

## Article

# Ethical analysis of purported risks and disasters involving suffering, extinction, or a lack of positive value

Simon Knutsson <sup>1,\*</sup><sup>1</sup> Stockholm University; Center for Reducing Suffering\* Correspondence: [simonknutsson@gmail.com](mailto:simonknutsson@gmail.com)

**Abstract:** I carry out an ethical analysis of how we should deal with the following purported risks and disasters: suffering on an astronomical scale, personal disasters, extinction, and the possibility that positive value will not be created. I consider them in relation to one another in part because measures to reduce one risk may increase another risk, and because there are opportunity costs and tradeoffs. I build largely on ideas from the ethics of risk. For example, it seems that risk imposition is especially morally problematic when an unconsenting individual is subjected to the risk of extreme ill-being for someone else's pure benefit. My findings include that there are ways to reduce personal disasters, such as illnesses, that create little risk. Measures to prevent extinction and ensure that positive value is created should generally be limited to measures that also reduce (or at least do not increase) the risk of large-scale severe ill-being. Examples of measures on the table include forms of moral improvement and helping victims of war in ways that prevent conflicts.

**Keywords:** risk of astronomical future suffering; s-risk; existential risk; x-risk; extinction, ethics of risk, ethical risk analysis, pure benefits, suffering-focused ethics

**Citation:** Knutsson, Simon. 2025.Ethical analysis of purported risks and disasters involving suffering, extinction, or a lack of positive value. *Journal of Ethics and Emerging Technologies* 35: 2.<https://doi.org/10.55613/jeet.v35i2.163>

Received: 20/02/2025

Accepted: 20/02/2025

Published: 10/03/2025

**Publisher's Note:** IEET stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Central to this paper are risks of disasters that involve suffering or ill-being on a large scale.<sup>1</sup> In particular, I will pay special attention to the worst kinds of risks and disasters that involve the most severe, large-scale suffering. That said, in part because measures to reduce one risk can increase another risk, and because there are opportunity costs (in the simple sense that by doing something we forgo doing something else), risks should reasonably be considered in relation to other risks and opportunity costs. In that spirit, I will deal with a variety of purported risks and disasters, namely personal disasters, a lack of purported positive value being created, extinction, and risks of astronomical suffering.

My main contribution in this paper is an ethical analysis of these purported risks and disasters. My approach is not to apply a specific general ethical theory such as a form of utilitarianism. Rather, I largely draw on ideas from the ethics of risk, such as ideas about the moral cost of imposing risk, who decides about risk-imposition, consent, and whether the risk-exposed individual is the one who would benefit. I also consider the morality of risking someone's extreme suffering for a pure benefit that does not reduce anything bad.

With that ethical background in mind, I analyse how we should address the different purported risks and disasters. Much of my analysis concerns opportunity costs, tradeoffs, and how to avoid morally problematic risk-imposition. My findings include that there are ways to reduce personal disasters such as illness that create little risk, and that there are moral problems with various measures to prevent extinction and ensure that positive value is created in the future. In sum, the considerations point towards choosing measures that

<sup>1</sup> I use the term 'ill-being' interchangeably with 'negative well-being', 'negative quality of life', and 'negative welfare'.

also reduce (or at least do not increase) the risk of large-scale severe ill-being. For example, moral improvement and preventing conflicts may be robustly beneficial.

Section 2 sets the stage with some preliminaries on disasters and risks. Section 3 explains what risks of astronomical suffering (s-risks) are, while section 4 describes risk factors and interventions related to s-risks that have been mentioned in the literature on s-risks. In that section, I also make some brief comments on such risk factors and interventions. Section 5 is about substantive ethics relevant to the analysis. The section is largely about the ethics of risk, and I present support for the view that some risk-taking is arguably morally problematic. Section 6 is dedicated to the more concrete ethical analysis of how to address purported risks and disasters involving personal disasters, a lack of positive value being created, extinction, or astronomical suffering. Section 7 concludes.

## 2. Preliminaries on disasters and risks

I understand disasters and risks to be negative. That is, the term ‘disaster’ typically refers to something very bad, and risks involve something unwanted, undesirable, or the like.<sup>2</sup> So I take disasters and risks to be value-laden (e.g. Hansson, 2010). As disasters and risks have a value component, the question arises as to what *should* be considered a disaster or risk (and whether the phenomenon in question is indeed very bad, undesirable, etc.). This question will recur later in this paper. I sometimes write ‘purported’ risks and disasters because the value question of whether something is a risk or disaster is not settled from the outset, and I especially use ‘purported’ to convey that I doubt it qualifies as a risk or disaster.

In addition, there are different kinds of disasters. For example, Raphael (1981) uses the term ‘personal disaster’ to describe intensely distressing personal experience such as bereavement, serious illness, and the experience of famine. I think of a personal disaster as a disaster for one individual.<sup>3</sup> One might think that personal disasters are not large-scale disasters, but if there are many personal disasters, that would be disasters on a large scale.

Different features of disasters such as temporal and spatial dimensions have also been discussed. For example, Hsu (2019) puts forth an approach according to which disasters can be sudden, gradual, or chronic, and says that there is a need to understand how disasters can be spatially diffuse. In my understanding of disasters, they may be chronic and spatially spread out.

When it comes to reducing risks, authors have noted that the cost of taking measures against some threat or risk includes the cost of *increasing* some other threat or risk.<sup>4</sup> That is a crucial consideration that will recur in this paper, and it is one reason why it makes little sense to merely think about how to reduce a specific risk without taking into account other risks, and whether the reduction of one risk would in turn increase others.

---

<sup>2</sup> On conceptualisations of disasters, see O’Mathúna and Gordijn (2018, p. 2) and Sandin (2018). On concepts of risk, see Hansson (2010, 2023, Section 1) and Boholm, Möller, and Hansson (2015).

<sup>3</sup> Canadian Red Cross (n.d.) speaks of personal disasters as affecting a *household*: “Personal disasters typically affect one household. House fires are one of the most common examples”.

<sup>4</sup> E.g. Munthe (2011, pp. 3–4): “Whatever unclarities are built into the notion of PP [the Precautionary Principle], this principle strives to say something about our (or society’s, or some particular acting party’s) responsibility for ‘taking measures’ against various potential ‘threats’. No matter what this amounts to more precisely, it means that what PP prescribes or recommends (i.e. precaution, in some sense of this word) will always have a price. This includes, of course, straightforward monetary costs for the measures in question, but not only that. First, the precautionary measures will always in themselves create some threats or risks of their own.” See also the literature on risk–risk analysis and risk–risk tradeoffs (e.g. Hansen et al., 2008; Hansson, 2017).

### 3. Introducing risks of astronomical suffering (s-risks)

In this section, I briefly explain what risks of astronomical suffering (s-risks) are. One can try to imagine an extreme s-risk by trying to imagine a world that is as terrible as possible. For example, one might imagine that such a world contains only horrible, relentless, sickening torment, cruelty, and slaughter. Such a world might have features such as that all victims are innocent and are begging to be killed. These horrors would occur at enormous, astronomical scales much larger than anything that has occurred on Earth. This imagined world is an extreme case in that the exercise was to imagine a world that is close to being as horrible as possible. One can turn down the horror one or more notches and still be left with horrible worlds. For example, perhaps somewhat fewer suffer, or do not suffer as severely, but are still in an awful state; such a scenario would still count as an s-risk.

More generally and concisely, an explanation of s-risk by Baumann (2022, pp. 8–9) reads: “s-risks are scenarios that involve severe suffering on an astronomical scale, vastly exceeding all suffering that has existed on Earth so far.”<sup>5</sup>

As alluded to in the previous section, we face value questions such as whether a mixed outcome that contains both many horrible lives and more fortunate lives *should* count as a risk or disaster given the purported positive value in the outcome. I think so, and s-risks are generally considered to include such mixed scenarios as long as they contain vast severe suffering.

