seems perfectly intuitive that that’s a big deal — the geophysical impacts we have on such timescales, we should pay attention to.

Or when we consider radioactive nuclear waste, again, we are thinking thousands of years into the future, where if you harm someone, it really seems to not matter intuitively whether you've harmed that person in a year's time or 10 years time, or 100 years time. Harm just matters morally, whenever that occurs. So that's the first idea.

The second is just that there could be enormous numbers of future people. So I think that we might stand at the very beginning of history, where there have been 300,000 years of Homo sapiens existence in the past. If we last as long as a typical mammal species, we've got about 700,000 years still to go. If we survive past that using technology to mitigate risks to our own extinction, well, the Earth will remain habitable for hundreds of millions of years.

And perhaps we could even extend humanity's lifespan beyond that if we take to the stars, and then we're thinking billions of years. And the key thing is not like any particular, precise estimate of how much — what humanity's life expectancy is. We don't know that for sure.

But on basically any reasonable estimate, if we don't cause our own extinction in the next few centuries, then the number of the future of people could be absolutely prodigious, tens of trillions of people, or maybe even trillions of trillions of people, such that the number of the future people, even on a very conservative estimate, outnumber us something like 1,000 to one.

So then the final claim, that we in the present really can influence how the course of the future goes, that's what I dedicate most of "What We Owe The Future" to discussing. But I think there are two main ways of impacting the very long-term. One is by reducing the risk of unrecoverable civilizational collapse or human extinction, such as by ensuring we don't develop next generation weapons of mass destruction, like bioweapons, or just trying to prevent a third world war that could end us all.

And then a second is by improving the quality of life of those in the future, such as by avoiding perpetual totalitarian state, or the takeover of AI systems that could devote — mean that future civilization is just alien and valueless from our perspective. So those are the three core ideas of longtermism.

Ezra Klein

So I want to explore them a little bit more. But before we get into them, I want to open some space for skepticism, because these all sound simple, but they're quite controversial, where they lead is quite controversial. And you yourself say that when you first heard the arguments for longtermism, they, quote, "left you cold." Why did they leave you cold?

William Macaskill

I think two main reasons why these ideas left me cold. So at the time, I was co-founding an organization called Giving What We Can and my interest was very much focused on improving the lives of the very poorest people in the world. So there are many extremely urgent and pressing problems, not least the fact that thousands of children every day die unnecessarily of easily preventable diseases.

And so when faced with problems of such salience and urgency, the thought of then taking action for people I haven't met that don't even exist yet, that's like an abstract argument clashing with the very real pull of problems around today. And, you know, I was initially resistant to that.

And then the second is just the question of, well, OK, even if I grant that there is an enormous amount at stake when it comes to the future of civilization, well, what can we even do about it, where the particular — at the time, and we're talking about 12 years ago — the particular avenues for making a difference seemed just extremely speculative to me.

And that just made me think, look, I want to be doing something that is more reliably making a difference, where I'm learning, I'm getting feedback loops. I'm able to tell that I'm actually making a positive impact, and then improve how well I'm doing from there.

Ezra Klein

What changed for you?

William Macaskill

Two things changed. One was just the weight of the arguments over time, where, in particular, the arguments for not really thinking about the long-term future and not really considering the interests of future people — well, those future people are disenfranchised. They are of necessity voiceless.

And it just started to really seem to me that concern for future generations was the next step of moral progress, that in the past, people in positions of power have made moral progress by taking seriously the interests of those who are unfairly disempowered. Now, this is a bit of a different case, because those in the future — there's nothing we can do to empower them, but we can at least take their interests more seriously.

And then in terms of the second aspect of, well, can you actually make a difference, over the last 10 years, I've just seen enormous progress to make me think that, yes, actually we can. So the field of A.I. safety — it was utterly fringe back in 2009 — is now a relatively mainstream area of machine learning research, at least comparatively mainstream and respectable. And you're able to see of the ways, even now, in which A.I. systems are doing things that we really don't intend and concrete work we can do to reduce those risks.

And then, to an even greater degree, one of the big risks that was being discussed all the way back in 2009 was risk from pandemics. And in the broader effective altruism community, we were encouraging people to work on pandemic preparedness, we were funding organizations in that area from about 2014.

And the fact that — I mean, we didn't make nearly as much progress as I would have liked, but I think we did something to help. And now, when I think about what the things we could be doing, such as early detection systems — so just scanning wastewater all around the world for new pathogens — or certain methods of being protected against respiratory diseases and new pathogens, like there's a certain type of lighting that is safe for humans but sterilizes airborne pathogens — so far-UVC, it's called.

It really does seem like with funding and a bit of activism, we could get that installed in most light bulbs around the world. And that really could make an enormous difference to reducing the chance of the next pandemic, which could be a lot worse, while also at the same time having enormous benefits for the present, potentially just eliminating airborne diseases. And so it really seems like these things actually are a lot more tactical than I thought they were about a decade ago.

Ezra Klein

So I want to go through a couple of these premises, and I want to start with the first one — future people count. It sounds obvious. The word you used was intuitive.

William Macaskill

Yeah.

Ezra Klein

And I think the evidence is it isn't intuitive. I think it sounds obvious, but we obviously make many decisions without really counting future people at all. And nor is it simple how you would do that. Should I think of somebody born hypothetically, potentially, 1,000 years from now? Should they have the worth of one child now, of 0.1 children now, of 0.01 children now? How do you decide how much future people count?

William Macaskill

So a thought experiment I give is just dig up some glass while hiking on a trail, and you then find out that someone has cut their hand on that glass. Now, does it make a difference if that occurred next year or if it was a message from the future?

