## 1

#### CW: NON-GRAPHIC MENTIONS OF THE WORD SUICIDE

#### Interpretation: Debaters must provide a content warning when using the word suicide- to clarify this is not a reps K of the word but rather a reason why their presentation is bad

#### Violation: cx when explaing intutions

#### Standards

#### 1] Safety and Psychic violence- suicide is a traumatizing and psychologically violent topic- I don’t need to explain why- not allowing debaters to prepare themselves mentally to read the word and not get triggered creates an unsafe model of debate that has a high risk of causing traumatizing thoughts and flashbacks. Saftey o/w and is an independent voter- debate has a prior obligation to be a safe space for high school competitors- pre-req to accessing the debate space

#### 2] I’ll pre-empt generic “content warning bad” arguments a) we only need to win why a content warning is good in this particular case b) circuit norms prove that suicide is a sensitive topic that should be addressed with a warning c) creates a culpability DA to people facing psychic violence to face the effects of people abusing trigger warnings to skirt clash- it’s not our fault d) context doesn’t solve because once a person sees the word the context is irrelevant and they are already triggered

#### Competing interps- a) you can’t be reasonably violent b) reasonability forces the judge to determine whether personal violence was small enough to go back to substance

#### DTD- a) I’ve already been triggered, and the round is irredeemable b) only way to create an incentive structure for people to change their norms

#### No RVIs- a) you shouldn’t win for winning you don’t need a content warning- its illogical b) baiting- debaters just read abusive/violent practices to win on the RVI- chills debaters from calling out bad practices

## 2

#### Agents must be practical reasoners:

#### First, inescapability – the exercise of practical rationality requires that one regards it as intrinsically good – that justifies a right to freedom.

Wood [Allen W. Wood, (Stanford University, California) "Kantian Ethics" Cambridge University Press, 2007, https://www.cambridge.org/core/books/kantian-ethics/769B8CD9FCC74DB6870189AE1645FAC8, DOA:8-12-2020 // WWBW]//rct st

Kant holds that the most basic act through which people exercise their practical rationality is that of setting an end (G 4:437). To set an end is, analytically, to subject yourself to the hypothetical imperative that you should take the necessary means to the end you have set (G 4:417). This is the claim that you rationally ought to do something whether or not you are at the moment inclined to do it. It represents the action of applying that means as good (G 4:414) – in the sense of “good” that Kant explicates as: what is required by reason independently of inclination (G 4:413). Kant correctly infers that any being which sets itself ends is committed to regarding its end as good in this sense, and also to regarding the goodness of its end as what also makes application of the means good – that is, rationally required independently of any inclination to apply it. The act of setting an end, therefore, must be taken as committing you to represent some other act (the act of applying the means) as good. In doing all this, however, the rational being must also necessarily regard its own rational capacities as authoritative for what is good in general. For it treats these capacities as capable of determining which ends are good, and at the same time as grounding the goodness of the means taken toward those good ends. But to regard one’s capacities in this way is also to take a certain attitude toward oneself as the being that has and exercises those capacities. It is to esteem oneself – and also to esteem the correct exercise of one’s rational capacities in determining what is good both as an end and as a means to it. One’s other capacities, such as those needed to perform the action that is good as a means, are also regarded as good as means. But that capacity through which we can represent the very idea of something as good both as end and as means is not represented merely as the object of a contingent inclination, nor is it represented as good only as a means. It must be esteemed as unconditionally good, as an end in itself. To find this value in oneself is not at all the same as thinking of oneself as a good person. Even those who misuse their rational capacities are committed to esteeming themselves as possessing rational nature. It also does not imply that a more intelligent person (in that sense, more “rational”) is “better” than a less intelligent one. The self-esteem involved in setting an end applies to any being capable of setting an end at all, irrespective of the cleverness or even the morality of the end setting. Kant’s argument supports the conclusion, to which he adheres with admirable consistency throughout his writings, that all rational beings, clever or stupid, even good or evil, have equal (absolute) worth as ends in themselves. For Kantian ethics the rational nature in every person is an end in itself whether the person is morally good or bad.