### 4. On proposed risk factors and interventions related to risks of astronomical suffering

This section is mainly background for readers who are unfamiliar with the literature on s-risk, although I will also make a few remarks that are not purely background. The literature on s-risk contains discussions of risk factors and interventions. I will first talk about risk factors (sec. 4.1) and then interventions (sec. 4.2).

#### 4.1. Risk factors related to s-risks

Risk factors make an outcome with astronomical amounts of severe suffering more likely or more severe (Baumann, 2019). Several risk factors have been mentioned, including certain forms of advanced technology, large-scale space colonisation, a lack of s-risk prevention, conflict, and malevolent actors (Althaus & Baumann, 2020; Baumann, 2019, 2022, Chapter 7; Sotala & Gloor, 2017).

It may be worth highlighting the capacity for large-scale creation — the ability to bring about vast numbers of new existences. If the number of existences is fixed, those lives could be horrible and potentially very long, but instances of severe misery would still be relatively limited compared to if the number of existences could be vastly increased. So the ability to create and increase the number of individuals seems to be a basic risk factor for s-risks.

A group of risk factors that may have received less attention in the literature include values and ideas in philosophy. A discussion of such risk factors related to extinction (not s-risks) is provided by Leslie (1996, pp. 10–13). He lists plenty of diverse “risks from philosophy”, including inalienable rights and avenging justice, for example, carrying out a threat of revenge. One can debate whether all the items on his list really are risk factors for extinction, and even if something is a risk factor for one type of purported risk such as extinction, that does not imply that the risk factor is all things considered more risky than beneficial. Anyway, we can ask what a corresponding list related to s-risks would look like. There would probably be some overlap with Leslie’s list. For example, ideas

---

<sup>5</sup> Similar formulations can be found in Sotala and Gloor (2017, p. 389) and Daniel (2017). A similar risk is described in terms of disvalue by Gloor (2018): “we can define downside risks as events that would bring about disvalue (including things other than suffering) at vastly unprecedented scales”.

favourable to revenge and carrying out threats are also plausibly risk factors for s-risk.<sup>6</sup> Still, there are presumably some differences.<sup>7</sup>

Without doing a thorough investigation here, a list of “risks from philosophy” or the like related to s-risks may include the idea that adding more and more positive lives is morally important even if it comes at the cost of increasing extreme suffering. Perhaps the list would contain a willingness to take risks of great harm, ideas of us having a cosmic purpose that involves exploring and settling the universe, or ideas that are favourable towards a flourishing science-fiction civilisation.

I bring up such potential risk factors for s-risks because it seems more fruitful work could be done on the topic, and because it is related to the topic of moral improvement as a way to reduce s-risks, which I will discuss below. To be clear, even if something like an idea or theory is a risk factor, it obviously does not follow that it should be silenced. Similarly, even if something is a risk factor for a certain kind of risk (in this case s-risk), that does not, by itself, imply that it is overall harmful or undesirable.

#### 4.2. *Interventions related to s-risks*

The literature on s-risks covers a wide range of possible interventions.<sup>8</sup> For example, some interventions concern improving space governance and international cooperation, as well as avoiding escalating conflicts and artificial intelligence arms races. Other interventions include safety mechanisms for artificial intelligence systems that are aimed at preventing s-risks.

Some interventions are related to morality, values, or personal characteristics. In general terms, there are ideas about improving our values and favouring careful moral reflection (e.g. Baumann, 2020, 2021, Section 1). One intervention is to ensure that there is moral concern for all sentient beings (Baumann, 2022, pp. 57–62). Related interventions concern agents who score high in traits such as Machiavellianism, narcissism, psychopathy, and sadism. Specific interventions to reduce the influence of such agents have been discussed, including political interventions to avoid such agents rising to power, and genetic enhancement to select against such traits (Althaus & Baumann, 2020, Section 4.2, 4.3.3; Vinding, 2022a, Section 14.6).

Relative to the existing literature, it seems that more could be said about the following (which I will discuss in sec. 6.4): If measures are taken to improve morality via, for example, moral bioenhancement or learning, then which morality-related characteristics are desirable from an s-risk perspective? For instance, compassion may be an especially desirable characteristic.

### 5. **Substantive ethics for my analysis**

In this section, I go through substantive ethical matters that are relevant to the analysis in the next section. In my analysis, I do not aim to apply a specific general ethical theory such as a form of consequentialism or contractualism. So I will not defend any such theory. Rather, I aim to note and sort out some ethical considerations that are relevant to the purported risks and disasters, in particular, considerations related to the ethics of risk and the moral importance of ill-being.

In the literature on the ethics of risk, we find essentially the point that it is morally costly to create or impose a risk, and that we should avoid doing so unless there are

<sup>6</sup> This has been discussed in the literature on s-risk (e.g. DiGiovanni, 2023, Section 3.2, 4.2).

<sup>7</sup> Although for almost anything, one can imagine a way in which it could make a risk or disaster more likely or severe, the following items in Leslie’s list do not seem to be important risk factors for s-risk (or not even risk factors at all): negative utilitarianism, Schopenhauerian pessimism, metaethical theories, and ideas about inalienable rights.

<sup>8</sup> Baumann (2017a, Section 6, 2017b, 2021, 2022, Part III), Gloor (2016), DiGiovanni (2023, Section 4), and Althaus and Baumann (2020). The Center on Long-Term Risk does work to address worst-case risks from artificial intelligence.

sufficient reasons to impose the risk.<sup>9</sup> Imposing risks on others can be morally irresponsible (Munthe, 2011, Chapter 5). There are many potential examples, such as storing explosives unsafely, drunk driving, and encouraging a small child to play on a cliff.<sup>10</sup> It seems plausible that even if things turn out fine, there is something morally problematic about the risk-imposition.<sup>11</sup>

Similarly, there is the idea that every person has a *prima facie* right (which can be overridden) of not being exposed to risk by others.<sup>12</sup> Each risk-exposed person is to be treated as a sovereign individual and it needs to be justified why it is acceptable that this individual is exposed to the risk (Hansson, 2000, p. 7). All that sounds reasonable to me, although I will not rely on the idea that there are rights. Rather, my analysis will reflect the general idea that there is a moral hurdle when it comes to justifying risk-imposition and that it is not enough to say that imposing unconsented risk of extreme ill-being is morally right simply because the benefits to others are greater.

In the literature on the ethics of risk, we also find specific aspects that have been taken to be relevant to the morality of imposing risk. These aspects include whether there is a risk of *great* harm; whether the person who is subjected to the risk consents; whether the risk-exposed person is different from the person getting a chance of benefit; and who decides about the risk-imposition.<sup>13</sup> For example, it has (in my opinion, reasonably) been considered less morally problematic when someone decides to subject themselves rather than others to a risk (Hansson, 2018). In real life, actions may often create some risks for others, but we can at least imagine a case where someone takes a risk for their own benefit with no effects on others. Imagine an old hermit without a family and any other relationships living alone in the wilderness who decides to swim for exercise and enjoyment although there is a risk of drowning. We stipulate that other beings will be unaffected regardless of what happens. By comparison to imposing risk on others, taking such a personal risk seems much less morally problematic. In contrast, it has been considered (again, reasonably, I think) especially morally problematic when someone is subjected to a risk without having a say in the matter, and when someone else gets the chance of a benefit. Similarly, it has reasonably been considered morally problematic if the decision-maker is also the beneficiary while someone else faces the risk. For example, it is more morally problematic if I decide to risk your life rather than mine in order to make a fortune for myself.<sup>14</sup>

In addition, an important idea that I find very plausible is that risk-taking is especially morally problematic when a risk of (great) harm is imposed for a pure benefit (Magnusson, 2022; Rulli, 2024; Shiffrin, 1999, 2012). I will understand a pure benefit to be a benefit that does not reduce anything bad. We need to be mindful when we think of pure benefits because something that might be thought of as a pure benefit might also

---

<sup>9</sup> E.g. Munthe (2011, p. 176): “The basic intuition underlying my theory of the morality of precaution is that creating risks always has a morally significant price. If there are no reasons in favour of the claim that paying this price is worthwhile, we should avoid creating risks as far as possible.”