I think as long as you've got the certainty that this harm has occurred, I actually do think it's common sense, morally speaking, to be just as concerned about that as if you'd, say, harmed a stranger. And when we don't pay attention to the interests of future generations, I think it's got more to do with things like uncertainty.

Ezra Klein

The next sentence there is there could be a lot of them. And this is where the math of your book is a lot to take on. So I want to quote here. "To illustrate the potential scale of the future, suppose that we only last as long as a typical mammal species, that is, around 1 million years, and assume that our population continues at its current size. In that case, there would be 80 trillion people yet to come. Future people would outnumber us 10,000 to one."

And I'll be honest that this is both very compelling to me — it's hard logic to deny — and is often where a longtermism leaves me a little unsettled, because it creates a kind of mathematical blackmail. There are so many hypothetical lives in the future — 80 trillion people — that anything that makes them on any calculation infinitesimally better is worth more than almost anything we could do now.

And something about that seems wrong. I think it can even seem dangerous. You can imagine really starving people of resources now to make very uncertain improvements in the future. How do you think about that trade-off?

William Macaskill

Yeah, so I think this idea of, oh, we should just be fanatically pursuing whatever is our best guess at promoting the fairly long-term is also something I'm like — even no matter how tiny the probability, is also something I would feel very nervous about. And if that were the case for longtermism, I just wouldn't be promoting it.

But that's OK, because some of the risks that we're facing, such as from a third world war, or from a worst case pandemic, or from the creation of AI agents that are even smarter than humans, these are really — they're as big as the risk of dying in a car crash, maybe even greater. And we should — so then definitely not in this vanishingly small kind of probability territory.

Then, the second thing I'll say, which is a more substantive way in which this reasoning I don't think always works is that sometimes people think, oh, well, the only way you can impact the very, very long-term future is by reducing the risk of extinction, and so nothing else, to a first approximation, matters apart from that.

And in "What We Owe The Future," I'm kind of arguing against that perspective. I'm actually arguing that there are lots of ways of impacting the long-term future. And in fact, you know, at the moment, the amount of resources that are being spent trying to protect the interests of future people is vanishingly small. Maybe it's like 0.01 percent of G.D.P. I don't know exactly, but it's very low. And I'm saying, it should be higher than that, maybe 0.1 percent, maybe even a whole percentage point.

Ezra Klein

Let me get at the uncertainty of some of this, though. So imagine we go back to an example you just gave, which is preventing a world war. Seems good. I'd prefer no world wars.

William Macaskill

Yeah.

Ezra Klein

But imagine trying to operationalize that question. So somehow I give you a button that could have prevented World War II. Would that have made the world better or worse off?

William Macaskill

Would — what happens — what's the counterfactual? Is this a case where —

Ezra Klein

That's a bit of my point, that — I mean, obviously, the famous version of this counterfactual is, do you kill baby Hitler?

William Macaskill

Yeah.

Ezra Klein

But at different points when World War II is looming in a more near-term way, I think we look back on people who are trying to prevent it. And we call them now appeasers, and we think they were making decisions that were foolish or potentially morally wrong. Now, obviously, it would be good to not have the Nazis rise up. Is it good that we invented nuclear weapons? I don't know.

I only ask this because it seems to me the band of uncertainty, even in things that seem really, really bad, in the very long-term analysis, is very hard to narrow, which makes me somewhat skeptical that it is so knowable even for things that seem quite good.

William Macaskill

So I think — I mean, you point to a very important issue, which is that you can have competing considerations on either side. So in the case of World War II, as a matter of fact, that war meant that nuclear weapons were developed and used, which is obviously increasing existential risk.

However, the alternative, the Nazis really did have aims at large-scale domination over the world. They had a horrific ideology. And they wanted that ideology to persist for a very long time, maybe even indefinitely. And again, in the book, I argue that they really might have succeeded.

I don't think they were close to ever winning the Second World War, let alone having world domination. But if you tweak history such that that really did happen, I think a global totalitarian state that was enacted now via future technology could persist indefinitely.

Now, we get into this really hard question of, OK, we've got perhaps potential increase in extensional risk from a war, but perhaps that is a war that is necessary to prevent the rise of some dangerous ideology. Maybe that's just an extremely hard question.

If so, that's just circumstance we're in and I think we should face that. But I think we can make progress, at least by starting to think precisely about this, for example, thinking how much worse would the future be if it were a global totalitarian state forever, compared to a liberal, democratic, thriving future civilization?

Secondly, how much is the risk of unrecovered collapse or all-out human extinction going to increase on the basis of some war? And I think we can at least start to make our reasoning more precise there. I know that you have worries about the use of precise probabilities. I think it's just extraordinarily helpful to advance debates to be able to use such kind of precise probability estimates.

And then, I think, we can start to at least reason to say — OK, yeah, no actually, I do think maybe the risk of ideological lock-in from fascism is actually sufficiently great and sufficiently bad that it would be worth, let's say, a 0.1 percent increase in extinction risk or something.

[MUSIC]

Ezra Klein

Something I found striking in the book is how important values change becomes for you under a long-term analysis. Tell me why.

William Macaskill

So the reason values become so important is twofold. One is that — and this is a way in which I differ somewhat from some other longtermist thinkers — when I look to the long-term future, I see — I think there could be a dramatic amount of contingency in what moral values guide the future. And I talk a lot about the abolition of slavery as a historical example of something that, at least arguably, was significantly contingent.

We could have a world which has today's level of technological development, yet widespread slavery, were it not for particular cultural changes and moral campaigns that occurred. And that's huge. I mean, if you look to a future that does or does not have legally permitted slavery, that makes an enormous difference to the value of those societies that are to come.

But then secondly, I think that future technology could enable what I call value lock-in, where a particular ideology or a particular set of moral beliefs gets entrenched indefinitely. And it's a more complex argument, and I think that there are many technologies or social developments that could make this happen.