#### Second, value theory – the existence of extrinsic goodness requires unconditional human worth.

Korsgaard (Christine M., “Two Distinctions in Goodness,” The Philosophical Review Vol. 92, No. 2 (Apr., 1983), pp. 169-195, JSTOR) OS \*bracketed for gen lang\* //rct st

The argument shows how Kant's idea of justification works. It can be read as a kind of regress upon the conditions, starting from an important assumption. The assumption is that when a rational being makes a choice or undertakes an action, he or she [they] supposes the object to be good, and its pursuit to be justified. At least, if there is a categorical imperative there must be objectively good ends, for then there are necessary actions and so necessary ends (G 45-46/427-428 and Doctrine of Virtue 43-44/384-385). In order for there to be any objectively good ends, however, there must be something that is unconditionally good and so can serve as a sufficient condition of their goodness. Kant considers what this might be: it cannot be an object of inclination, for those have only a conditional worth, "for if the inclinations and the needs founded on them did not exist, their object would be without worth" (G 46/428). It cannot be the inclinations themselves because a rational being would rather be free from them. Nor can it be external things, which serve only as means. So, Kant asserts, the unconditionally valuable thing must be "humanity" or "rational nature," which he defines as "the power set to an end" (G 56/437 and DV 51/392). Kant explains that regarding your existence as a rational being as an end in itself is a "subjective principle of human action." By this I understand him to mean that we must regard ourselves as capable of conferring value upon the objects of our choice, the ends that we set, because we must regard our ends as good. But since "every other rational being thinks of his existence by the same rational ground which holds also for myself' (G 47/429), we must regard others as capable of conferring value by reason of their rational choices and so also as ends in themselves. Treating another as an end in itself thus involves making that person's ends as far as possible your own (G 49/430). The ends that are chosen by any rational being, possessed of the humanity or rational nature that is fully realized in a good will, take on the status of objective goods. They are not intrinsically valuable, but they are objectively valuable in the sense that every rational being has a reason to promote or realize them. For this reason it is our duty to promote the happiness of others-the ends that they choose-and, in general, to make the highest good our end.

#### Third, practical reason – ethical principles must be derived from the structure of reason:

#### [1] Regress – we can always ask why we should follow a theory, so they aren’t binding because they don’t have a starting point. Practical reason solves – When we ask why we should follow reason, we demand a reason, which concedes to the authority of reason itself, so it’s the only thing we can follow

#### [2] Action Theory – every action can be broken down to infinite amounts of movements, i.e. me moving my arm can be broken down to the infinite moments of every state my arm is in. Only reason can unify these movements because we use practical reason to achieve our goals, means all actions collapse to reason

#### Fourth, epistemology – ethics must begin a priori, meaning they can’t be derived from our experience.

#### [A] Representations of space – we can only access our experiences if we can interpret the space around us, but that requires the a priori. Thinking of the absence of space is impossible – we can think of empty space but never the lack of space itself. Imagining space through a priori thoughts is the only way we can even begin to have a conception of interpreting experience; we need to be able to construct space through our minds.

#### [B] Separateness – if space is based on experience, it must be formed from objects separate to us outside of our reasoning abilities. But to represent objects as separate from us, we would already need to assume space exists in the first place to have a concept of “separateness,” so to represent space as something separate from us would be incoherent.

#### [C] Uncertainty – every person has different experiences so we can’t have a unified perspective on what is good if we each have different conceptions of it – even if we can roughly aggregate it’s not enough because there’ll always be a case when it fails so the framework o/w on probability.

#### [D] Is/Ought Gap – experience in the phenomenal world only tells us what is, not what ought to be. But it’s impossible to derive an ought from descriptive premises, so there needs to be additional a priori premises within the noumenal world to make a moral theory.

#### Practical reason means we all have a unified perspective: What can be justified to me can be justified to everyone who is a practical reasoner. If I can conclude that 2+2 is 4, then I understand not only that I know 2+2 is 4, but that everyone around me can arrive at the same conclusion. These things are temporally consistent: I know that me adding two numbers now and taking that sum will not result in me adding the same two numbers in the future and getting a different sum. Our unified perspective does not change but rather stays consistent.