<sup>10</sup> Munthe (2011, p. 51) presents an example with explosives.

<sup>11</sup> See Munthe (2011, pp. 51–52).

<sup>12</sup> See Hansson (2000, p. 7). Hansson (2003, p. 303) proposes an exemption problem: “It is a *prima facie* moral right not to be exposed to risk of negative impact, such as damage to one’s health or one’s property, through the actions of others. What are the conditions under which this right is overridden, so that someone is allowed to expose other persons to risk?”

<sup>13</sup> See Hermansson and Hansson (2007), Hayenhjelm and Wolff (2012), and Hansson (2018). On *great* harm, see e.g. Munthe (2011).

<sup>14</sup> This example is essentially from Hansson (2003, p. 302).

reduce bads. For example, convenience or entertainment need not be pure benefits because they may reduce frustration and boredom. A perhaps hypothetical pure benefit may be that someone's pleasure when eating is increased in a way that has no other effects in terms of reduced badness. Another candidate of a pure benefit is the purported benefit to a newly created individual of existing with positive well-being (we set aside any effects on others).

Let us use an example to consider the morality of imposing risk of great harm to reduce a severe ill versus to provide a pure benefit. Suppose that we encounter a severely injured person who needs urgent medical treatment to avoid a chronically painful condition. We quickly drive them to the hospital, and since we are in a hurry, we are less cautious than when we normally drive, and we thereby increase the risk to, for example, nearby pedestrians. Provided that the situation involves only marginally increased risk for pedestrians and a good chance of avoiding chronic pain for the person in need, it seems morally right to urgently drive the person to the hospital, even at the cost of the increased risk for others. In contrast, it seems plausible that we should not drive like that and expose pedestrians to the risk of severe injury only to increase the intensity of someone else's pleasure. That would be unjustifiable. At least, it would be more morally problematic than driving like that to prevent chronic pain.

There are issues of the moral weight or importance of harm, badness, or suffering that are similar to the just-mentioned issue of imposing a risk of harm for a pure benefit. For example, in the context of risk, Munthe (2011, Section 5.5-5.7) presents a theory that gives extra moral weight to harms. In other words, when it comes to morally responsible risk impositions, badness has extra weight compared to goodness. Not specific to risk, there is the broad ethical topic of how morally important it is to alleviate and prevent severe suffering (e.g. Knutsson & Vinding, 2024).<sup>15</sup> One need not accept the following kinds of ideas to hold that suffering or badness is especially morally important, but these ideas are relevant. The ideas are that the creation of positive value is not morally important, that nothing has positive value for its own sake, and that positive well-being does not exist (e.g. Knutsson & Vinding, 2024, Section 3.2-3.3). I will sometimes write 'purported' positive value because its existence is open to debate (and I doubt it exists). That said, my analysis in section 6 is not a mere application of the idea that reducing suffering is especially important, and it is definitely not a mere application of the more controversial idea that there is no positive value.

When thinking about how to address risks and disasters, it is natural to think of interventions. Another angle is to ask, among other things, how we should respond and be in relation to the risks and disasters (Knutsson & Munthe, 2017; see e.g. Munthe, 2011, Section 4.2). For example, what should we notice and pay attention to? How should we react? What should we recognise as a morally relevant consideration? What should our emotional and cognitive patterns be like? Which motives and intentions are appropriate? Such things can be assessed ethically and they can be practically relevant, for instance, by making certain actions more likely and by forming a robust basis from which an agent is better able to handle new situations and decisions.

Before we turn to our analysis, I will end with some comments on the following issue, which I see more as a matter of substantive ethics rather than mere terminology: I have talked about *imposing* risk, *creating* risk, and *exposing* or *subjecting* someone to risk. For ethical purposes, what counts as imposing, creating, exposing, or subjecting? There are seemingly clear cases where all these four terms apply, such as unnecessarily storing explosives unsafely in an apartment building. Other cases seem less clear, such as when it comes to counterfactuals and omissions. Let us consider omissions. Some omissions

---

<sup>15</sup> In addition, there are ideas in population ethics to the effect that it is more important to avoid adding lives with negative welfare than to add lives with positive welfare (see e.g. the Asymmetry and the idea that lives with positive well-being have diminishing marginal value). See also Wolf's (2004, p. 63) Misery Principle.

seem straightforward to evaluate, for example, if one's job responsibility is to assess aeroplane safety and one skips critical assessments. Other cases of omission seem less clearly like risking in the morally relevant senses. For example, consider an adult's decision not to have a child; that is, if one has no children, the decision is to not have any children, and if one already has three children, the decision is to not have a fourth child. Suppose that the child would have reduced some risks and increased others during their lifetime. Given that the adult did not have a child, did the adult thereby create, impose, expose, or subject others to the risks that will not be reduced due to the absence of the child? It sounds a stretch to say that the adult did that. This idea that some omissions may not count or count less as risk-imposition and the like is not a premise that I will rely on in my analysis. Rather, I see it as something that *might* make a moral difference and I will mention it in my analysis (although I do not here *claim* that it makes a moral difference).<sup>16</sup>

## 6. Ethical analysis of how we should address various purported risks and disasters

I now proceed to the ethical analysis of how we should address various purported risks and disasters. I will go through personal disasters (sec. 6.1), a lack of purported positive value being created (sec. 6.2), extinction (sec. 6.3), and then end with risks of astronomical suffering (sec. 6.4).

Since the analysis includes opportunity costs and the possibility that reducing one risk may increase another, my division of these respective risks and disasters into separate subsections does not indicate that they are isolated. On the contrary, there are strong connections between them.

I do not aim to do an exhaustive ethical analysis that takes every ethical consideration into account. My ambition is rather to include some of the perhaps most important points that are worth making here.

### 6.1. Personal disasters

To repeat, I think of a personal disaster as a disaster for one individual, such as experiencing a serious illness. The separation of personal disasters and risks of astronomical suffering (s-risks) is not strict because s-risks can be thought of as personal disasters on an extremely large scale if we understand personal disasters broadly enough. Because I will talk about s-risks later, I will in this subsection pay more attention to personal disasters that do not constitute s-risks or at least not the most typical examples of s-risks.

Personal disasters matter a lot morally, of course. They are associated with there being a victim, severe distress, a life's being ruined, and so on.<sup>17</sup> In practice, there are serious opportunity costs to doing something other than alleviating and preventing personal disasters. We can think of reducing personal disasters as a reference point and keep in mind whether something is more important than that.

If we consider aspects mentioned in section 5 that arguably make risk-imposition more morally problematic, we can note that the risk of *great* harm is present when we subject others to the risk of personal disasters (essentially by the definition of 'personal disasters').

When it comes to the characteristics of an agent, several generally desirable characteristics seem to be in line with concern for, attention to, and efforts to alleviate and prevent others' personal disasters. Consider, say, compassion, care, empathy, and resolve to make a difference. Similarly, it sounds undesirable to be uncompassionate or heartless, or to downplay others' misery.

Let us discuss whether addressing personal disasters substantially increases other risks. It is useful to make a few distinctions. One is the distinction between scenarios in

<sup>16</sup> See e.g. Sartorio (2009) on causation and ethics, and Beebe (2004) on causation by absence.