But a crucial one is artificial intelligence, where if there's at some point of time — and that time might be much faster than you might naturally think — when the ruling beings are digital rather than human, well, then power could be in the hands of beings that are in principle immortal. And so a significant reason for why we get moral change over time — namely, that people die, and including harsh dictators die, that would be removed.

And I actually think it could well be that the values that predominate today, or over the coming centuries, could actually be the same values that guide the entire course of the future.

Ezra Klein

Let's go to your slavery example for a bit here. And I'd like you to tell the story you tell there in somewhat more detail, because I think it's a good way of grounding this idea that values and values change can be contingent. So why do you think the abolition of slavery was contingent?

William Macaskill

Sure. So for context, slavery has been — for most of history since the agricultural era, slavery is almost ubiquitous. It was prevalent in most societies in one form or another. And so we should ask — and obviously, there was always enormous resistance, the slaves themselves —

Ezra Klein

And it should be said, there's still pockets of slavery today, but —

William Macaskill

There's still pockets of slavery today.

Ezra Klein

— in terms of large-term values, it isn't seen as acceptable.

William Macaskill

In the world today, especially if you include forced marriage, something like 0.5 percent of the world's population is in some form of slavery. But for context, in 1700, that number was like three quarters of the world's population. So the sheer prevalence of slavery in history is often underappreciated. I didn't appreciate it until I really started learning about this.

Yet, that changed in the course of the 19th century, in particular, although — there was an abolitionist movement that was successful. And the major empires of the world abolished slavery, though some countries still had slavery well into the 20th century — legally permitted slavery.

And so we might ask, why did that happen? And again, before learning about this, I thought, oh, it was economic matters. Mechanization and automation meant that slavery became obsolete, rather than the result of moral campaigns from those who opposed slavery. I think that argument doesn't work. And actually, the mainstream view among historians who study the topic is that the economic argument for why slavery ended does not work.

And there's many reasons for this, but just a couple of the — firstly, the slave trade was booming at the time of abolition.

The number of enslaved people transported across the Atlantic was increasing. There was an intense lobby from slave owners who were making significant profits. So that's kind of one line of argument.

The second is just that, even in the world today, there's enormous amount of unmechanized labor. There's also just many uses of enslaved people historically that — such as sex slavery or being a household servant — that just will not for a long time be threatened by mechanization. And so this kind of argument just doesn't really explain why it happened when it did, nor why it happened in such a kind of further going way.

In contrast, you can look to moral changes that were happening and look to the particular abolitionist campaign that arose out of the Quakers, originally, actually starting in the 17th century, but then were very successful in convincing both the British public and the British elites to change their laws and attitudes to enslavement, and then undertake really quite striking costs to abolish slavery, where the British empire suffered costs of about two percent of their gross national product for about 60 years, which is several times — four times, in fact — what Britain currently spends on foreign aid.

And so the overall picture is just that what happened here was there was cultural change, there was moral activism and that was the main driving force. And the economics of it, while also obviously relevant to what happened, are not the thing that ended slavery.

Ezra Klein

Tell me about Benjamin Lay.

William Macaskill

Benjamin Lay was one of the early abolitionists, born in the late 17th century, and an activist in the early 18th century. And he was a Quaker. He was also a dwarf, and referred to himself as little Benjamin, likening himself to little David, who slew Goliath.

And he is just this remarkable character who became well known across all of Pennsylvania for engaging in guerilla theater and very bold activism because he deeply opposed slave owning and thought it was inconsistent with Quaker Christian principles. And so he did just — there are just dozens of incredible stories about his activism.

So he would shout and harass slave owners at Quaker meetings, get kicked out of the meeting, and just lay face down in the mud so that everyone leaving the meeting just had to walk over him in order to leave, or he would just stand in the snow in kind of bare feet. And then when people were asked, why is he doing that, he would point out that enslaved people would have to suffer like that the entire winter.

Or at one point, during a Quaker meeting, he brought a fake Bible into the meeting, and said that by owning an enslaved person, you are committing as grievous a sin in the face of God as you would be by — as he does by stabbing this Bible, which then sprayed fake blood over everyone in the audience. And so he became incredibly well known.

And it's not entirely documented, the kind of paths for this influence, but it seems they were very great, in particular on the next generation of Quaker activists, and in particular Anthony Benezet, who was kind of the much milder, much meeker — repackaged many of these ideas in ways that made them continuous with Enlightenment thought, and then was enormously influential on getting the abolitionist campaign started in Britain.

Ezra Klein

You argue, in part based on this story, but also based on this larger idea, that values can be changed and that value change can ripple forever. For the value of being a moral weirdo — what is a moral weirdo to you? And who are some great moral weirdos today?

William Macaskill

So a moral weirdo for me is someone who believes things that are not widely accepted by society. Perhaps they're even looked down upon or ridiculed by wider society. Nonetheless, that person stands up for what they believe in. And that doesn't necessarily mean that the right thing to do is to engage in guerrilla theater. Perhaps it's like Benezet — you should take a more cooperative or diplomatic

approach.

But nonetheless, you're willing to go out on a limb for what you believe in morally. And I think that can be enormously impactful. So in the book, I talk about a movement that inspires me in particular, and I know inspires you, which is the animal welfare movement.

I talk about Leah Garcés — there's a story — who was the president of Mercy for Animals. But she's just a representative of a much wider array of people who are speaking out for concern for the interests of non-human animals, and have actually had enormous success at getting, essentially, all the major fast food chains, all the major retailers, to move away from caged eggs and make pledges to say they will no longer use caged eggs.