#### But, willing an action that violates the freedom of others is a contradiction: If I decide to kill someone, that action is not universalizable because that would justify other people killing me too. If I die, I cannot exercise my freedom to kill someone else. This is a contradiction: I both justify extending my freedom to kill others and limiting my own freedom.

#### Thus, the standard is respecting freedom.

## Offense

#### Negate:

#### Acquisition of property can never be unjust – to create rights violations, there must already be an owner of the property being violated, but that presupposes its appropriation by another entity.

Feser 1, (Edward Feser, 1-1-2005, accessed on 12-15-2021, Cambridge University Press, "THERE IS NO SUCH THING AS AN UNJUST INITIAL ACQUISITION | Social Philosophy and Policy | Cambridge Core", Edward C. Feser is an American philosopher. He is an Associate Professor of Philosophy at Pasadena City College in Pasadena, California. [https://www.cambridge.org/core/journals/social-philosophy-and-policy/article/abs/there-is-no-such-thing-as-an-unjust-initial-acquisition/5C744D6D5C525E711EC75F75BF7109D1)[brackets](https://www.cambridge.org/core/journals/social-philosophy-and-policy/article/abs/there-is-no-such-thing-as-an-unjust-initial-acquisition/5C744D6D5C525E711EC75F75BF7109D1)%5bbrackets) for gen lang]//phs st

There is a serious difficulty with this criticism of Nozick, however. It is just this: There is no such thing as an unjust initial acquisition of resources; therefore, there is no case to be made for redistributive taxation on the basis of alleged injustices in initial acquisition. This is, to be sure, a bold claim. Moreover, in making it, I contradict not only Nozick’s critics, but Nozick himself, who clearly thinks it is at least possible for there to be injustices in acquisition, whether or not there have in fact been any (or, more realistically, whether or not there have been enough such injustices to justify continual redistributive taxation for the purposes of rectifying them). But here is a case where Nozick has, I think, been too generous to the other side. Rather than attempt —unsatisfactorily, in the view of his critics—to meet the challenge to show that initial acquisition has not in general been unjust, he ought instead to have insisted that there is no such challenge to be met in the first place. Giving what I shall call “the basic argument” for this audacious claim will be the task of Section II of this essay. The argument is, I think, compelling, but by itself it leaves unexplained some widespread intu- itions to the effect that certain specific instances of initial acquisition are unjust and call forth as their remedy the application of a Lockean proviso, or are otherwise problematic. (A “Lockean proviso,” of course, is one that forbids initial acquisitions of resources when these acquisitions do not leave “enough and as good” in common for others.) Thus, Section III focuses on various considerations that tend to show how those intuitions are best explained in a way consistent with the argument of Section II. Section IV completes the task of accounting for the intuitions in question by considering how the thesis of self-ownership itself bears on the acqui- sition and use of property. Section V shows how the results of the previ- ous sections add up to a more satisfying defense of Nozickian property rights than the one given by Nozick himself, and considers some of the implications of this revised conception of initial acquisition for our under- standing of Nozick’s principles of transfer and rectification. II. The Basic Argument The reason there is no such thing as an unjust initial acquisition of resources is that there is no such thing as either a just or an unjust initial acquisition of resources. The concept of justice, that is to say, simply does not apply to initial acquisition. It applies only after initial acquisition has already taken place. In particular, it applies only to transfers of property (and derivatively, to the rectification of injustices in transfer). This, it seems to me, is a clear implication of the assumption (rightly) made by Nozick that external resources are initially unowned. Consider the following example. Suppose an individual A seeks to acquire some previously unowned resource R. For it to be the case that A commits an injustice in acquiring R, it would also have to be the case that there is some individual B (or perhaps a group of individuals) against whom A commits the injustice. But for B to have been wronged by A’s acquisi- tion of R, B would have to have had a rightful claim over R, a right to R. By hypothesis, however, B did not have a right to R, because no one had a right to it—it was unowned, after all. So B was not wronged and could not have been. In fact, the very first person who could conceivably be wronged by anyone’s use of R would be, not B, but A himself, since A is the first one to own R. Such a wrong would in the nature of the case be an injustice in transfer—in unjustly taking from A what is rightfully his—not in initial acquisition. The same thing, by extension, will be true of all unowned resources: it is only after some- one has initially acquired them that anyone could unjustly come to possess them, via unjust transfer. It is impossible, then, for there to be any injustices in initial acquisition.7