<sup>17</sup> For a text on such features, although not explicitly about disasters, see e.g. Knutsson and Vinding (2024, Section 3).

which existence is held constant (e.g. someone feels less badly but exists regardless) from those in which existence is variable (e.g. due to death or the creation of new beings). Another distinction is between risks with *attempts* to reduce personal disasters (e.g. risks from using a technology) versus risks that may come with *successfully* avoiding personal disasters (e.g. risks from survival or better health). I will focus on risks from successfully avoiding personal disasters because there are myriad ways to try to reduce them ranging from standard medical care to new technologies, which would require separate risk analyses depending on the measure taken.

Let us first consider when existence is held constant and then turn to when existence is variable. If the existence of an individual is unaffected and they avoid a personal disaster (e.g. they exist regardless and avoid an accident or illness), one possible mechanism of increased risk is that an individual in a better condition might do more harm or create more risk (e.g. by working productively on a dangerous technology). That said, there seem to be plenty of personal disasters whose aversion or alleviation would create relatively little risk when existence is constant. For example, consider a middle-aged or older human who avoids chronic joint pain or a devastating accidental house-flooding.<sup>18</sup> Another example is the prevention of crime that is not lethal but still personally disastrous. Similarly, when it comes to non-human animals, we can think of measures that seem to have relatively isolated effects and keep existence relatively constant. Such measures include preventing the abuse of animals who would exist at any rate, as well as curing and preventing painful non-lethal diseases among wild animals.<sup>19</sup>

Preventing death could result in the surviving individual causing harm, for example, if a surviving human purchases animal products.<sup>20</sup> And when we save a being who will contribute to the creation of new beings in the future, we enable the harm and risks that those future beings will face. If one is preventing one's own death (or a group is acting to prevent the death of its members) at the expense of risk to others, such as future generations, then there is the potentially morally problematic aspect of creating risk for others at least in part for one's own sake.

With regard to not preventing death, it is obvious that not trying to prevent death risks death as a consequence, which can obviously be considered a personal disaster. And there may be risks that come with the absence of a given individual.

If one in part addresses the occurrence of personal disasters by not creating new beings in the first place, personal disasters can be reduced in part because there is no risk imposed on or caused by such new beings. One could counter by saying that there are risks that come with non-creation such as negative societal effects of lower birth rates. However, as mentioned in section 5, one may think that omitting to create is not, or is to a lesser extent, morally problematic or irresponsible risk-imposition, so non-creation may be more acceptable from that perspective of the ethics of risk-imposition.

All in all, there seem to be plenty of ways to address personal disasters that create relatively little risk. This limited risk-imposition can be a consideration in favour of allocating attention and resources to personal disasters. Measures with little risk can, for example, keep existence relatively constant while improving health conditions or preventing accidents and crime. Addressing personal disasters by not creating the beings in the first place is, of course, an exceptionally effective way to avoid personal disaster for them and any individuals they would have harmed. And, as mentioned, such rather direct avoidance of personal disasters may be considered morally more weighty than the undesirable things that may exist if the being is not created.

## 6.2. *A lack of purported positive value being created*

<sup>18</sup> See e.g. the work of the Organisation for the Prevention of Intense Suffering.

<sup>19</sup> See e.g. Animal Ethics (n.d.) for information on vaccinating and treating wild animals.

<sup>20</sup> See e.g. Benatar and Wasserman (2015, pp. 78–112) on the harms humans cause humans and other animals.



This subsection is about new, purportedly positive value not being created.<sup>21</sup> A scenario *with* such purported positive value being created could involve large-scale space colonisation (to make room for all the purported positive value).

I would not use the terms ‘disaster’ and ‘risk’ for the lack of purported positive value being created. In my view, it does not qualify as a disaster or risk. There is nothing negative or bad about it. For example, if we consider a lifeless planet, it seems to me unreasonable to say that the state of the planet is a disaster, or that it would be a disaster if it remained like that. Similarly, if nothing bad might be added and the existence of the planet has no negative effects, there is no risk involved (and there is no threat). The lifeless planet is not something that needs to be addressed, in my view.

Regardless of whether one agrees with that, when it comes to questions such as whether the creation of purported positive value is important or would make things better, there is a difference between theoretical reasoning and how we should act and be in the real world. For example, if the world were free of ills and we could create a new community with no problems and risks and if everything is consented to by all the beings involved, then it would be no big deal, morally speaking. There would, by stipulation, be no risk and no opportunity cost in terms of ills that could have been reduced instead. But that hypothetical scenario is obviously very far from the real world. In real life, we need to consider serious opportunity costs and risks. Whenever there are risks and opportunity costs in terms of, for example, others’ extreme involuntary suffering, the question arises as to how much such suffering we should risk or accept in order to, in this case, contribute to the creation of purported positive value (e.g. Vinding, 2020, Chapter 3, 2022b). (In my view, the more such suffering is risked or accepted, the more morally problematic it is.)

The issue of opportunity costs and increasing other risks depends on which activity we are talking about. For example, if we are talking about an ongoing and long-term endeavour to achieve large-scale space colonisation in part by the development of new technologies, then it is easy to identify increased risk and opportunity cost in terms of severe suffering (e.g. risks for future beings and opportunity costs related to current beings in need). If, on the other hand, we are talking about an isolated practical intervention meant to increase the likelihood of vast positive value being created in the future, say, an intervention to promote global stability and avoid the most dangerous conflicts, then there may be much less risk created and less opportunity cost since the measure may coincide with measures that would be taken to reduce other risks and ills (see e.g. the interventions in sec. 4.2). In principle, the intervention could even be optimal when it comes to reducing extreme suffering and avoiding the risk of extreme suffering, even if it is aimed at the creation of positive value.

From an ethical risk perspective, we can note morally problematic features, roles, and distributions of risks and benefits. Suppose that we are talking about endeavours to colonise space and drastically increase the number of beings. Risk can be imposed on currently existing beings, future beings in the near term, or future beings in the farther future. Several factors that plausibly make risk-imposition more morally problematic (see sec. 5) are relevant: the risk of great harm; that the benefit would be a pure benefit; and that the individual being exposed to the risk is not making the decision, has not consented, and is not the one who would benefit. These factors seem to be present in our situation. That is, when it comes to subjecting current and near-term beings to risks in the process of trying to realise vast amounts of purported value, at least some current beings plausibly do not consent (and future beings cannot consent), the potential harms would be great (e.g. death and worse), and it would be for future pure positive value. Given that the positive value would not benefit the current or near-term beings, it would be for someone else.

---

<sup>21</sup> I speak in terms of value here, but one could instead speak about whatever it is that would not be created, such as pleasure or satisfied preferences.

When it comes to beings in the farther future, we should note both the extreme harm at risk, as well as other morally relevant aspects. I will start with the extreme harm at risk. There is the potential for ill-being to become more intense in the future compared to what is currently possible (e.g. due to new beings having a greater capacity for suffering or more effective torture methods) and for ill-being to occur on a much larger scale (e.g. due to space colonisation). So the risk of great harm in the far-future case becomes a risk of even more extreme harm on an even greater scale. That makes imposing such risks morally highly significant.

Moreover, when it comes to far future scenarios of which some may contain vast amounts of value and some may contain extreme suffering and ill-being, it is plausibly not the case that a given individual will be created and face the risk of extreme suffering and the chance of the positive well-being. Rather, the individuals in the favourable scenario are plausibly not the same as the individuals who would exist in the unfavourable scenario. The scenarios are alternative, very different, long-term futures, and hence the individuals in these respective scenarios would plausibly be different as well. If the unfavourable scenario materialises, the victims in that scenario would suffer while the purportedly positive lives and value would have been realised among other individuals in an alternative future that did not materialise.<sup>22</sup> So it is an unconsented subjection of risk of extreme harm for some individuals for the purportedly pure benefit of others (or for the sake of an impersonal value), which seems to be an especially morally problematic kind of risk-imposition (cf. Hansson, 2018).

The purportedly positive value may be taken to be personal (value *for* some individuals) or impersonal (just value, period, even if it is not value for anyone). For example, one may think that it is valuable for someone to experience pleasure or have their preferences satisfied. Impersonal value might be thought to be, for instance, scientific progress, there being meaning in the universe, or a civilisation's flourishing, which would all purportedly be good above and beyond what is good for individuals. Regardless of whether the purported positive value is personal or impersonal, there is the just-mentioned moral problem that the benefit would not be for the risk-exposed individuals. Still, it is relevant to distinguish personal and impersonal purported positive value because additional moral issues may arise, depending on which sort of value we are talking about. For example, suppose that the purportedly positive value is impersonal. There seems to be something morally problematic about accepting, causing, or risking others' severe suffering for individuals for purported positive impersonal value. The cost is tangible suffering that matters acutely for someone in exchange for the purported value of something above and beyond what matters for individuals. We can seemingly find morally problematic real-world examples of, for example, accepting misery and death among innocent, unconsenting victims for national glory. It seems challenging to explain why such behaviours would be morally permissible when the impersonal value is due to the properties of, say, all of human civilisation rather than a nation.

When considering the moral question of how to deal with the possibility that purportedly positive value would not be created, one needs to hold separate self-related features, such as the following: its giving someone meaning or purpose to think that they are part of a grand future or project, or the person's having a psychological need for hope or positivity, or finding a certain future fascinating. If such self-related reasons are part of why the person tries to ensure the creation of vast positive value, and if the person in effect exposes others (e.g. future beings) to risk, then we have the morally problematic role of being decision-maker plus beneficiary while others are exposed to risk (cf. Hansson, 2018).

The following are some conclusions about the practical issue of addressing the possibility that purported positive value will not be created (or that less such value will

<sup>22</sup> A similar point is made in Knutsson (2022).

be created). I would say that it is not a risk or disaster, and nothing that should be addressed, but let us set that opinion aside. It seems that someone who finds the creation of positive value in the future to be important would need to look for measures that, generally speaking, limit the risk of others' severe ill-being, especially such risks among unconsenting individuals who are not the beneficiaries. Similarly, the opportunity costs in terms of helping very badly off individuals should be kept low. Morally permissible or advisable measures would seem to largely coincide with measures to avoid s-risks (see e.g. measures related to s-risks mentioned in secs. 4.2 and 6.4) and personal disasters (secs. 6.1 and 6.4). Generally sensible measures such as preventing the worst global conflicts and improving political institutions seem advisable both for the creation of positive value and limiting risks of severe ill-being (see e.g. Baumann, 2020).

### 6.3. Extinction

When discussing the ethics of extinction, it is worth asking: Extinction of what, what sense of 'extinction', how would extinction occur, and which aspects of extinction would matter? For example, we can talk about the extinction of a current plant or non-human animal species, humans, or a kind of being that does not yet exist. According to *Merriam-Webster*, meanings of 'extinction' include "the act of making extinct", "the condition or fact of being extinct", and "the process of becoming extinct".<sup>23</sup> Our evaluations of these different senses of 'extinction' might differ substantially. For instance, one can hold that the process of becoming extinct involves disvalue in so far as it involves suffering, while the condition of being extinct is not bad. Relatedly, a voluntary extinction that occurs by a process that is free from ill-being might morally be very different from an involuntary extinction (e.g. Bergström, 1978/2022, Section VIII). Finally, there are different reasons why one may hold that extinction matters, such as the instrumental effects of the extinction; the death of existing beings; the fact that there will not be new, happy individuals in the future; or the loss or absence of an impersonal value such as scientific achievement or the impersonal purportedly positive value of a species. (One could also hold that extinction does not matter morally.)

With these distinctions in mind, let us continue our ethical analysis. If extinction is, in part, thought to be bad because of the ill-being or death of existing individuals, then that consideration would be reducible to personal disasters (which we already dealt with in sec. 6.1), provided that death counts as a personal disaster. The number of personal disasters occurring during extinction may be small or large. If the last individual of a species dies, it is only one personal disaster (setting aside that individuals of another species might be affected). If billions die, it is billions of personal disasters. Either way, when it comes to the consideration of any death and ill-being involved in extinction, it makes sense to think of it as a certain number of personal disasters and to treat the death and ill-being accordingly.<sup>24</sup>

Similarly, if extinction is claimed to matter because purported positive value would not be created in the future, we have already dealt with that consideration too (sec. 6.2).

<sup>23</sup> <https://www.merriam-webster.com/dictionary/extinction>. See also Torres (2023, Chapter 7, n.d.).

<sup>24</sup> When it comes to any death and ill-being involved in extinction, plenty could be said about how much of a priority they should be and how special they are compared to other instances of death and ill-being. First, if measures to prevent involuntary extinction are not the most effective at preventing involuntary death, then there is an opportunity cost in terms of preventable deaths associated with focusing specifically on extinction. Second, extinction has the feature that survivors are not around to mourn the deceased, so that would be a way in which such deaths cause less ill-being compared to other deaths. Third, in terms of the number of deaths and the amount of ill-being on Earth, the death of everyone may not be so different from the usual state of affairs. For example, given how many deaths normally occur among all species, a sudden mass death may not be that different (Horta, 2017).

The same moral and practical conclusions apply to this concern when it relates to extinction.

Is there anything noteworthy about extinction that we have not already covered? Extinction could be considered important because there already is a value that would be lost if extinction occurs. An obvious example is familiar from environmental ethics and concerns the value of a species itself.<sup>25</sup> Other potential examples might be that an existing culture, tradition, or civilisation would no longer exist, or perhaps that knowledge, meaning, or consciousness would no longer exist. I do not think that any such things have impersonal positive value for their own sake, but it would be a large project to argue that here.<sup>26</sup>

As in the previous subsection, an issue that arises for purported impersonal values is how large risks and tradeoffs they justify in terms of others' extreme ill-being.<sup>27</sup> Even if there is some moral reason or value connected to, for example, a species above and beyond individuals, that would need to be weighed against any risks and opportunity costs involving extreme suffering for individuals.

There are some risks associated with preventing death (as mentioned in sec. 6.1), as well as with the perpetuation of a species. For example, by sticking around and enabling the creation of new generations, we enable the ill-being, death, and risks that those future beings will face (Bergström, 1978/2022, Section VII–VIII).

Self-interest can be an interesting feature of measures to prevent extinction. If extinction prevention is pursued in part for one's own benefit (e.g. because of one's own projects or to avoid one's own death), then there is the morally problematic role of potentially exposing future beings to risk of great harm in part for one's own benefit (cf. Hansson, 2018). This seems to be the case also if a group (e.g. humanity) acts to prevent its own extinction in part for the benefit of its current members while thereby exposing future individuals to the risk of great harm. There seems to be something selfish and inconsiderate about trying to ensure one's own or one's group's continued existence if it happens at the expense of an increased risk of harm for future beings (compare the example in sec. 5 of risking someone else's life to make a fortune for oneself).

I will end with some practical upshots regarding extinction. If we set aside some instrumental effects of extinction, then involuntary extinction via mass death should in itself seemingly be treated roughly as any involuntary mass death event (including any suffering that would be involved). For example, ten million deaths and ten billion deaths differ in quantity and the difference matters even though extinction is not a separate moral concern in itself. The remaining practical upshots seem similar to those regarding positive value in the previous subsection, for example, keeping down risks and costs in terms of unconsenting individuals' extreme ill-being. One may look for measures that both prevent extinction and reduce the risk of terrible unconsented ill-being. Such measures may include moral improvement, certain artificial intelligence safety mechanisms that also reduce s-risks, and the avoidance of great wars (see the measures in secs. 4.2 and 6.4).

<sup>25</sup> Even if we would grant that extinction matters beyond instrumental effects and the fate of individuals, there is the question of which species or kinds of beings matter more or less and why. If the extinction of species matters in itself, one may expect the primary concern to be, for example, the extinction of many species of plants and insects. If *Homo sapiens* or some future species is claimed to be much more important in itself than those species, then there is the challenge of explaining why. There are also counterarguments, for example, if higher intelligence is claimed to make humans more important or valuable, then more intelligent humans would seemingly be more important than other humans, and a more intelligent species would seemingly be more important or valuable still.

<sup>26</sup> See e.g. Bykvist (2014) and the literature on sentientism or sentiocentrism.

<sup>27</sup> Related ideas on weighing extinction or species preservation against suffering can be found in Tännsjö (2023) and Fanciullo (2024).

#### 6.4. *Risks of astronomical suffering*

Given a wide enough understanding of the nature of personal disasters, s-risks are merely very large numbers of personal disasters. In this subsection, I try to avoid repeating the points from the subsection on personal disasters (sec. 6.1), and instead focus on what is special about s-risks.

That said, a few points from the preceding subsections are also relevant to s-risks, so I will briefly mention them now for completeness. First, ill-being may become more intense in the future and it may occur on a much larger scale, so there is a risk of more extreme harm on an unprecedented scale. Second, in so far as the future beings who would have terrible ill-being do not consent and are not the ones with the chance of benefit, we find more features that tend to make risk-imposition morally problematic, especially if the new existences are enabled for the sake of pure benefits and in part for the decision-makers' benefit. Third, when it comes to measures to reduce s-risks, there can be tradeoffs, opportunity costs, and other risks that may be created. How much risk is created and how great the opportunity costs are depends on the measure in question. For example, improving political institutions may create relatively little risk and have rather favourable opportunity costs also from the perspective of positive value being created in the future, avoiding extinction, and reducing personal disasters. Although safety mechanisms for artificial intelligence systems that prevent s-risks may also be beneficial from the perspective of extinction and positive value in the future, some could perhaps claim that there is some opportunity cost because if positive value in the future and extinction has enough moral weight, the s-risk measure is suboptimal. If there are such tradeoffs, we get into the moral balancing of s-risks versus a lack of positive value being created or extinction, as we have touched upon in previous subsections.

Let us briefly consider potentially ethically problematic aspects of s-risk reduction. The aim of reducing s-risks does not involve increasing risks of harm for the sake of pure benefits, so it does not have that potentially morally problematic feature. Reducing s-risks could be accused of being partly self-serving in so far as the measure is also meant to prevent us from enduring extreme suffering. However, that does not mean that it exposes others to risk for one's own benefit. It is rather that current beings, including potentially the decision-makers, are among those who would face less risk. In part because s-risk reduction does not involve the creation of new beings (who would face risks and cannot consent), I expect that a main moral objection to s-risk reduction would not be that such reduction involves irresponsible risk-imposition.<sup>28</sup> Rather, I picture that a primary objection to s-risk reduction would concern opportunity costs and tradeoffs. In other words, I picture that a main objection would be that we should try to accomplish something else that should be an even higher priority.

A potential tradeoff that we have talked less about so far is s-risks versus smaller-scale personal disasters that occur nowadays. It may well be the case that the most effective way to address one of them is not the most effective way to address the other. Someone might morally object to focusing on reducing s-risks over (other) personal disasters even if a focus on s-risks is the most effective way to reduce severe suffering. Such a moral objection could, for example, appeal to the immediacy of a current being's plight and the moral importance of exhibiting empathy in the face of that plight (cf. Slote, 2007, 2010). (To be clear, I do not endorse that moral objection against a strong focus on s-risks, although I do think that we should devote significant resources toward the reduction of ongoing suffering.) In addition to moral reasons, one might find that

---

<sup>28</sup> One can imagine attempts to reduce s-risks that involve the creation of new beings (e.g. perpetuating some life on Earth to reduce s-risks created elsewhere in the universe). But enabling the creation of new beings does not seem central to s-risk reduction in the way that it does to efforts to prevent extinction and create vast amounts of positive value in the future.

addressing s-risks is less motivating than addressing current personal disasters. If one has doubts about a focus on s-risks for moral or motivational reasons, one could try to address both s-risks and current personal disasters at the same time. Several of the interventions to reduce s-risks mentioned in section 4.2 plausibly benefit both s-risks and other personal disasters, for example, improving political institutions, ensuring that there is moral concern for all sentient beings, and reducing the influence of malevolent agents. Still, one may consider such measures too indirect or slow given the many individuals being in desperate need now. We can ask whether there are ways to concretely help current victims in a way that creates little risk, has little opportunity cost, and also reduces, say, a key risk factor for s-risk. I will get back to that shortly in my discussion of conflicts.

In the remainder of my text, I will focus on two measures, namely moral improvement and the prevention of the most dangerous conflicts. I will focus on these two measures to try to say something useful about them, although there are plenty of other measures to reduce s-risks (see sec. 4.2).

Moral improvement (including the reduction of moral decline) could occur via, for example, moral bioenhancement or societal changes that facilitate moral development and learning. Which morality-related characteristics are desirable from an s-risk perspective? As has been mentioned in the literature (sec. 4.2), it seems desirable to reduce Machiavellianism, narcissism, psychopathy, and sadism (Althaus & Baumann, 2020), as well as retributive tendencies (e.g. DiGiovanni, 2023). Similarly, aggression seems undesirable. Having concern for all sentient beings makes sense (Baumann, 2022, pp. 57–62); otherwise, vast suffering may occur among a neglected group of beings.<sup>29</sup> Not being tribal also seems beneficial (Baumann, 2022, p. 67). Vinding (forthcoming, Chapter 9) lists plenty of virtues to cultivate to reduce suffering, including integrity, honesty, curiosity, agency, self-discipline, and being cooperative. This illustrates how wide-ranging and diverse the set of desirable traits may be.

I will next spend some extra time on a few potentially desirable characteristics. The most obvious one is perhaps compassion. Compassion is prominent in Schopenhauer's philosophy. He writes, among other things:

I wish to propose the following rule: for every human being with whom one comes into contact,... focus alone on his suffering, his distress, his fear, his pain – then you will always feel kinship with him, sympathize with him and ... sense that compassion for him... (Schopenhauer, 1851/2015, p. 184)

According to Fox (2023, Note 2), Schopenhauer identifies compassion “specifically with a direct desire to relieve the *suffering* of others” and understands compassion “in a way that tracks the literal meaning of the German term *Mitleid*: suffering with.” This sense of compassion may be especially important to s-risk reduction, for example, as a foundation for being concerned about s-risk reduction, prioritising it, paying attention to it, and weighing considerations appropriately when making decisions.

Another relevant characteristic is to have a sense of equality or an aversion to inequality or unfairness (e.g. Horta, 2014; Vinding, 2023). Similarly to compassion, this can lead to a concern for those who are worse off compared to others.

In addition, characteristics such as the following may be beneficial for reducing s-risks: contentment, satisfaction, a tendency towards peace of mind, and an appreciation of serene and untroubled states of the world. The forms of these characteristics that I have in mind do not involve being unconcerned or passive. Rather, I have in mind characteristics such as a tendency towards sufficiency in the sense of, for instance, having the attitude that one's life is going sufficiently well as long as one avoids various severe ills; that there are enough people and purportedly valuable objects; and that a planet free

<sup>29</sup> Although see e.g. Vinding (2018).

of sentient beings is enough in its own right. Such characteristics may lead to less of a drive to maximise, create beings, and increase the intensity and amount of pleasure. In that way, s-risks may be reduced.

The final characteristic I wish to talk about is caution. For s-risks, caution may be more or less important depending on what it concerns. Still, if we do wish to avoid s-risks, it would be helpful were agents generally more cautious.

Let us move on from moral improvement and turn to conflict. Perhaps the most prominent risk factor for s-risks is conflict, in particular, conflicts that risk resulting in astronomical suffering (e.g. an escalating war between global powers with potent technology). Given the importance of such conflicts for s-risks, and assuming that one may also want to do something for beings who are very badly off nowadays, one can look for ways to reduce personal disasters that are also broadly and perhaps robustly beneficial, including for s-risk (and, in particular, for such conflicts). For example, perhaps reduced oppression and poverty now leads to less dangerous conflicts later. Another example would be helping current victims of wars of aggression or revenge in a way that increases the victims' resiliency and defensive capacity. This may also discourage future wars of aggression or revenge, for instance, by making such wars less rewarding for the aggressor and contributing to norms of standing up for the victims of such wars and norms against aggression and retribution. In addition, such interventions to reduce oppression or help victims of war seem favourable from perspectives concerned with avoiding extinction and creating positive value.

## 7. Conclusions

There are ways to address personal disasters that seem to create relatively little risk, for example, when an individual exists in any event and intense distress is reduced. This low risk can be seen as an advantage of such interventions. One can also effectively prevent personal disasters by not creating new beings in the first place. If extinction occurs via involuntary death, then that should reasonably be addressed as other cases of involuntary death. Beyond that, the practical conclusions about extinction seem similar to those about the possibility that purported positive value will not be created (or that less such value will be created). Measures to prevent extinction and ensure that positive value is created should be limited to measures that also reduce, or at least do not increase, personal disasters and risks of astronomical suffering (s-risks). A variety of interventions to reduce s-risks have been mentioned in the literature, such as reducing the influence of malevolent individuals, avoiding large-scale conflicts, and working on safety mechanisms for artificial intelligence systems that are aimed at preventing s-risks. I have paid special attention to moral improvement and preventing the most dangerous conflicts. Regarding moral improvement, there is a long list of seemingly desirable characteristics, including compassion; caring about equality and fairness; contentment and an appreciation of what it is; and caution. Conflict is a top risk factor for s-risks and, with that in mind, I have suggested the following intervention as an intervention that may be widely acceptable and beneficial from a range of perspectives: helping current victims of wars of aggression or revenge in a way that both reduces their ill-being and improves their resilience and defensive capacity. That would address current personal disasters and it could help prevent future conflicts, which seems beneficial for s-risks, extinction risk, and the creation of purported positive value in the future.

**Funding:** This research received no funding of any kind.

**Acknowledgments:** I am grateful to Teo Ajantaival and Tobias Baumann for their comments on earlier drafts. I also wish to direct a special thanks to Magnus Vinding for his very helpful and skilled comments.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

- Althaus, D., & Baumann, T. (2020). *Reducing long-term risks from malevolent actors*. Center on Long-Term Risk. <https://longtermrisk.org/reducing-long-term-risks-from-malevolent-actors/>
- Animal Ethics. (n.d.). *Vaccinating and healing sick animals*. Retrieved 20 December 2024, from <https://www.animal-ethics.org/vaccinating-healing-sick-injured-animals/>
- Baumann, T. (2017a). *S-risks: An introduction*. Center for Reducing Suffering. <https://centerforreducingsuffering.org/research/intro/>
- Baumann, T. (2017b). *Focus areas of worst-case AI safety*. Reducing Risks of Future Suffering. <https://s-risks.org/focus-areas-of-worst-case-ai-safety/>
- Baumann, T. (2019). *Risk factors for s-risks*. Center for Reducing Suffering. <https://centerforreducingsuffering.org/research/risk-factors-for-s-risks/>
- Baumann, T. (2020). *Common ground for longtermists*. Center for Reducing Suffering. <https://centerforreducingsuffering.org/research/common-ground-for-longtermists/>
- Baumann, T. (2021). *How can we reduce s-risks?* Center for Reducing Suffering. <https://centerforreducingsuffering.org/research/how-can-we-reduce-s-risks/>
- Baumann, T. (2022). *Avoiding the Worst: How to Prevent a Moral Catastrophe*. [https://centerforreducingsuffering.org/wp-content/uploads/2022/10/Avoiding\\_The\\_Worst\\_final.pdf](https://centerforreducingsuffering.org/wp-content/uploads/2022/10/Avoiding_The_Worst_final.pdf)
- Beebe, H. (2004). Causing and Nothingness. In L. A. Paul, E. J. Hall, & J. Collins (Eds.), *Causation and Counterfactuals* (pp. 291–308). Cambridge, MA, USA: MIT Press. <https://philarchive.org/rec/BEECAN>
- Benatar, D., & Wasserman, D. T. (2015). *Debating procreation: Is it wrong to reproduce?* Oxford University Press.
- Bergström, L. (2022). *The consequences of pessimism* (S. Knutsson, Trans.). <https://www.simonknutsson.com/the-consequences-of-pessimism/>. (Original work published 1978)
- Boholm, M., Möller, N., & Hansson, S. O. (2015). The Concepts of Risk, Safety, and Security: Applications in Everyday Language. *Risk Analysis*, 36(2), 320–338. <https://doi.org/10.1111/risa.12464>
- Bykvist, K. (2014). Killing and extinction. In S. Luper (Ed.), *The Cambridge Companion to Life and Death* (pp. 316–329). Cambridge University Press. <https://doi.org/10.1017/CCO9781139149129.023>
- Canadian Red Cross. (n.d.). *Personal Disasters Infographic*. Retrieved 28 October 2024, from <https://www.redcross.ca/how-we-help/emergencies-and-disasters-in-canada/be-ready-emergency-preparedness-and-recovery/personal-disaster-infographic>



- Daniel, M. (2017). *S-risks: Why they are the worst existential risks, and how to prevent them* (EAG Boston 2017). Center on Long-Term Risk. <https://longtermrisk.org/s-risks-talk-eag-boston-2017/>
- DiGiovanni, A. (2023). *Beginner's guide to reducing s-risks*. Center on Long-Term Risk. <https://longtermrisk.org/beginners-guide-to-reducing-s-risks/>
- Fanciullo, J. (2024). Why prevent human extinction? *Philosophy and Phenomenological Research*, 109(2), 650–662. <https://doi.org/10.1111/phpr.13066>
- Fox, J. I. (2023). Does Schopenhauer accept any positive pleasures? *European Journal of Philosophy*, 31(4), 902–913. <https://doi.org/10.1111/ejop.12830>
- Gloor, L. (2016). *Suffering-Focused AI Safety: In Favor of "Fail-Safe" Measures*. Center on Long-Term Risk. <https://longtermrisk.org/suffering-focused-ai-safety/>
- Gloor, L. (2018). Cause prioritization for downside-focused value systems. *Center on Long-Term Risk*. <https://longtermrisk.org/cause-prioritization-downside-focused-value-systems/>
- Hansen, S. F., von Krauss, M. K., & Tickner, J. A. (2008). The precautionary principle and risk-risk tradeoffs. *Journal of Risk Research*, 11(4), 423–464. <https://doi.org/10.1080/13669870801967192>
- Hansson, S. O. (2000). *Seven Myths of Risk*. Talk at the conference Stockholm thirty years on. Progress achieved and challenges ahead in international environmental co-operation. Swedish Ministry of the Environment, June 17–18, 2000. <https://people.kth.se/~soh/SevenMythsTalk.pdf>
- Hansson, S. O. (2003). Ethical Criteria of Risk Acceptance. *Erkenntnis*, 59(3), 291–309.
- Hansson, S. O. (2010). Risk: Objective or subjective, facts or values. *Journal of Risk Research*, 13(2), 231–238. <https://doi.org/10.1080/13669870903126226>
- Hansson, S. O. (2017). Five caveats for risk-risk analysis. *Journal of Risk Research*, 20(8), 984–987. <https://doi.org/10.1080/13669877.2016.1147493>
- Hansson, S. O. (2018). How to Perform an Ethical Risk Analysis (eRA). *Risk Analysis*, 38(9), 1820–1829. <https://doi.org/10.1111/risa.12978>
- Hansson, S. O. (2023). Risk. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Summer 2023). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/sum2023/entries/risk/>
- Hayenhjelm, M., & Wolff, J. (2012). The Moral Problem of Risk Impositions: A Survey of the Literature. *European Journal of Philosophy*, 20(S1), E26–E51. <https://doi.org/10.1111/j.1468-0378.2011.00482.x>