And throughout history, I mean — a striking thing, actually — it's a digression — is lots of the early abolitionists were also vegetarian. So this was true of Benjamin Lay. It was true of Benezet. I think there's a lesson there. But yeah, animal welfare activists, of which Leah is just one, are just these, in my view, moral heroes, where tides are changing and model attitudes are improving, I think quite rapidly, with respect to the interests of non-human animals.

But even very recently, even 15 years ago when I first became vegetarian, it was just easily the subject of ridicule. And people found it kind of laughable that you would care about the interests of non-human animals. And so I'm really excited to see that change.

Ezra Klein

I mean, it still is, soy boy.

William Macaskill

Exactly.

Ezra Klein

But I want to push you on this point about moral weirdos, because I would say you had to have, as an example, the weirdos of PETA, so that Leah Garcés — who I admire enormously — she was on this podcast back when it was at Vox — so that Garcés could be normal, right? So she could actually have as mainstream a role as she does. And so I want to push you again on this. I want to know who you think are the moral weirdos today. Who are out there saying things that are not mainstream, that are maybe even a little bit reviled, and who refuse to pipe down? I think it's a bit of a dodge to only locate these people in the deep past. Who do you look at today and think, huh, in 50 years, we may look at them and say, they were on the right side of this and we should have listened?

William Macaskill

Sure. I mean, someone who clearly leaps to my mind here is Eliezer Yudkowsky, where I have lots of disagreements with Eliezer, but he was advocating for taking seriously the risk of advanced artificial intelligence for years and years, just kind of in the wilderness on his blog, with some followers. And I think he would be complimented by the title of a weirdo, and he certainly is one. And lots of people find him very weird.

And like I say, I have disagreements with him. But I think the fact that he was able and willing to advocate for what is, I think, a very important cause over the course of decades, that's just very admirable, I think.

Ezra Klein

The moral philosopher Peter Singer once argued to me that moral reasoning is a social activity, by and large. We mostly figure out right and wrong in terms of how those around us think and act about right and wrong. Do you think there's truth to that?

William Macaskill

Yes, I think that — and this is something I'm almost wanting to push against a little bit — is I think that, by and large, for morality — I mean, for many things, but morality in particular, we tend to just go with a group and we tend to assume the beliefs of our peer group. And it really is quite costly.

I mean, if you're a lefty, and all your friends are on the left, and you're having a conversation about politics, it just often is — depending on your group — but can be costly to then voice that heterodox opinion.

And that's got nothing to do with the left in particular. That's also true on the right. It's also true of many, many political interest groups.

If you're part of the environmentalists, and then you come out in support of nuclear power, then people have a hostile reaction to you, at least in many circles.

And so I think we're very natural to pick up the beliefs of our peers. And I think the power of reason that Peter Singer points to is that we can actually try and move beyond that, and just really think things through, and then with courage stand up for those cases where we think that the people around us are making a mistake.

Ezra Klein

Well, one thing I take from that — because I do think it's true — and I think, look at how we treat dogs versus look at how we treat pigs, if you doubt it. But then it really matters not just how individuals reason morally, but how cultures and groups do.

Benjamin Lay, as you mentioned, was a Quaker. And I think the Quakers have an unusually impressive history of being on the right side of big issues quite long before others are. They seem very moral. What in your view makes Quakers so good at collective moral reasoning?

William Macaskill

It's a great question. And I think there are two aspects to desirable moral social epistemology, you could call it — culture that is good for making moral progress. One is just everyone is unified in really wanting to act in accordance with what's morally correct. That's not true for all cultures.

And then the second — and so kind of just taking morality seriously is a big core of it. And then a second is being willing to tolerate dissent. And in fact, you can even encourage diversity of ideas, and intellectual engagement, and so on.

I think with the move towards Protestantism within Christianity, one distinctive part of that was there was a culture that was more about anyone can access the word of God by reading the Bible, by thinking things through for themselves, rather than it being dictated on high. And I think that was a completely important cultural transition.

And so that doesn't single out the Quakers in particular, but then I think they do have this — at least historically — had this distinctive thing of really taking morality seriously. I mean, the idea of — the word Quaker comes from quaking before God. I mean — and also then, there's these familiar Quaker meetings where it's in silence, and anyone can speak, and the word of God can speak through that person.

This, again, is this intellectually egalitarian ethos, where, in principle, anyone could have a new moral insight. That could be God speaking through you. And that, I think, is hugely important. I mean, not the appeal to gods, but the intellectual egalitarianism I think is hugely important in terms of being able to make moral progress, because you get a greater diversity of ideas. You get a greater clash of good arguments.

Ezra Klein

Something you wrote that I've been reflecting on is that, quote, "We might naturally think that slave owners really knew deep down that what they were doing was wrong and that they didn't care. We should be careful not to imagine the views and values of other people as being more similar to our own than they really are."

And you go on to note that quite a few famed thinkers who dedicated much of their lives to moral reflection accepted slavery. So that was true for Plato, for Aristotle, for St. Augustine, for Aquinas, for Immanuel Kant. What are some ideas that come to mind to you that are not only accepted today socially, but defended by prominent moral thinkers, but you suspect may come to be seen as abhorrent in 300 or 3,000 years?

William Macaskill

So we've already talked about animals. And that's, in my experience, when I raise this thought experiment with people, they immediately point to animals. So —

Ezra Klein

Which is damning, we already know it.

William Macaskill

Which is damning, absolutely. A second, I think, is the treatment of prisoners. So think of the abhorrence with which we think about corporal punishment. So if in the U.S., there was someone who was beaten as a punishment, there would be absolute outrage.

But now ask yourself, supposing that you were rightly or wrongly convicted of a crime, and you had the choice — you could spend four years in jail, separated from your partner, from your kids, from everything you know, from your work, or you could be whipped 20 times. Which would you choose? I think probably you would choose to be whipped. I mean, I'm not sure — I certainly would.