#### To own yourself and use your own freedom is to be able to interact with external objects. Anything else makes you unable to exercise your own freedom on other things and creates a contradiction.

Feser 2, (Edward Feser, 1-1-2005, accessed on 12-15-2021, Cambridge University Press, "THERE IS NO SUCH THING AS AN UNJUST INITIAL ACQUISITION | Social Philosophy and Policy | Cambridge Core", Edward C. Feser is an American philosopher. He is an Associate Professor of Philosophy at Pasadena City College in Pasadena, California. [https://www.cambridge.org/core/journals/social-philosophy-and-policy/article/abs/there-is-no-such-thing-as-an-unjust-initial-acquisition/5C744D6D5C525E711EC75F75BF7109D1)[brackets](https://www.cambridge.org/core/journals/social-philosophy-and-policy/article/abs/there-is-no-such-thing-as-an-unjust-initial-acquisition/5C744D6D5C525E711EC75F75BF7109D1)%5bbrackets) for gen lang]//phs st

There is. An alternative, soft-line approach could acknowledge that the initial acquirer who abuses a monopoly over a water hole (or any similar crucial resource) does commit an injustice against those who are disad- vantaged, but such an approach could still hold that the acquirer never- theless has not committed an injustice in acquisition —his acquisition was, as I have said, neither just nor unjust. Nor does he fail to own what he has acquired; he still cannot be said to have stolen the water from anyone. Rather, his injustice is an unjust use of what he owns, on a par with the unjust use I make of my self-owned fist when I wield it, unprovoked, to bop you on your self-owned nose. In what sense does the water-hole owner use his water unjustly, though? He doesn’t try to drown anyone in it, after all— indeed, the whole problem is that he won’t let anybody near it! Eric Mack gives us the answer we need in what he has put forward as the “self-ownership proviso” (SOP).28 This is a proviso not (as the Lock- ean proviso is) on the initial acquisition of property, but rather on how one can use his property in a way that respects others’ self-ownership rights. It is motivated by consideration of the fact that the talents, abilities, capac- ities, energies, etc., that a person rightfully possesses as a self-owner are inherently “world-interactive”; that is, it is of their very essence that they are directed toward the extra-personal environment.29 Your capacity to use your hand, for instance, is just a capacity to grasp and manipulate external objects; thus, what you own in owning your hand is something essentially grasping and manipulating.30 Now if someone were to cut off your hand or invasively keep you from using it (by tying your arm against your body or holding it behind your back), he would obviously be violating your self-ownership rights. But there are, Mack suggests, other, noninvasive ways in which those rights might be violated. If, to use an example of Mack’s, I effectively nullify your ability to use your hand by creating a device that causes anything you reach for to be propelled beyond your grasp, making it impossible for you ever to grasp or manip- ulate anything, I have violated your right to your hand as much as if I had cut it off or tied it down. I have, in any case, prevented your right to your hand from being anything more than a formal right, one that is practically useless. In the interests of guaranteeing respect for substantive, robust rights of self-ownership, then, “[t]he SOP requires that persons not deploy their legitimate holdings, i.e., their extra-personal property, in ways that severely, albeit noninvasively, disable any person’s world-interactive powers.” 31 The SOP follows, in Mack’s view, from the thesis of self-ownership itself; or, at any rate, the considerations that would lead anyone to accept that thesis should also, in his view, lead one to accept the proviso.32 A brief summary of a few of Mack’s thought experiments should suffice to give a sense of why this is so.33 In what Mack calls the Adam’s Island example, