Does longtermism depend on questionable forms of aggregation?

Moreno, Marina

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

gesis
Leibniz-Institut
für Sozialwissenschaften

Mitglied der

Leibniz-Gemeinschaft

Does longtermism depend on questionable forms of aggregation?

Marina Moreno

We are constantly making choices about how to invest our time and resources. From a moral perspective, we must ask which moral concerns are most deserving of our attention. Longtermism, as e.g. defined by Greaves and MacAskill, holds that our moral focus should be on the long-term future, and that current and medium-term moral problems are comparatively insignificant. This theory is centrally based on the assumption that the moral importance of individuals can be aggregated. Since the number of individuals of future generations far exceeds the number of current individuals and those closer in time, future generations are to be morally prioritised, according to longtermism. This paper explores the implications of rejecting the premise of moral aggregation of individuals and adopting a strongly non-aggregationist position instead. It is argued that, according to strong non-aggregationism, the magnitude of the probability with which our intervention actually make a difference, as well as whether we look at the available interventions from an ex ante or ex post perspective, are relevant factors in their moral assessment. Ultimately, the conclusion is reached that strong non-aggregationism does likely not support strongly longtermist conclusions.

Keywords: *Longtermism; aggregation; extinction risks; unaligned AI*

Introduction

Ever since humanity has gained the technological potential to destroy itself, extinction risks have increasingly been the focus of political and philosophical debates which have in turn sparked efforts to minimise these risks. One important line of argument supporting this focus is suggested by recent literature, according to which it is one of the most important moral imperatives to let our actions be guided by a long-term perspective. The basic argument for this claim runs as follows: Humanity has the potential to go on existing for a very long time before going extinct. During this time, a very large number of morally considerable individuals (humans, animals, and potentially digital minds) could come into existence. Since these individuals are just as morally important as similar individuals living today, their aggregated moral importance outweighs the moral importance of present generations by far. Thus, ensuring the existence and influencing the welfare of the large number of generations in the far future is among the

> Longtermism holds that future individuals are just as morally important as similar individuals living today, therefore their aggregated moral importance outweighs the moral importance of present generations by far. Thus, ensuring the existence and influencing the welfare of the large number of generations in the far future is among the most important moral goals, if not the singular most important one.

most important moral goals, if not the singular most important one.

The above argument relies on a crucial moral assumption; namely that welfare can be aggregated across individuals to form better goods and worse bads. The thesis that moral importance is relative to the number of individuals affected lies at the heart of aggregationism. With his famous paper *Should the numbers count?* in 1979, John Taurek challenged this thesis and sparked a debate that is still on-going today (see Taurek 1977). For instance, in her recent monograph *Non-Aggregationismus: Grundlagen der Allokationsethik*, Weyma Lübbe defends a rigorous non-aggregationist position (see Lübbe 2015) – and various other authors have proposed alternative ways to morally count or not count the number of people affected by a decision as well.

In this paper, I aim to explore the challenge a strongly non-aggregationist position poses for longtermism and its implications, such as a strong moral focus on extinction risks, and point to the relevant research questions in this area. To this end, I will first locate and contrast the version of longtermism I am concerned with and show how it implies a certain anti-presentism in section 2. In section 3, I will then motivate non-aggregationism and defend it against an obvious objection. Subsequently, I will move on to section 4, where I consider what non-aggregationism implies regarding the longer-term future and large-scale decisions, such as where to invest resources and which career path to choose. At this point I will show that non-aggregationism seems to suggest a long-term focus as well. However, in section 5 I will raise and discuss complications regarding the probabilities with which individuals today and in the future can be affected. These complications, I will ultimately argue, suggest that non-aggregationism does not in fact recommend a long-term focus after all. Rather, it points to a near- to medium-term focus instead. Section 6 concludes with a summary and an outlook for further research into the issues discussed.

Aggregation, longtermism and (anti-)presentism

In *Reasons and Persons*, Derek Parfit asks us to consider and morally compare the following three scenarios: a) peace, b) a war that kills 99% of the human population and c) a war that kills 100% of the human population. While it stands to reason that a) is better than b) and b) is better than c), Parfit offers the striking thought that the difference in goodness between b) and c) is vastly larger than the difference between a) and b). This is due to the fact that c) implies the loss of all of humanity's possibly astronomically large future, while b) still allows for a full human recovery (Parfit 1984: 453-454). To the extent that the goodness of outcomes ought to crucially influence our moral decision-making, Parfit's reasoning supports a priority of the far off long-term over the short-term.

Based on such considerations, many longtermists have followed Parfit to further develop this argument. For instance, Nicholas Beckstead defends the claim that influencing the far future is morally *overwhelmingly* important, i.e., more important than most other moral goals, because the expected value of the far future is astronomically great (Beckstead 2013: 2, ch. 3). Similarly, Nick Bostrom suggests that existential risks reduction to avoid the above scenario c) ought to be our global priority and serve as a strongly action-guiding principle, more strongly so than other considerations exclusively affecting the well-being of present generations. That is, he argues that there is a moral case for existential risks reduction being more important than any other global public good, as well as that many existential risks will not arise right now, but rather in the foreseeable future, such that our global priority ought to be to build resilience against them. These two claims taken together imply that other policies which affect the well-being of present generations that might, for instance, starve within the next couple of years, are globally less important than existential risks reduction, whether this can be achieved in the short- or medium-term (Bostrom 2013: 1). A more comprehensive case for this kind of priority of the long-term over the near-term and medium-term has recently been developed by William MacAskill and Hillary Greaves (2021). They define the notion of the so-called "strong longtermism" which, in contrast to softer versions of longtermist thinking, holds that the long-term future is the *most important* moral feature of our present actions (Greaves and MacAskill 2021: 3-4, 26-27). Similarly, Toby Ord has recently argued that the prevention of extinction risks, whether presently relevant or likely to hit in the further future, ought to be one of, if not *the* most important of our global priorities (Ord 2020).

As I have outlined the arguments in the longtermism literature so far, longtermism seems to imply a certain anti-presentism, i.e., the view that the impact of our actions on present generations matters little in contrast to the impacts on the far future.[1] This anti-presentist implication could be resisted in at least three straightforward ways. Firstly, we might argue that even if their axiological claims regarding the importance of the far future, given its astronomical stakes, were true, this would not imply that our all-things-considered deontic duties must follow these axiological claims. That is, even if the consequences of our actions are actually best if we follow the maxim of seeking to optimise its long-term consequences, the goodness of those outcomes may not (exclusively) determine what we ought to do. Secondly, one might deny that there are in fact trade-offs between affecting the near-term and long-term future, i.e., that whatever benefits the present generations most will also benefit the far future most. While there would thus still be a conceptual priority of the far future and future generations, the priority would not matter much in practice. Lastly, one might reject the outlined arguments for longtermism altogether and champion a different version of longtermism, based upon considerations other than the astronomical aggregated stakes of humanity's (far) future. Neither one of these strategies for argueing against longtermism are the ones employed in this essay, but I will briefly address them now in turn in order to set up the context of my own debate.

Firstly, Greaves and MacAskill explicitly distinguish between axiological and deontic strong longtermism. Importantly, however, they defend both the axiological and deontic version of strong longtermism, i.e., they hold that the impact on the far future most importantly determines both the goodness of the respective outcomes as well as what we, all things considered, ought to do (Greaves and MacAskill 2021: 2).

Similarly, Bostrom explicitly asserts that the impacts existential risks have on the generations of the far future ought to be our central *action-guiding* principle, i.e., is deontically very relevant as well (Bostrom 2013: 1). Generally, central proponents of the version of strong longtermism that is based on the aggregative premise outlined above thus do indeed often make deontic claims. However, my arguments can be equally applied to both the axiological as well as the deontic versions of strong longtermism, since it will crucially concern itself with the underlying aggregative premise, which underlies both the axiological as well as the deontic version.

Furthermore, even if we grant that axiological strong longtermism does not directly translate into deontic strong longtermism, important questions on the deontic level are still likely to remain, to which my subsequent discussion will be relevant. Beckstead (2013: 9) and more recently MacAskill (2022) have left open the possibility that their aggregative axiology may not translate into deontic guidance directly. For instance, MacAskill allows for other important factors that should influence our decisions from a moral perspective, namely special relationships and reciprocity (MacAskill 2022: 8). However, consider the following possible scenario: A philanthropist has a certain sum of money that they would like to donate to a specific cause. They have already fully discharged both their duties regarding special relationships as well as reciprocity, such that these factors do not influence which cause the philanthropist is required to select to donate to. According to longtermists such as MacAskill, the deontic permissibility of their cause selection would then be entirely determined by how it influences the generations of the far future. Even this more limited deontic claim seems somewhat doubtful, however, given that many people (to whom the philanthropist has no special relation or reciprocal duties) alive today are in dire need of help as well.

This leads us to the second point above, namely whether the near- and long-term perspectives might not in fact simply converge on a more practically applied level. For instance, consider the example of nuclear disarmament. Ending the cold war was both important from a long-term perspective in order to mitigate the extinction risk posed by nuclear war, as well as beneficial for present generations on several levels; they no longer had to live in fear and a resulting peace dividend freed up resources which could in turn be spent on other humanitarian efforts.[2] To the extent that this kind of empirical convergence applies more broadly, the anti-presentism implied by strong longtermism would not actually make a relevant practical difference.

However, consider the following example inspired by Richard Pettigrew (n.d.: 2). Say, again, a philanthropist is choosing between two charities to which to donate.

On the one hand, they are considering donating to the cause of preventing malaria by net distributions. On the other hand, they are considering to fund researchers working on AI alignment.[3] From the perspective of future generations, it is mostly the latter that matters, since preventing malaria today is comparatively unimportant to them, given that it is unlikely to interfere with humanity's greater trajectory by which they will turn out to be influenced. AI alignment, however, is possibly very relevant to

future generations, insofar as it constitutes an extinction risk or a risk of very bad lock-in scenarios crucially shaping humanity's trajectory. Strong longtermists must therefore strongly advise the philanthropist to donate to AI alignment, given that the expected value of doing so will be larger by a significant amount in view of the vast number of people in the future influenced by AI but not by malaria. However, from the perspective of present generations, the case is not as clear at all. While it is nearly certain that malaria will affect people existing today within a short time period, risks resulting from misaligned AI may very well not affect people alive today at all, but certainly not the people who will die from malaria within, say, the next year if the necessary nets are not distributed. The priorities of present and future generations are thus in a trade-off here, simply as a matter of empirical reality. The same basic argument applies, mutatis mutandis, to time spent working on various causes as well as setting global priorities and developing respective policies. In general, it would seem surprising if the actions which have the best consequences for generations many thousands of years from now would converge exactly with the actions that have the best consequences for present generations. A certain anti-presentism thus does not seem to be avoidable within the framework of strong longtermism as outlined and defended by Greaves and MacAskill.

This leads to our third point, i.e., to considering other forms of longtermism which may be able to resist the anti-presentist implication. Note that none of these theories can be based as strongly on the aggregative premise outlined above, since this would lead us back to some form of anti-presentism. However, the general idea of introducing a long-term perspective has found its uptake in various frameworks, which I do not explicitly address in the present text. For instance, Roman Krznaric has suggested the notion of "long-term thinking", which refers to a type thinking which is aimed at mitigating the myopic short-term thinking that is pervasive in today's society (Krznaric 2020). It stands in sharp contrast to strong longtermism as outlined above insofar as it is not a fully comprehensive normative theory of evaluation but rather a guide to better incorporate the future perspective into our present thinking. The fact that Krznaric proposes a horizon of consideration of 100 years (Krznaric 2020: 14) further showcases the radical differences to strong longtermism in its normative underpinnings, where the relevant horizon includes up to several thousands if not billions of years (Greaves and MacAskill 2021: 6). In contrast to strong longtermism, Krznaric's long-term thinking can almost be viewed as presentist, i.e., biased towards the next 100 years.

Another broadly longtermist approach is concerned with the representation of future generations within the present political institutions. For instance, Jörg Tremmel (2021) suggests a four-branches model of government, where the three traditional branches are supplemented by an office for future generations. Tremmel states explicitly, however, that this office is not to have any veto power since that could significantly slow down the political process, but rather is intended to just propose policies benefitting future generations (Tremmel 2021: 18). Another political approach to the inclusion of future generations has been proposed by Dennis Thompson, who argues that democracies tend towards a harmful presentist bias which can be mitigated by the principle that "present generations should act to protect the democratic process itself over time" (Thompson 2010: 14). That is, present generations ought to act as trustees of the democratic process. Thompson likewise explicitly rejects theories which sacrifice the well-being of present generations in favour of some temporally distant collective good (Thompson 2010: 1).

Insofar as it can be expected of these political approaches that their practical implementation would still give a certain priority to the present generations, i.e., not weigh the interests of each person of the vast future generations exactly equally in an aggregative way, which would again lead to swamping out the interests of the present generations, they thus avoid the anti-presentist implications. I will, however, examine this latter strong longtermism for at least two pertinent reasons. Firstly, I am crucially considering how the implications of a non-aggregative normative theory differ from the explicitly aggregative theory which leads to the strong form of longtermism. Hence, the natural argumentative opponent of my arguments is the form of longtermism based on aggregative premises. Secondly, as Beckstead suggests, the strong form of longtermism would, if true, have the most radical and controversial implications. Greaves and MacAskill have recently argued that if longtermism were true, it would imply a strong focus on decreasing extinction risks and on influencing the path of high and long-term impact developments – such as the development of a superintelligent AI – (Greaves and MacAskill 2021: 13-14), since those are among the most promising interventions influencing the long-term future. By contrast, they argue that focusing e.g. on fighting malaria or factory farming today is of little moral importance, since the number of individuals affected by any of these is lower by several magnitudes. These implications are so counterintuitive that Beckstead devotes an entire section to make the claim plausible that most of what we believe about moral priorities might be wrong (Beckstead 2013: 25-53). Given the possibly radically revisionary nature of this theory, I believe it to be worthwhile to engage with it at length.

> Greaves and MacAskill have recently argued that if longtermism were true, it would imply a strong focus on decreasing extinction risks and on influencing the path of high and long-term impact developments, such as the development of a superintelligent AI.

However, since such a rigorous anti-presentism is initially counterintuitive, it is frequently met with scepticism. Can a moral theory that implies anti-presentism, i.e., that we ought to essentially abandon efforts to help any currently existing individuals really be considered plausible? Is a very strong focus on long-term developments such as extinction risks really reasonable rather than fanatical, as longtermism is sometimes accused of being (Greaves and MacAskill 2021:24)? Longtermists may answer these questions by arguing that the mere fact that some conclusions are counterintuitive should not be sufficient to undermine a theory, at least not to the extent that we actually agree with its premises. If the argument sketched above proves sound upon reflection, philosophy might play a legitimate revisionary role here.

However, the fact that these conclusions do seem so counterintuitive ought to give us at least good reason to examine the premises upon which they rest. This paper critically discusses the aggregative premise and explores the implications its rejection has for the conclusions of strong longtermism. In what follows, I will refer to the position of strong longtermism based upon an aggregative

premise and as defended e.g. by Greaves and MacAskill as well as Beckstead and Bostrom, as "longtermism".

Strong non-aggregationism: the rationale

In what follows, I will first outline the rationale and arguments motivating strong non-aggregationism. The literature on these issues has grown over the past years and I will only attempt to outline the specific view that I will be working with in this paper, which is inspired by Taurek, Lübbe and Mannino.

To illustrate his view, Taurek asks us to consider the following interpersonal rescue conflict: Six people are in need of medication to survive and you have a certain supply of said medication. Five of those six only need one fifth of the amount available to survive, while the sixth, call him David, needs all of it, if he is to survive. Is there a moral duty to give the medication to the five people and letting David thereby die? And relatedly, is it better in any relevant sense to save the five and let David die? To answer questions of the latter sort, Taurek argues one ought to ask: Better for whom? (Taurek 1977: 299) If the five people receive the medication, that is better for each of the five people and worse for David. If David receives the medication, it is better for David and worse for each of the five other people. However, Taurek argues that there is no one for whom it is better if five people are being saved instead of only one. There is no one perspective from which the fact that the greater number is saved is better.

> Taurek's thought experiment: Six people are in need of medication to survive and you have a certain supply of said medication. Five of those six only need one fifth of the amount available to survive, while the sixth, call him David, needs all of it, if he is to survive. Is it better in any relevant sense to save the five and let David die?

Taurek goes on to ask what it would even mean to evaluate the death of five people as worse than David's death (Taurek 1977: 299-300). What do evaluative judgements such as better or worse really mean if there is no person for whom they are better or worse? It is precisely the fact that something can be bad from someone's own perspective that makes them ethically relevant individuals in the first place. However, there is no meaningful level of evaluation above and beyond this personal perspective. According to Taurek, this point extends to all numbers cases, i.e., that the number of affected people never morally counts. For instance, even if, say, 50 family members were to be suffering if the five people die (10 each), and only 10 in case David dies, the same argument applies again: Since there is no meaningful aggregated perspective from which it is worse that more people suffer (i.e., 50 instead of 10), there is no corresponding reason to prioritise the greater number of separate people regarding the family.[4] Note that this example also shows that the lack of an impersonal perspective Taurek alludes to here is not an implication of the fact that the people *die* if they are not saved, i.e., the problem is not that their perspectives vanish once they die. The example would still hold if we consider the grieving family or if we modified the example such that each of the six people get a severe (but not deadly) illness without the medication. While receiving the medication is better for each of them from their own personal perspective, there is no one perspective from which five prevented illnesses are better than one prevented illness.

Contrast the above case with a rescue situation in which there is a unique pareto-optimal solution. Suppose you could either save the five people or save the very same five people and David. Here it is clear that the best option is saving everyone, because that is better from everyone's perspective. However, in a genuine conflict situation it seems unclear what it would even mean that one outcome is worse, period, than the other. According to defenders of this "no worse claim", there is thus no reason to save the greater number based on any notion of realising better world states (Taurek 1977: 300). In this sense, Taurek's non-aggregationism undermines both axiological as well as deontic versions of how numbers may morally count.

This argument bears a similarity to Rawls' critique of utilitarianism, according to which utilitarianism does not take the separateness of persons sufficiently seriously (Rawls 1971: 23, 163). One way to interpret this idea is exactly what motivates non-aggregationism: Since people's perspectives are separate, and do not in this sense aggregate into one superbeing whose well-being consists of all the separate individuals, there cannot be an equivalent aggregate moral value. Call this view regarding the strict non-aggregation across persons strong non-aggregationism.

However, strong non-aggregationism ought to offer a different solution to interpersonal rescue conflicts in order to constitute a genuine alternative to aggregationism. How exactly this solution is spelled out differs for various theories. What most of these suggestions have in common is the following general thought: In interpersonal rescue conflicts, all of the people affected equally hold a very strong claim to being helped. As such, all claims involved ought to be respected. Since the situation is such that we cannot actually fully satisfy all of the claims jointly, we have to instead find another way to express our equal concern for all the people involved. Now if we simply aggregated and opted to save the greater number, the people in the majority group would receive all of our concern and the people in the minority group would receive none. Instead, many strong non-aggregationists embrace a lottery solution: Since the good of survival cannot be given to everyone who has a claim on it, one ought to instead at least distribute the chances of survival equally among everyone.

There again exist different versions of how the chances are distributed exactly. Due to the scope of this paper, I cannot go into detail about all the advantages and disadvantages each of those versions offer. Rather, I will present and further discuss one particular version defended by Lübbe (2015). The reason I chose this version is twofold: Firstly, I personally believe it to best capture the spirit of equal and maximal concern, and secondly, because I believe it poses the most interesting challenge to longtermism, which is what I aim to explore in the present paper.

Consider again Taurek's case of five people versus one person. Lübbe holds that each of the people involved in this rescue conflict ought to be respected equally and maximally. Hence, she argues, the maximal chance of survival to be allocated that is consistent with equality is a 50-50 lottery between the five people and David. This way, each individual person involved receives a chance of 50% of survival, which is as good as it can get for everyone without decreasing anyone else's chance for survival.

Even though not all of the claims involved can thus be respected to the degree that they will eventually be satisfied, as this is unfortunately simply the empirical reality we find ourselves in, every claim is still equally respected to the maximal possible degree.