But then that suggests that the harm that we would be doing to you by putting you behind bars for four years is greater than the harm of being whipped.

But what's the difference? Well, it's much less salient. It's much less visceral. And so, yeah, I think it's very plausible that the way we will regard the prison system today is with utter abhorrence. Whereas, other people defend it. It's good. People deserve to be punished.

And then a final one I think is greater weight given to co-nationals, where very many people endorse nationalism. And there are some ways in which that can be good. I'm Scottish. I love Scotland. I love the cultural traditions. But that's very different from saying that the people in my nation are more morally important than the people in other nations.

Whereas, that is a fairly common view. And yeah, again, I think a future world where we've made moral progress will take much more seriously the moral idea that everyone's interests should count equally.

[MUSIC]

Ezra Klein

Tell me about your idea of plasticity.

William Macaskill

The idea of plasticity is that there are moments in time when change is much easier. So an analogy I give is of history as molten glass. When one is glass-blowing, while the glass is hot, it can be molded into many different shapes. And the shapes could be beautiful in the end. They could be disfigured. They could have cracks in them. But then when the glass is harder, it's much harder to change things.

And I think you often get periods of plasticity on a small scale, where after the Second World War, I think there was enormous plasticity in the global order, where there was the formation of institutions like the v There were decisions to be made about the use of — how to handle politically the use of nuclear weapons, where one plan on the table was that the U.N. would just control all fissile material and no one would have nuclear weapons.

Now, that's just so outside the Overton window of what policy proposals for handling nuclear weapons — that's obviously not going to happen now, but it at least was on the table after World War II. So I think in periods of crisis, or when technologies are new, or where institutions are new — so after the formation of the United States, there were many constitutional amendments. In the last 50 years, we have only had one constitutional amendment, and even that one was proposed in the late 18th century.

At those periods, things can be plastic. Things can be moved and actually change. But then you get entrenchment of norms, of laws, of interests and that makes change much harder to occur. And this, I think, is very relevant, because actually I think that maybe the world as a whole is in this period of plasticity.

But it's also relevant for what we choose to focus on, where — Bill McKibben, a leading environmentalist, talks about climate change. He says things were much more plastic decades ago, 50 years ago, when we really — we had still a reasonable understanding of the science then. We could have taken more action. We could have made a bigger difference with less effort then. And change is much harder now that things are much more politicized.

When I look to new issues that are developing, such as the rise of powerful AI systems, such as advances in biotechnology, we have this unusually good opportunity to take action to make sure that things go in a positive direction, where a decade after the development of technology, perhaps things have gotten politicized and it's much more difficult to make change.

Ezra Klein

How do you think about the tension between these two directions of devoting our moral attention? So one version is look for where there is plasticity, look for — we can do more right now on pandemics than we could do at other times in human history. People are aware of it. They're afraid of them. There's maybe some appetite, though not as much as I wish, for action. Biotech is another good example, A.I. safety.

On the other hand, we just had this whole conversation about something that probably looked very rigid, but was, in fact, not, which was slavery.

As you said, that had been the norm in human cultures for a very long time. And then all of a sudden, at least within the scale of human history, it wasn't.

So if you were looking at some of the other things you just mentioned, which are entrenched — the way we treat animals, the way we treat prisoners, the way we treat people outside of our own borders — I think a plasticity analysis would say that's probably too hard. And the sort of more values-oriented analysis you were offering would say history shows things can change very quickly. How do you think about that?

William Macaskill

It's — yeah, a great thing to push on. In the case of the abolitionist movement, I actually think that there was a moment of plasticity that was being taken advantage of. And this is the analysis by historian Christopher Leslie Brown at Columbia that I'm very influenced by.

He argues that, in particular, American War of Independence created this moment of plasticity for how Britain and the British Empire conceptualized itself, where essentially he argues that this was just an enormous blow to the British self-image and the British project. And so they used abolition of slavery as a way of regaining what he calls moral capital, regaining moral status, and giving a way to look down at the people in the US who were still slave owners.

So there is an argument there. I also think that the burgeoning Industrial Revolution, and, in particular, the much greater — the ease of spreading information was, again, a kind of moment of plasticity, where the abolitionists, they were the first social campaign as we really think about it. They were the first people to engage in mass petitions.

And I think they were — had the first strike, as it were, where they could engage in this mass mobilization campaign, doing petitions, taking them to Parliament. And the kind of natural opponents, like the plantation lobby, didn't really know what they were getting hit with. But then people learn over time. So this new social tactic or even social technology, that they could then — now you have misinformation coming from the people that activists campaign for.

So in this case, I actually do think it, at least plausibly, occurred at this moment of plasticity. With these other things, you're absolutely right, that things are quite entrenched. Nonetheless, moments of plasticity can occur. And I think the correct plan, if you are an activist, is that you build up these general purpose resources and then, at key moments, then you can actually put them into action.

And so I think if you study the history of the neoliberal movement, this is what happened. The period of the '50s and the '60s was really spent just growing the power of neoliberals, those who believe in free market economics, such that in the '70s, when stagflation happened, then they just had this enormous portfolio of — they could say, basically, I told you so, push forward their new economic agenda, and, for better or worse, that was enormously influential.

And so when we're looking at animal well-being, if we're looking at prisoners, if we're looking at people in other nations, then looking for these moments of plasticity and prepping in advance I think is very important. And just taking the animal case, I think the introduction of in vitro meat — that is, meat grown in labs — is just going to be one of these moments of plasticity where, for example, can you call it meat, legally? That is something that existing meat producers are not going to want you to be able to do, but could be very important for how widespread adoption is.