- Hermansson, H., & Hansson, S. O. (2007). A Three-Party Model Tool for Ethical Risk Analysis. *Risk Management*, 9(3), 129–144.
- Horta, O. (2014). Egalitarianism and Animals. *Between the Species*, 19(1). <https://digitalcommons.calpoly.edu/bts/vol19/iss1/5>
- Horta, O. (2017). Animal Suffering in Nature: The Case for Intervention. *Environmental Ethics*, 39(3), 261–279. <https://doi.org/10.5840/enviroethics201739320>
- Hsu, E. L. (2019). Must disasters be rapidly occurring? The case for an expanded temporal typology of disasters. *Time & Society*, 28(3), 904–921. <https://doi.org/10.1177/0961463X17701956>
- Knutsson, S. (2022). *Disvaluable acts and traits from now to the far future*. Unpublished manuscript. An earlier version of this manuscript is titled *Pessimism about the value of the future and the welfare of present and future beings based on their acts and traits* and is available at <https://www.simonknutsson.com/pessimism-value-future-welfare-acts-traits/>
- Knutsson, S., & Munthe, C. (2017). A Virtue of Precaution Regarding the Moral Status of Animals with Uncertain Sentience. *Journal of Agricultural and Environmental Ethics*, 30(2), 213–224. <https://doi.org/10.1007/s10806-017-9662-y>
- Knutsson, S., & Vinding, M. (2024). *Introduction to suffering-focused ethics*. <https://centerforreducingsuffering.org/research/introduction-to-suffering-focused-ethics/>
- Leslie, J. (1996). *The end of the world: The science and ethics of human extinction*. Routledge.
- Magnusson, E. (2022). On Risk-Based Arguments for Anti-natalism. *The Journal of Value Inquiry*, 56, 101–117.
- Munthe, C. (2011). *The price of precaution and the ethics of risk*. Springer. <https://doi.org/10.1007/978-94-007-1330-7>
- O'Mathúna, D. P., & Gordijn, B. (2018). Conceptualizing and Assessing Disasters: An Introduction. In D. P. O'Mathúna, V. Dranseika, & B. Gordijn (Eds.), *Disasters: Core Concepts and Ethical Theories* (pp. 1–9). Springer International Publishing. [https://doi.org/10.1007/978-3-319-92722-0\\_1](https://doi.org/10.1007/978-3-319-92722-0_1)
- Raphael, B. (1981). Personal disaster. *Australian and New Zealand Journal of Psychiatry*, 15(3), 183–198. <https://doi.org/10.3109/00048678109159434>
- Rulli, T. (2024). Effective Altruists Need Not Be Pronatalist Longtermists. *Public Affairs Quarterly*, 38(1), 22–44. <https://doi.org/10.5406/21520542.38.1.03>

- Sandin, P. (2018). Conceptualizations of Disasters in Philosophy. In D. P. O'Mathúna, V. Dranseika, & B. Gordijn (Eds.), *Disasters: Core Concepts and Ethical Theories* (pp. 13–26). Springer International Publishing. [https://doi.org/10.1007/978-3-319-92722-0\\_2](https://doi.org/10.1007/978-3-319-92722-0_2)
- Sartorio, C. (2009). Causation and Ethics. In H. Beebe, C. Hitchcock, & P. Menzies (Eds.), *The Oxford Handbook of Causation* (pp. 575–591). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199279739.003.0027>
- Schopenhauer, A. (2015). *Parerga and Paralipomena: Short Philosophical Essays* (A. Del Caro & C. Janaway, Eds.; A. Del Caro, Trans.; Vol. 2). Cambridge University Press. <https://doi.org/10.1017/CBO9781139016889> (Original work published 1851)
- Shiffrin, S. V. (1999). Wrongful life, procreative responsibility, and the significance of harm. *Legal Theory*, 5(2), 117–148. <https://doi.org/10.1017/S1352325299052015>
- Shiffrin, S. V. (2012). Harm and its moral significance. *Legal Theory*, 18(3), 357–398. <https://doi.org/10.1017/S1352325212000080>
- Slote, M. (2007). *The ethics of care and empathy*. Routledge.
- Slote, M. (2010). Famine, Affluence, and Empathy. In David. Boonin & G. Oddie (Eds.), *What's wrong?: Applied ethicists and their critics* (2nd ed., pp. 146–154). Oxford University Press.
- Sotala, K., & Gloor, L. (2017). Superintelligence As a Cause or Cure For Risks of Astronomical Suffering. *Informatica*, 41(4), Article 4.
- Tännsjö, T. (2023). Does It Matter if We Go Extinct? In T. Tännsjö (Ed.), *From Despotism to Democracy: How a World Government Can Save Humanity* (pp. 59–74). Springer Nature. [https://doi.org/10.1007/978-981-99-5559-6\\_5](https://doi.org/10.1007/978-981-99-5559-6_5)
- Torres, É. P. (2023). *Human Extinction: A History of the Science and Ethics of Annihilation*. Routledge.
- Torres, É. P. (n.d.). *On the Extinction of Humanity*. Unpublished manuscript. [https://f2fa52f4-c77c-4e5d-8cbb-4747dd34696e.filesusr.com/ugd/d9aaad\\_27bbb0fdac8b4725b1f8cb2b673e0117.pdf](https://f2fa52f4-c77c-4e5d-8cbb-4747dd34696e.filesusr.com/ugd/d9aaad_27bbb0fdac8b4725b1f8cb2b673e0117.pdf)
- Vinding, M. (2018). Moral Circle Expansion Might Increase Future Suffering. *Magnus Vinding*. <https://magnusvinding.com/2018/09/04/moral-circle-expansion-might-increase-future-suffering/>
- Vinding, M. (2020). *Suffering-focused ethics: Defense and implications*. Ratio Ethica. <https://magnusvinding.com/2020/05/31/suffering-focused-ethics-defense-and-implications/>
- Vinding, M. (2022a). *Reasoned Politics*. Ratio Ethica.
- Vinding, M. (2022b). Reply to Chappell's "Rethinking the Asymmetry". *Center for Reducing Suffering*. <https://centerforreducingsuffering.org/reply-to-chappells-rethinking-the-asymmetry/>

Vinding, M. (2023). *A convergence of moral motivations*. <https://magnusvinding.com/2023/06/15/a-convergence-of-moral-motivations/>

Vinding, M. (forthcoming). *Compassionate Purpose*. <https://magnusvinding.com/books/#compassionate-purpose>

Wolf, C. (2004). O repugnance, where is thy sting? In T. Tännsjö & J. Ryberg (Eds.), *The Repugnant Conclusion* (pp. 61–80). Springer Netherlands. <https://faculty.sites.iastate.edu/jwcwolf/files/inline-files/Wolf%202004%20O%20Repugnance%20%28galley%29.pdf>