Such proposals indeed seem to be in line with many societal practices in place today. In many jurisdictions, for instance, it is illegal to deprioritise patients with an increased need of a certain scarce medical resource on the grounds that doing so would likely save a greater number of people. For instance, the Organ Procurement and Transplantation Network in the United States functionally assigns a higher chance to be saved to people who need two organs than to people who only need one organ (OPTN 2020: 137). Similarly, the policies from Eurotransplant do not opt to save the greater number either, but rather specifically assign an equivalent chance to multi-organ patients as to single-organ patients (Eurotransplant 2019: 10, 23, 36). That is, everyone's claims are being respected by granting everyone access to medical resources by virtue of an organ waiting list, i.e., a fair queue. The fair queue arguably implements a type of natural lottery, where the random process of admission to the queue distributes the respective chances equally from an ex ante perspective (see John and Millum 2020). The triage law recently developed and passed in Germany likewise does not ensure that the greater number is saved by e.g. deprioritising patients who will need scarce medical resources such as ventilators for a longer period of time than others. If e.g. two patients each need the last ventilator for a week to survive and a third patient needs it for two weeks, the triage law will not make the former two patients a priority, but will give an equal chance to all three of them by equally admitting to the fair queue (see Deutscher Bundestag 2022).

> Consider again Taurek's case of five people versus one person. Lübbe holds that each of the people involved in this rescue conflict ought to be respected equally and maximally. Hence, she argues, the maximal chance of survival to be allocated that is consistent with equality is a 50–50 lottery between the five people and David. This way, each individual person involved receives a chance of 50% of survival, which is as good as it can get for everyone without decreasing anyone else's chance for survival.

Quantitative Catastrophes and Unequal Stakes

In their paper, Greaves and MacAskill (2021) briefly consider a non-aggregationist view according to which an aggregate good or aggregate bad may not be considered sensible notions, conceding that such a view might challenge their longtermist conclusions.
They reject this line of thought by the following argument: Any minimally plausible theory must be able to explain the fact that special norms seem to apply to situations where there is a huge amount at stake. For instance, if millions of people are threatened by a natural catastrophe, other comparatively minor moral constraints get justifiedly overriden. Greaves and MacAskill (2021:27-28) hold that any attempt at explaining this fact from a non-aggregationist perspective will lead back to supporting their own conclusions.
Their argument can be reformulated to pose a challenge to non-aggregationism more straightforwardly: Is it not simply entirely implausible to not believe that a large-scale natural catastrophe affecting and possibly killing billions of people is not any worse than a small-scale natural event killing one person would be? And relatedly: Is it not obvious that in catastrophic circumstances, saving as many people as possible is the right thing to do?

Indeed, emergencies are often thought to be a context in which a thoroughly consequentialist aggregative ethics is most adequate, given the huge number of people affected (e.g. Rakić 2018). While one may be inclined to follow Taurek in cases where the number differences are small, such as one versus five people, the bigger the number differences become, the harder it is to swallow Taurek's claim. While our intuitions in the case of catastrophic circumstances may still be able to be accounted for by the particular badness catastrophic circumstances may come with (which I will come back to below), even just in comparing, say, 10,000 car accidents killing one person each versus one other car accident killing one other person, it seems very counterintuitive not to claim that the former is clearly worse. If non-aggregationism is unable to account for these intuitions in any way, it may simply be implausible at the outset.

> Is it not simply entirely implausible not to believe that a large-scale natural catastrophe affecting and possibly killing billions of people is not any worse than a small-scale natural event killing one person? And relatedly: Is it not obvious that in catastrophic circumstances, saving as many people as possible is the right thing to do?

While these intuitions do indeed pose a great challenge to non-aggregationism, they do not completely undermine its plausibility. This is due to at least three features a plausible version of non-aggregationism ought to include.
Firstly, while non-aggregationists do not consider the numbers intrinsically morally relevant in cases of interpersonal conflicts, there are ways in which they do let the numbers count after all. In the case in which we could either save the five people or the five people and David, additionally saving David is the right thing to do even for a non-aggregationist, as it is the unique Pareto-optimum. A non-aggregationist can thus make sense of the intuition that additionally saving more people from a catastrophe is better than saving fewer – so long as saving these additional people is not in direct conflict with saving other people with equal claims.
Secondly, people who are in a sufficiently "veiled" ex ante position may reasonably agree on an aggregative *policy* without necessarily referring back to aggregative *moral premises*. (The notion of "veiling" employed here is Rawls's (1971), but the ex ante situation is supposed to be empirically real, not a hypothetical "original position".) Consider a government which ought to decide whether, as a general policy, they send help to a place where there are more people to be saved versus to a place where there are less people to be saved, if both of them are affected by the same devastating catastrophe. From an ex ante perspective, each person living under the government is generally more likely to end up in the place where there are more people, assuming that where you end up is sufficiently chancy, and uniformly so. Non-aggregationists may thus jointly commit to an aggregative emergency policy to the extent that it is beneficial to everyone from an ex ante perspective because this policy gives everyone the greatest chance to be among the ones who will be saved. Thus they adopt an aggregative policy not because saving more people is better in an impersonal sense, but because the policy benefits everyone separately.
For instance, the six people in Taurek's medication case might have precommitted to an aggregative policy in advance, i.e., be-

fore they knew how much of the medication they would need. With such a policy, all of them have a higher chance to end up being saved, since it is, *ceteris paribus*, likelier for everyone to end up in the bigger group of five people who only need a fifth of the medication. As such, this precommitment benefits each of the people separately, by virtue of giving each of them a higher ex ante survival chance, and it is thus not necessary to hold that saving five people is simply better from some objective, impersonal perspective. While this reply is not able to completely answer the objection raised by large moral catastrophes, it is able to weaken its strength, as it is able to account of some aggregative policies without referring back to aggregative premises, i.e., without giving up the central idea of Taurekianism.

Note, however, that the same argument is not analogously available in the context of future generations for at least two reasons. Firstly, present generations cannot agree to aggregative policies that include all future generations, since the situation is not sufficiently veiled in this case: People living today would know in advance that they will be among the deprioritised ones and at no point is such an aggregative policy in their interest. Secondly, it is hard to make sense of an ex ante consensus regarding people who have not been born yet. For an analogous argument to work, one would have to argue that for each person that has not been born yet, any time in the future is roughly equally likely to be their actual birth time. If this is even a sensible proposition, it is certainly at least metaphysically doubtful. An argument for aggregation based on a veiled ex ante situation does not seem very promising if the veiled situation vanishes as soon as the relevant individuals *come into existence*.

Alternatively, one may interpret the veiled ex ante situation as a "mere" operationalisation of impartiality (see Rawls 1971), and thus argue that it must apply to future generations as well. But this would beg the question against non-aggregationists, who hold that impartiality implies that equal rescue chance must be given to the individuals actually involved in a rescue conflict, even if saving the greater number of people would be preferable behind a hypothetical veil.

Finally, non-aggregationism also has resources to account for the particular badness such catastrophes often come with. While non-aggregationists do not consider the interpersonal quantitative dimension of a catastrophe, they should consider the *interpersonal comparability* of the quantitative dimension of a catastrophe. That is, for non-aggregationism to be plausible, it ought to consider the stakes of the affected individuals as a relevant criterion to determine the strength of a claim these individuals may hold. For instance, if the five people in the above example only need the medication to cure their broken hand while David needs all of it to survive, a 50-50 lottery seems hardly adequate. Instead, non-aggregationists would have reason to prioritise David. This prioritisation may take the form of a lottery weighted in his favour or choosing him directly. I will consider both of these types of prioritisation below. These cases with unequal stakes are rarely comprehensively addressed by non-aggregationists, but e.g. Mannino (2021) argues that a plausible theory of non-aggregationism allows for interpersonal comparability of this sort, and prioritises people according to their claim strengths, which are dependent on their stakes. One reason why non-aggregationists can include this type of interpersonal comparability is the fact that this type of comparison does not necessarily leave the person-relative framework: When comparing a headache of person A with torture-level suffering of person B, one can plausibly hold that B's suffering is worse *for* B than the headache is *for* A.

It is in this sense that non-aggregationists can make sense of the particular badness of many intuitively catastrophic states of affairs. Many intuitively catastrophic events come with devastating consequences for the affected people, such that their claim to be helped is particularly strong. Even though it is not the number of people affected that constitutes the catastrophe, non-aggregationists can still account for the fact that these catastrophes are very bad for the separate individuals and create a strong reason to help them. Vice versa, if the number of people affected by an event is very high but the individual stakes are sufficiently low, the intuition of it being a catastrophic event is usually way less strong. This seems to suggest that our understanding of catastrophes is not as far away from the non-aggregationist picture as may be thought initially.

While interpersonally quantitatively large catastrophes may thus still pose a challenge to non-aggregationism, I do not believe this objection to undermine it sufficiently such that it ought not to be considered a serious theoretical rival to aggregationism.

In this context, it is also worth noting that Heikkinen has recently argued that non-aggregationism poses a challenge for longtermism *even if* one accepts that interpersonal quantitative catastrophes make for a *worse* outcome in an axiological sense than e.g. smaller scale accidents. This is because accepting an aggregative axiology does not force us to adopt an aggregationist deontic stance according to which this axiological fact gives us reason to prioritise avoiding the axiologically worse outcome (Heikkinen 2022: 13-14). In fact, Greaves and MacAskill also note that many non-aggregative theories are mostly concerned with deontic claims, rather than axiological claims (Greaves and MacAskill 2021: 26-27). While they provide a separate deontic argument along the lines mentioned above, according to which very high stakes always ought to override deontic side constraints, Heikkinen shows convincingly that their argument is not sufficient to reject non-aggregationism in the deontic sense, even when we accept the corresponding axiological stance. While the non-aggregationist theory I am defending here includes the rejection of aggregation both in terms of value axiology *and* deontic obligations, Heikkinen's reasoning showcases even further how pervasive the challenge for longtermism posed by non-aggregationism seems to be. The arguments in the following sections can, mutatis mutandis, likewise be applied to non-aggregationist theories which are concerned with the deontic rather than the axiological questions. Before moving on to considering the implications non-aggregationism likely has for longtermism, I will briefly summarise the particular non-aggregationist picture I have outlined so far and will be working with in the following sections: Non-aggregationism holds that evaluative statements involving notions such as "better" or "worse" can only be made from personal perspectives. For this reason, saving the greater number of people cannot be said to be better from anyone's perspective, unless it is indeed better for everyone involved, and is thus a Pareto-optimal solution. In the case of an interpersonal conflict, however, everyone involved ought to be respected equally by giving everyone a maximal chance of having their claim satisfied that is consistent with everyone else receiving the same chance. This way, each and every person is given maximal and equal concern. However,

even though non-aggregationists do not prioritise people on the grounds that they are part of a majority group, they do consider the individual stakes at hand. That is, they prioritise people according to the strength of their claims, rather than according to their number. Hence, a non-aggregationist ought not to be on the lookout for the largest groups to save, but rather for the person with the strongest claim to satisfy. In what follows, I will further elaborate on what factors determine the individual strength of a claim.

> Non-aggregationism holds that evaluative statements involving notions such as "better" or "worse" can only be made from personal perspectives. For this reason, saving the greater number of people cannot be said to be better from anyone's perspective, unless it is indeed better for *everyone* involved. In the case of an interpersonal conflict, however, everyone involved ought to be respected equally by giving everyone a maximal chance of having their claim satisfied that is consistent with everyone else receiving the same chance.

Locating the Strongest Claims

The world importantly consists of an enormously large and complex interpersonal rescue conflict. An astronomical number of beings across the space-time continuum have held, currently hold or will hold claims to be helped. As individuals and collectives, we have limited resources to affect those beings. For instance, a philanthropist can spend her money to support certain causes and not others. As individuals, we also have a limited amount of time we can invest to contribute to providing help. As such, we constantly face an interpersonal rescue conflict. The conflict examined in the present context pertains to the trade-off between interventions benefitting people of the present, near- or medium-term future versus people of the long-term future. As outlined above, Greaves and MacAskill argue that it is a moral imperative to invest one's resources into affecting the long-term future, since the most beings who are morally relevant will be affected by that. However, how does a non-aggregationist resolve this enormous conflict?

As argued above, the non-aggregationist moral imperative plausibly holds that one ought to locate the strongest individual claims and prioritise accordingly. How could one go about identifying the strongest claims? I argue that at least two criteria ought to be taken into account: magnitude of the benefit and priority to the worst-off. Let us consider their implications in turn.

Magnitude of the Benefit and Priority to the Worst-Off

Firstly, as explained above, the stakes at hand for each individual matter. The first question to be answered is thus the following: To which individuals across time and space can we offer the largest benefits? To answer this question, we first need to ask whether increasing someone's happiness is equally as important as saving someone from harm. That is, should we consider the stakes for someone for whom we can increase their level of well-being from, say, 0 to +10 as equally as high as the stakes for someone for whom we can increase their level of well-being from -10 to 0? This same debate is being held among aggregationists as well, with some suggesting that there is a certain asymmetry between suffering and happiness which warrants the prioritisation of suffering. Many of

the arguments brought forward by aggregationists in this debate will likely also apply to the analogous debate for non-aggregationists. However, there may be an additional argument pointing to a suffering-focused view for non-aggregationists: Even if it is indeed true that we tend to disvalue suffering more than we tend to value happiness, aggregationists may still be able to argue that happiness ought to be prioritised in certain cases wherever there is a sufficient amount of people whose happiness can be affected, such that their number outweighs the suffering on the other side. For non-aggregationists, however, this line of argument is not available since it is not the aggregate happiness that matters, but the individual claims. Thus, non-aggregationists may more often have reason to focus on suffering than aggregationists do, at least to the extent that we assume a certain asymmetry between happiness and suffering to hold.

> Is increasing someone's happiness equally as important as saving someone from harm? Should we consider the stakes for someone for whom we can increase their level of well-being from 0 to +10 as equally as high as the stakes for someone for whom we can increase their level of well-being from -10 to 0?