Ezra Klein

In vitro meat, if it becomes a widespread reality, which I very much hope it will, is an interesting alternative way to think about this question, because we've been talking about values and we've been talking about activists. And we've primarily used the example of slavery abolition here. But what about technology? And what about technologists? New technologies can create plasticity that didn't exist or create opportunities that didn't exist before. What is the longtermist approach to being a technologist?

William Macaskill

So yes, exactly. There are two things. One is that new technologies gives this enormous moment of plasticity, because you need to regulate and have norms around the technology. So what should the regulations and norms of an artificial intelligence be as it gets more and more powerful? What should the norms and regulations be around being able to make new pathogens in labs, given technology like CRISPR-Cas9 and other gene editing technology.

So one thing that is at least distinctive about the community of people that I would call a longtermist community is that there's a lot of horizon scanning for — OK, what are the areas of technology that are making very rapid progress — where AI and biotech are the two most obvious ones — that really could be transformative, even if the most powerful technology is not yet here, but will be in five, 10 years, and then in advance trying to prepare.

Thinking through, OK, what does a good regulatory environment look like? What are the social norms we should have about this? And then finally, what are other technology that maybe we should be trying to bring forward in order to protect against some of the risks?

So this is a concept called differential technological progress, where some sorts of technology, like vaccines, are defensive — or masks are defensive against pathogens, or this far-UVC lighting that I talk about, that could make a room safe from airborne pathogens. But others, such as the ability to make viruses more deadly or more transmissible, or even to make viruses essentially de novo, that is at least dual use and has the ability to create new pathogens of unprecedented, disruptive power.

And so, yeah, the two attitudes I think would be think about regulation and social norms well ahead of time, and then secondly, wherever possible, try and advance safer technology, defensive technology. And if it's possible, though often that's extremely hard is, yes, slow or minimize the development of more technology that at least poses very significant risks.

Ezra Klein

It's been my observation that it is more common for people to have a thought through agenda for values change, for political change, for the programs they would like to see emerge, the transfers and subsidies they would like to see happen, than for technological change. So I'm curious, from your longtermist, moral philosopher perspective, what are the technologies you would like to see focused on?

William Macaskill

The two technologies, by far, are artificial intelligence, where things are moving very quickly. I don't know yet, but if anyone listening, if you just check out some of the latest models, they're incredibly impressive. So one model, for example — it's a language model. You can give it a puzzle, like my grandmother was in hospital and I felt a certain color. Then I got caught up in traffic, and I felt another color.

If you take those two colors and put them together, merge them, what color would you get? And the language model will say, well, if your grandmother was in hospital, you would feel a sad color, like blue. And if you got caught up in traffic, you would feel an angry color, like, red. When you combine blue and red, you get purple. So the answer is purple. So we're starting to see these models that are engaging in something that looks a lot like quite impressive, in fact, human reasoning.

Obviously, it's still at a fairly early stage, but the progress is dramatic. They're also able to do proofs at a reasonably high level. They're also showing evidence of generality. So not just one model data was not just able to do a single task, but able to do very many tasks. And this is going to have just an enormous array of implications, some of varying degrees of familiarity. They could be very potent at entrenching power among elites.

But then it also means that we might get to the stage where we develop artificial systems that are as powerful, or much, much more powerful than human beings, and more capable across a wide variety of domains, where if you survey experts on this, they are thinking, like, this is going to occur with really quite significant probability over the coming few decades.

That's something we should be really prepared for, in particular because on leading economic models, it really follows quite naturally that you should expect much faster rates of technological progress as a result. Basically, once you get to the point where A.I. systems can automate innovation, including innovation about A.I. systems themselves, then things could start really happening very quickly. And so what we might think of as things that would be designed only after centuries and centuries of progress we could actually get within decades.

At the limit, then, we could even lose control of A.I. systems to the A.I. systems themselves, the argument being that, well, the reason humans are so powerful on the planet, so ecologically dominant, is because of our collective intelligence. And that's something that could be surpassed by A.I. systems themselves.

Now, these are just gigantic problems and gigantic issues that we're facing.

And that requires foresight, both in terms of the technical side — so ensuring that A.I. systems do what we want them to do, ensuring that even if we did have A.I. systems that were far more intellectually capable than us, nonetheless, we would still kind of be in control. And then secondly, having foresight about regulation here, where how do we ensure that these technologies don't get used to entrench power. How do we ensure that we're safely navigating what might be one of the most important technological transitions in history?

Ezra Klein

I think somebody listening to this would note that various technologies have come up a number of times here, A.I. in particular, but also biotechnologies, nuclear weapons and primarily negatively. That we have invented or we are inventing a bunch of technologies that you seem to think pose a quite profound risk to our future.

And a reasonable question might be is the longtermist view that we should stop? That we don't understand the stuff well enough, we shouldn't be putting our long-term future at risk. When there is uncertainty, we should err on the side of caution. My sense is that's not your view. That you're actually quite worried about technological stagnation. But talk through that tension here, because technology comes up a lot for you, but not very positively.

William Macaskill

Sure, I mean, the thing I should say is like, if you ask me in general what do I think about technology, then I'm enormously pro. I mean, if we look to the last 300 years, people's quality of life has increased enormously. I don't need to suffer pain during surgery. I don't die of syphilis, or tuberculosis, or typhus. I'm able to travel quite widely. There's just enormous benefits from technology. And I expect in the future there to be enormous benefits again.

But you're right. Technology is a double-edged sword. Look at the power to split the atom. It gave us both nuclear power, which is enormously beneficial, and the bomb, which is a very scary development. So I'm quite sympathetic to the idea that we do want to delay the dangerous technology, or at least slow the point at which we get it, until we are adequately prepared.

And the big worry is, is this feasible at all? Where if the most responsible countries say, oh, we're not going to do certain technology, and then the irresponsible ones race ahead, you might have made things worse. And the idea of globally slowing the development of a particular technology, that could be very hard. However, it does happen.

We have the technological power to clone people, for example. We've already cloned monkeys. We could clone human beings. That has not happened. And I don't actually expect it to happen anytime soon. That's not a technological hurdle. That's a matter of global social norms. So for some technologies, I think we could slow them.

If we're slowing growth in general, I do start to get more hesitant. So in chapter seven of "What We Owe The Future," I talk about stagnation, the idea that maybe growth wouldn't just slow, but actually even come to a halt. And there is reason for thinking that that might well happen within something like 100 years.

And I actually think that would be quite a scary thing, because I think that given the level of technological progress that we have at the moment is not sustainable. So if we think about — imagine if we stagnated at 1920s level of technology. The only way we could have supported society is by burning coal and other fossil fuels. We didn't have any alternatives.

So eventually, we would have just ran out of fossil fuels. We wouldn't have been able to continue. And we would have an utter climate catastrophe, worst possible case climate catastrophe on our hands.

In the future, if we have dangerous technologies — I mean, we already have them now with nuclear weapons — but we will quite potentially be able to develop extremely powerful bioweapons. Well, then we'll be at this period of heightened risk.

And if we're able to move through that period and get out on the other side, hopefully safely, then we've navigated this unsustainable state. Whereas, if we stagnate and we have a technology that could destroy us all, then over the long run, it becomes very, very likely indeed that we will, in fact, destroy everyone.

And so that's why we need to take this nuanced approach to technology and growth, where in the long run it's just very important that we maintain the engine of technological progress. But the most risky technologies, if we can push them into the future a little bit, at least, and ensure that we're appropriately handling them and that we'll handle them in a wiser way, then that would seem pretty good.

Ezra Klein

How do you think about biodiversity loss and the broader extinction crisis from a longtermist perspective?

William Macaskill

Yeah, this is something where I think there's a relatively — the longtermist perspective makes biodiversity loss much greater in importance. It's something that's irrevocable. So longtermists typically talk about extinction of humanity as it's lost forever, because if humans go extinct, we can't come back. And that's true for other sorts of extinctions, too. So I wrote the book during the pandemic, and I had these kind of — I sometimes went down very long rabbit holes, one of which was about the Quaternary megafauna extinction, the fact that human beings seemed to cause the extinction of a very large number of megafauna about 10,000 years ago.

And we would have had a much wider diversity of incredible animals — Glyptodonts, which were these car-sized armadillos, or Megatherium, which is this ground sloth that weighed like four tons.

We could have had them roaming the planet, but they went extinct. And my reading of the evidence is that humans were the most important culprit. And so the loss of some species today is not just the loss from the next few generations, but that's actually something that all people in the future would not be able to benefit from or value. And so I think it does just increase the value of this quite significantly.

Ezra Klein

If you're wrong about something big, either in the long-term framework or just in your moral philosophy, what do you think it's most likely to be?

William Macaskill

The thing that I still feel most confused about, or would feel least surprised about, having completely changed my mind in 10 years, is probably A.I., where it's just a particularly complex issue. So if you take the risk of biotechnology — in a sense, it's quite simple. Viruses kill people. We might be able to develop the power to create viruses that could kill far more people. That's really bad. That's really pretty simple. I feel like I know what's going on there.

Whereas with AI, it's more complex, not only in questions of when do we develop A.I., but also just at the point in time that we have systems that are as powerful as human beings or even greater. What does that society even look like? Because in that society, we've had enormous A.I. advances in the past.

And then secondly, how exactly do we want to navigate that? Because it's not as simple as, oh, we've just — there's one framing, which is, like, oh, we've got to lock in human power. That would also kind of seem bad. What I would want ideally is that A.I. could be used to further moral progress, that you can have these A.I. philosophers and they help contribute to the conversation.

And then also, there's just so many ways it can go. I worry about A.I. being helpful to individuals or countries or even companies with authoritarian impulses. I think the set of actions that you might do if you're particularly worried about that is different than the set of actions you might do if you're solely focused on A.I. takeover scenarios.

And the list could keep going on. Questions like, should we speed or slow A.I. progress? It's actually just extremely difficult. And so yeah, if you told me in 10 years, I've got this radically different perspective, I wouldn't really feel that surprised.

Ezra Klein

Let me end here on the question of what somebody does with all of this. The future is big — 80 trillion people, try to reduce great power conflict, or change A.I. — you know, most people are just doing their best. What does it imply for just somebody listening, who maybe has a job or doesn't have their choice of all jobs in the world, or has kids and is busy. What does having a longtermist perspective mean for living a normal life?

William Macaskill

So having a longtermist perspective, I think it's very natural to think, wow, these are just huge issues. They are above my pay grade. Maybe governments should be tackling them, but that doesn't really impact what I, as an individual, can do. And I think that would be very natural, and I think it's incorrect. I actually think anyone listening to this podcast can make an enormous difference, and an enormous difference not just on the present generation, but for all generations to come, for our grandkids and their grandkids. And I think that's in two main ways.

So one way is by donating. So if there's someone who just — they want to stay in their career. They want to contribute, well, you can donate to highly effective non-profits. You can find some examples at the EA — the effect of altruism funds, including ones that are targeted on benefiting the long-term future, within areas that include A.I. and biotech as well as climate change charities. And so one thing I particularly recommend is the Giving What We Can Pledge. This is an organization I co-founded that encourages people to give at least 10 percent of their income to the causes they believe will do the most good. And I know that you, Ezra, recently took the 10 percent pledge. And so I'm very happy and thankful that you've done this.

Ezra Klein

I am a signatory.

William Macaskill

Yeah.

Ezra Klein

I always think of it too — I think it's a funny — not a rebranding, because I think it's great. And as you said, I'm part of it. But this goes way back. I used to do this long before it was called Giving What We Can. It was tithing, right? And it's a deep tradition in many religions.

William Macaskill

Absolutely, there's a reason we chose 10 percent as the figure to aim for when making your donations, which is that if people say, oh, that seems a lot, seems like too much, well, one thought is, look, people are much poorer than us have been doing this for thousands of years. And so yeah, I think maybe if you're listening to this, you don't believe in God, you're not religious, well, that doesn't mean you shouldn't still be trying to live a fully ethical life. And I would love it to be a norm where donating 10 percent of your income to effective non-profits, that's just part of living a good life.

Ezra Klein

And now, always our final question, what are three books you would recommend to the audience?

William Macaskill

So three books I'd recommend — the first is "Moral Capital" by Christopher Leslie Brown, which I mentioned earlier. It's just the single best book that I know of on interpretation of the abolitionist campaign for the end of slavery. And he argues quite forcefully in favor of the argument that it really was an event that could have gone either way. Second is "The Precipice" by my colleague Toby Ord. I certainly couldn't not mention it. I really see "What We Owe The Future" as like a complement, like a sister book to "The Precipice," which details in incredible detail the existential risks that we face, including from bio and A.I., and great power war.

And then the final book I'd recommend is "The Scout Mindset" by Julia Galef, which is just a wonderful book for expressing a mode of reasoning that really resonates with me, which is rather than a soldier mindset, where you have a particular point of view, when you're trying to marshal all of your arguments in defense of it, instead you're just trying to explore. And you're trying to figure out, what does the landscape look like? What is true? And be cautious and questioning about your own point of view. And that's — I don't know if I always achieve that, but it's something I try to aspire towards.

Ezra Klein

Will MacAskill, thank you very much.

William Macaskill

Thank you so much.

[MUSIC]

Ezra Klein

So that's my conversation with Will MacAskill. I'm going to start trying to use the outros here to offer a few recommendations for where you can follow up, or hear challenges, or go deeper on the ideas in these episodes. I want to try and do a better job of connecting the various explorations we do on the show to one another. So A.I. risk came up a lot in this conversation. I decided not to go too deep into it here. But if you do want to dig deeper into it, you should check out my conversation with Brian Christian, author of the really great book on this subject, "The Alignment Problem."

If you want to hear more about effective altruism and why this moment might be an unusual hinge in human history, my conversation with Holden Karnofsky is a great place to start. And if you want an alternative perspective, particularly on the relationship human beings should have to the natural world, the novelist Richard Powers is really worth listening to. I loved that conversation, and it is one worth holding in tension with these others. We're going to put links to all of these in the show description. And as always, you can support the show by sending this episode to a friend, or if you hated it, an enemy, or leaving us a review in whatever podcast app you're using right now. "The Ezra Klein Show" is produced by Annie Galvin and Rogé Karma. Fact-checking by Michelle Harris, Mary Marge Locker and Kate Sinclair. Original music by Isaac Jones, mixing by Sonia Herrero and Isaac Jones. Audience strategy by Shannon Busta. Special thanks to Kristin Lin and Kristina Samulewski.

Three Sentences That Could Change the World — and Your Life



The Ezra Klein Show



The philosopher William MacAskill lays out the case for longtermism: "Future people count. There could be a lot of them. And we can make their lives better."

0:07 / 1:11:13



More episodes of **The Ezra Klein Show**



Transcript

August 9, 2022 • 1:10:57
Three Sentences That Could Change the World — and Your Life



August 5, 2022 • 1:18:27
Gender Is Complicated for All of Us. Let's Talk About It.



Aug. 9, 2022

Produced by 'The Ezra Klein Show'

Today's show is built around three simple sentences: "Future people count. There could be a lot of them. And we can make their lives better." Those sentences form the foundation of an ethical framework known as "longtermism." They might sound obvious, but to take them seriously is a truly radical endeavor — one with the power to change the world and even your life.

That second sentence is where things start to get wild. It's possible that there could be tens of trillions of future people, that future people could outnumber current people by a ratio of something like a million to one. And if that's the case, then suddenly most of the things we spend most of our time arguing about shrink in importance compared with the things that will affect humanity's long-term future.

[You can listen to this episode of "The Ezra Klein Show" on Apple, Spotify, Amazon Music, Google or wherever you get your podcasts.]

William MacAskill is a professor of philosophy at Oxford University, the director of the Forethought Foundation for Global Priorities Research and the author of the forthcoming book, "What We Owe the Future," which is the best distillation of the longtermist worldview I've read. So this is a conversation about what it means to take the moral weight of the future seriously and the way that everything — from our political priorities to career choices to definitions of heroism — changes when you do.

We also cover the host of questions that longtermism raises: How should we weigh the concerns of future generations against those of living people? What are we doing today that future generations will view in the same way we look back on moral atrocities like slavery? Who are the "moral weirdos" of our time we should be paying more attention to? And what are the areas we should focus on, the policies we should push, the careers we should choose if we want to guarantee a better future for our posterity?

You can listen to our whole conversation by following "The Ezra Klein Show" on Apple, Spotify, Google or wherever you get your podcasts. View a list of book recommendations from our guests here.

(A full transcript of the episode is available here.)



Matt Crocket

"The Ezra Klein Show" is produced by Annie Galvin and Rogé Karma; fact-checking by Michelle Harris, Mary Marge Locker and Kate Sinclair; original music by Isaac Jones; mixing by Sonia Herrero and Isaac Jones; audience strategy by Shannon Busta. Special thanks to Kristin Lin and Kristina Samulewski.