Be that as it may – who among those individuals whose fate we may be able to influence are the ones with the strongest claims regarding the magnitude of the benefit they might receive by us? And where are they located across all of time and space?

Even though non-aggregationists do not consider the number of people as directly morally relevant, the enormous amount of people that may come into existence in the future does play a relevant role in answering this question. Since the vast majority of people holding a claim will live in the future, it is overwhelmingly likely that the people with the strongest claims in terms of the magnitude of the benefit will also live in the future. Considering the magnitude of the benefit we may be able to provide for individuals thus seems to point to a long-term focus. Note, however, that this argument does not take into account the probability by which these benefits will actually be conferred, but rather just the net benefit that would be conferred in case our help is successful. I will consider complications regarding the probabilities of success in section 4.

Let us turn to the second criterion. Many theories of distributive justice share a certain basic normative feature according to which – *ceteris paribus* – the worse-off ought to receive a certain priority. For instance, if a certain minority group is particularly bad off in a society, this constitutes a reason to redistribute and allocate resources to them disproportionately, i.e., to give them more than their arithmetically equal share.

Similarly, there may be a reason to prioritise someone whose well-being we can affect such that they will go from -20 to -10, rather than helping someone for whom our help would make a difference in terms of going from -10 to 0. Even though the magnitude of the benefit is the same, and we are concerned with avoiding suffering in both cases, there may still be a reason to prioritise those in particularly bad situations. If this criterion to determine the strength of a claim is correct, what does it suggest for the identification of the strongest claims across time and space? The same argument from above seems to apply here too: Given

the presumably enormously large number of people living in the future, it is overwhelmingly likely that the being living the worst life ever lived, i.e., the worst-off, will live in the future. Thus, in isolation, this criterion again seems to suggest a long-term focus for non-aggregationists.

How to Prioritise

Considering the two criteria just discussed, it may thus initially seem as though non-aggregationism also suggests a moral long-term focus. Since the strongest claims of people – in terms of the magnitude of the stakes and of how bad off they are – are likely located in the far future, non-aggregationists have moral reasons to prioritise the far future.

Before I turn to complications regarding the probabilities by which our actions can actually affect different people, I would like to address the question of how exactly non-aggregationists would prioritise the individuals in the far future, if they were to do so. Firstly, it is important to distinguish between conflict cases regarding divisible goods and conflict cases regarding indivisible goods. For instance, if A and B are both very poor and need monetary support, but A needs more than B, A might have to be prioritised. If there is a sufficient amount of money available, however, a sensible prioritisation of A does not necessarily consist of giving all the money to A, but rather of splitting the money according to their needs, i.e., the strength of their claims. If it is an indivisible good, however, and A's claim is stronger than B's, then one might either only support A, or perform a lottery that is weighted in A's favour in order to respect each of the claims according to their strength. Which of these two options is sensible may depend on the exact claim strengths: It is plausible that e.g. the claims resulting from one broken arm and two broken arms should both be considered in a conflict, and thus in a case of indivisible resources a lottery weighted in favour of the two broken arms should be performed. On the other hand, in a conflict between a broken arm and a threat of death, the claim resulting from the potential death should perhaps be prioritised outright.

How exactly non-aggregationists should prioritise the far future thus depends on at least two factors: Firstly, it depends on whether the respective resources invested are divisible; and secondly, it depends on whether the claims of individuals in the far future are likely to be sufficiently strong such that present and nearer-term claims should be ignored or whether present- or nearer-term claims are also sufficiently strong, such that they should merely be given a lower weight. There are certainly cases both in which the respective resources are divisible, such as when money is to be donated to charities, as well as cases in which they are not divisible, such as when choosing certain career paths that limit one's ability to contribute to causes other than the chosen ones.

Depending on the exact details of the non-aggregationist position and the empirical realities of different claim strengths and the resources available, the exact nature of the prioritisation may look different; an area of research certainly worth exploring further. While some positions might suggest an outright prioritisation of the strongest claims, others might suggest that we perform a complex (weighted) lottery between all relevant claims, possibly including the weaker near-term claims and the stronger long-term claims.

In this context, it is also interesting to note that an allegation often raised against consequentialist moral theories, such as var-

ious forms of utilitarianism, seems to apply just as much to this inherently non-consequentialist theory, namely the problem of cluelessness. It is often claimed that e.g. utilitarian theories are implausible because it is too difficult to determine all the empirically relevant factors influencing which actions will have the best overall outcomes. However, determining which claim strengths across time and space all deserve our consideration and respect, and potentially even performing an extremely complex lottery in order to determine which of those individuals eventually receive our help, seems to be at least as difficult, if not more so, as making a comprehensive consequentialist impact assessment.

Relatedly, both the aggregationist and the non-aggregationist must rely on complex empirical analyses regarding the quantification of individual (and collective) stakes. That is, it presupposes some form of ethical calculability of benefits and needs which is interpersonally comparable. To the extent that such analyses are too empirically difficult, they pose a problem for both long-termism and the non-aggregationist alternative I am sketching here. However, this is a problem any possible theory of large-scale moral prioritisation must ultimately solve, and a whole different canon of literature is already dedicated to it, particularly in health economics where concepts such as such as the QALY (quality-adjusted life years) or the DALY (disability adjusted life years) have been developed.

Probability Discounting

As hinted at above, an important factor has not been addressed so far. Even though the individuals with the strongest claims in terms of the two criteria outlined above may indeed live in the far future, the probability with which our actions will actually make a difference for these individuals may be vanishingly small. Consider, for instance, a rescue conflict between A and B where you can either decrease the probability of death for A by 0.00001 or certainly save B from becoming paraplegic. Even if A has the stronger claim in terms of all two aforementioned criteria, the fact that the probability of successfully helping A is so small must have an influence on A's claim strength. Indeed, it seems that in this case, B ought to prioritised, since, discounted by the respective probability, B has a lot more at stake regarding our help. Since there is a lot of uncertainty regarding the potential help we can provide to people in the far future, such cases may be analogous to the rescue conflict between individuals of the present- and the near-term future and individuals of the far future: We can be fairly sure that the right donation will in fact save someone from contracting malaria or from starvation, while an attempt to influence people in the far future may very likely not have an effect at all. Thus, if we discount the assistance by its probability of success, the actual benefit we can offer people in the far future may be much lower than the benefit we can offer people in the present or nearer-term future. For aggregationists such as Greaves and MacAskill, this fact does not play as much of a role, particularly given their endorsement of expected value theory. Since even very small probabilities of affecting a vast future population by e.g. reducing the risk of extinction results in a large expected value given aggregationist premises, discounting every individual single claim with a very small probability does not undermine their longtermist conclusions. In what follows, however, I will discuss the factors which determine whether the argument does undermine longtermist conclusions for non-aggregationists.

Consider a rescue conflict between A and B in which you can either decrease the probability of death for A by 0.00001 or certainly save B from becoming paraplegic. Even if A has the stronger claim in terms of all two aforementioned criteria, the fact that the probability of successfully helping A is so small must have an influence on A's claim strength.

Stochastic (In)dependence

Expected value theory, by itself, does not distinguish between the following two cases.

Case 1: There are 1000 people on a ship which is in danger of sinking, thereby killing all the people on it. You have the possibility of negatively affecting the probability of it sinking by 0.001. In expectation, you save one person with this action. You give thus each of the people on the ship an ex ante benefit of 0.001 of survival.

Case 2: There are 1000 people threatened to be killed by a vicious disease. You can give each of those 1000 people a medication which lowers each of their respective probabilities of dying by 0.001. In expectation, you save one person with this action. You give thus each of the sick people an ex ante benefit of 0.001 of survival.

In Case 1, the probabilities of providing help for the 1000 people are stochastically dependent: you either save all of them or you save none. In Case 2, on the other hand, the probabilities are stochastically independent, i.e., some may die and some may survive. For aggregationist expected value theorists such as Greaves and MacAskill, there is little reason to distinguish between these two cases. However, there does seem to be a morally important difference between them. There are two important perspectives to be taken into account here: ex ante and ex post. Ex ante, i.e., before the action has taken place, it is indeed the case that the actions seem to benefit each of the people in Case 1 and Case 2 equally. They each get an ex ante benefit of 0.001 of survival and the ex ante expected value is the same, namely one person saved. However, ex post, i.e., after the action has played out, the probabilities of which outcomes have actually occurred are radically different: While in Case 1, the probability that at least one of the people has actually survived is 0.001, while the probability that at least one person survived in Case 2 is higher than 0.6, since the outcomes for the stochastically independent chances are more spread out. From an ex post perspective, we can be much more sure that we will indeed have satisfied at least one claim in Case 2, which arguably constitutes a reason to prioritise the group in Case 2 over the group in Case 1. Even though from the ex ante perspective, each person in Case 1 and Case 2 gets the same expected benefit, having satisfied at least one claim ex post must be relevant for non-aggregationists as well. If we hold both the ex ante and post perspective to be important, then both of these perspectives have to be taken into account when evaluating the relevance of probability discounting.

Longtermism Ex Ante and Ex Post

Consider a non-aggregationist who attempts to compare the claim strengths between individuals of the far future and individuals in the present- or near-term future. As outlined in the previous section, notwithstanding success probabilities, it may look as though individuals in the far future hold the stronger claims. However, what does it look like if we add the discounts due to the probabilities of success? This depends crucially on whether we take an ex ante or an ex post perspective.

As explained above, it seems that the claims of future individuals would have to be discounted by a lot, when considered from an ex ante perspective. The benefit we can offer to an individual of the present or near-term future when discounted by the probability of success is a lot higher than the benefit we can offer an individual of the far future, given that it is extremely uncertain whether our help will affect them at all. Individuals of the present and nearer-term future thus clearly seem to hold a stronger claim from an ex ante perspective, once the discounts given the probabilities of success are added.

However, it may look differently from an ex post perspective. This depends crucially on whether affecting the longer-term future is more analogous to Case 1 or analogous to Case 2. While the numbers do not count in and of themselves for a non-aggregationist, the number of people influenced does make a difference even for a non-aggregationist if the benefits offered to them are stochastically independent. To see why, consider two further cases:

Case 3: You can either relieve individual A of medium pain with a probability of 1, or provide a probability of 1/100 of a relief of intense pain for individual B.

Case 4: You can either relieve individual A of medium pain with a probability of 1, or provide a probability of 1/100 of a relief of intense pain for each person of a group of 100,000 individuals.

From an ex ante perspective, individual A probably ought to be prioritised in both cases according to non-aggregationism, because their claim is stronger when probability discounting is added. However, from an ex post perspective, it is incredibly likely that at least one of the 100,000 individuals in Case 4 will have been spared from intense pain, and thus the large number of the group in Case 4 provides a reason to prioritise them from an ex post perspective. Ex post, it is simply very likely that at least one person in the large group will have benefited more from your helping them than A would have, if you had helped them.

Returning to the question of longtermism, non-aggregationists would have to ask the following question: Are there actions that grant the vast populations of the future probabilistically small but stochastically independent benefits? Or is influencing the future more analogous to Case 1, such that the probability of success for having helped any of the people in the future is very small ex post? If our situation is analogous to Case 1, the ex post perspectives would converge with the ex ante perspective and suggest a nearer-term focus for non-aggregationism. However, if our situation is analogous to Case 2, the ex post perspective likely supports a long-term focus, since there would be a very high likelihood that a strong claim will have been satisfied ex post.

So which of the two cases are we actually faced with? Unfortunately for longtermists, I believe that influencing far off future populations will turn out to be mostly analogous to Case 1. In particular, their focus on interventions such as reducing extinction risks or influencing high-impact developments such as work-

ing on AI alignment, as Greaves and MacAskill suggest, seems to influence future populations in a stochastically dependent way. Extinction influences the future (non-)existence of all future populations simultaneously and different AI alignment scenarios influence the shape of the world and universe for all future populations at the same time. For if humanity goes extinct, this extinction event will have been the cause for the non-existence of all future populations simultaneously. Similarly, the world created by a misaligned superintelligent AI will influence future populations simultaneously. To the extent that this reasoning is correct, it also does not matter much which exact extinction risk we are considering, for *if* humanity does go extinct, it will have influenced the nonexistence of future populations simultaneously, no matter which exact extinction risk was ultimately realised.

One may object that some extinction risks influence existing people differently, rather than simultaneously, before extinction actually occurs. For instance, climate change will have adverse effects on different parts of the global population in different ways, and various interventions will influence people differently. However, this is not an argument the longtermist view has at its disposal, since its focus on extinction risk is primarily justified by the vast number of people who will not come into existence in case of extinction, rather than the people who are influenced by these risks before extinction occurs.

If this reasoning is correct, neither extinction risks nor AI alignment influence people in a stochastically independent way. Thus, if the probabilities of success regarding respective interventions is sufficiently small, the vast number of people who are influenced by them is still not relevant for a non-aggregationist from an ex post perspective. For this reason, I believe that non-aggregationism does ultimately not support longtermist conclusions. Consequently, a strong focus on reducing extinction risks must be justified in a different manner to non-aggregationists, if at all.

There may be a certain sweet spot between the present and the far future, such that the probabilities of success are still sufficiently high and the number of people sufficiently large such that some of the people affected are likely going to have very strong claims in terms of the two criteria outlined above. For instance, the number of people who will be living in the next few centuries is likely going to be a lot larger than the number of people living today. Hence it is also likely that some of these people will hold some of the strongest claims. At the same time, the probability of success regarding possible help for these future people may be sufficiently high for it to still be worth it, even if the respective help is stochastically dependent. In particular, non-aggregationists may get behind efforts to reduce extinction risks or other high-impact developments to the extent that they are sufficiently tractable, such that the probability of success is sufficiently high to be worth it. Non-aggregationism may thus support a much weaker form of anti-presentism, according to which the very strong focus on present issues is viewed critically, but not replaced with a focus on the very long-term future but a near- to medium-term focus instead.

Conclusions and Outlook

In this paper, I have argued that strong non-aggregationism is a relevant alternative to aggregationism. It has a compelling rationale and can be defended against obvious objections, such that it should be taken seriously when considering which large-scale moral imperatives we are faced with. For this reason, it is im-portant to consider what exact moral recommendations follow from non-aggregationism. In this paper in particular, I explored its implications for longtermism and a focus on extinction risks and other high impact long-term developments. To this end, I outlined one version of non-aggregationism and argued that such non-aggregationists ought not to be on the lookout for the largest groups to save, but rather for the individuals with the strongest claims to satisfy. I have furthermore argued that at least two criteria are relevant for determining claim strengths: magnitude of the benefit and priority to the worst-off.

Considering these two criteria in isolation, non-aggregationism seems to suggest a long-term focus as well, and thus influencing the existence and welfare of generations in the far future would be a central imperative of non-aggregationism, too.

However, when taking the probabilities of success into account, this result cannot be maintained. From an ex ante perspective, the discounts resulting from the uncertainty of being able to affect the far future suggest that non-aggregationists ought not to prioritise the far future. From an ex post perspective, it depends on whether the probabilities with which people in the far future can be influenced are stochastically dependent or independent. In the former case, the ex post perspective likewise suggest more of a near- to medium-term focus, rather than a long-term focus. In the latter case, longtermist conclusions look more attractive. However, since the empirical reality seems to be more similar to the former case, I have tentatively concluded that non-aggregationism likely does not recommend a long-term focus. In particular, reducing extinction risks is an example of stochastically dependent probabilities with which future generations can be influenced. However, to the extent that the extinction risks are sufficiently tractable, non-aggregationists can support a respective focus as well, even if less strongly so.

There is ample space for further research in this area. All the different versions of non-aggregationism could be examined regarding their implications for longtermism. Furthermore, more research into the criteria which determine claim strengths might turn out to be very important in answering the questions I have outlined. In this context, investigating both different versions of prioritisation and the strength of the cluelessness objection for non-aggregationist theories may likewise be very interesting. Finally, further work on the correct analysis of the implications of stochastically dependent and independent probabilities of success, and the correct analysis and combination of the ex ante and ex post perspectives will also likely significantly influence the answers to the questions I have outlined.

Notes

1 Note, however, that the term "anti-presentism" in this context is not meant to imply that there is normative discounting of present interests. Rather, insofar as present and future generations all count equally, the future generations just vastly outweigh the present generations such that the latter end up mattering much more. Thus, to the extent that there are trade-offs in benefitting the present generations and future generations, one ought to opt for benefitting the vastly greater future generations. "Anti-presentism" in this context is not meant to refer to anything other than this basic implication of the classic longtermist argument.

2 I thank an anonymous reviewer for this helpful example.

3 The former is often referred to as one of the most cost-effective

ways to save lives in the present (Greaves and MacAskill 2021: 2), while the latter is one of the explicit priorities of many longtermist organisations, such as e.g. the career advice centre "80000 hours".
4 I thank an anonymous reviewer for the suggestion to clarify this point.

References

Beckstead, Nicholas (2013): On the Overwhelming Importance of Shaping the Far Future.
Ph.D. thesis, Rutgers University, New Jersey. https://rucore.libraries.rutgers.edu/rutgers-lib/40469/PDF/1/play/. Viewed 9 December 2022.

Bostrom, Nick (2013): Existential Risk Prevention as Global Priority. In: Global Policy, 4 (1), 15-31. https://existential-risk.org/concept.pdf. Viewed 9 December 2022.

Deutscher Bundestag (2022): Entwurf eines Zweiten Gesetzes zur Änderung des
Infektionsschutzgesetzes. https://dserver.bundestag.de/btd/20/038/2003877.pdf. Viewed 9 December 2022.

Eurotransplant (2019): Eurotransplant Manual. https://www.eurotransplant.org/allocation/eurotransplant-manual/. Viewed 9 December 2022.

Greaves, Hillary, and William MacAskill (2021): The case for strong longtermism.
GPI Working Paper no. 5-2021. Global Priorities Institute. https://globalprioritiesinstitute.org/wp-content/uploads/The-Case-for-Strong-Longtermism-GPI-Working-Paper-June-2021-2-2.pdf. Viewed 9 December 2022.

Heikkinen, Karri (2022): Strong longtermism and the challenge from anti-aggregative moral views. GPI Working Paper no. 5-2022. Global Priorities Institute. https://globalprioritiesinstitute.org/wp-content/uploads/Karri-Heikkiken-GPI-working-paper-1.pdf. Viewed 9 December 2022.

John, Tyler M. / Millum, Joseph (2020): First Come, First Served? In: Ethics, 2 (130), 179–207.

Krznaric, Roman (2020): The good ancestor: How to think long term in a short-term world. London: Ebury Publishing..

Lübbe, Weyma (2015): Nonaggregationismus: Grundlagen der Allokationsethik. Münster: Mentis.

MacAskill, William (2022): What we owe the future: A million-year view. London: Oneworld Publications.

Mannino, Adriano (2021): Wen rette ich, und wenn ja wie viele? Stuttgart: Reclam.

OPTN Organ Procurement and Transplantation Network (2019). Ethical Implications of Multi-Organ Transplants. Richmond. https://optn.transplant.hrsa.gov/media/eavh5bf3/optn_policies.pdf. Viewed 9 December 2022.

Ord, Toby (2020): The Precipice. Existential Risk and the Future of Humanity. New York: Hachette Books.

Parfit, Derek (1984): Reasons and Persons. Oxford: Oxford University Press.

Pettigrew, Richard (2022): Should Longtermists Recommend Hastening Extinction Rather Than Delaying It? Unpublished manuscript.

Rakić, Vojin (2018): Disaster Consequentialism. In: O'Mathúna, Dónal P. / Dranseika, Vilius / Gordijn, Bert. (eds.): Disasters: Core Concepts and Ethical Theories. Cham: Springer International Publishing, 145–156.

Rawls, John (1971): A Theory of Justice. Cambridge, MA: The Belknap Press of Harvard University Press.

Taurek, John (1977): Should the Numbers Count? In: Philosophy and Public Affairs, 6 (4), 293–316.

Thompson, Dennis F. (2010): Representing Future Generations: Political Presentism and Democratic Trusteeship. In: Critical Review of International Social and Political Philosophy, 13 (1), 17–37.

Tremmel, Jörg (2021): The Four-Branches Model of Government: Representing Future Generations. In: Cordonier Segger, Marie-Claire / Szabó, Marcel / Harrington, Alexandra R. (eds.): Intergenerational Justice in Sustainable Development Treaty Implementation: Advancing Future Generations Rights through National Institutions. Cambridge University Press, 754–780.

Marina Moreno is a PhD student at the Munich Center for Mathematical Philosophy at LMU Munich and holds an MA in logic and philosophy of science.
Email: marinaestrellamoreno@gmail.com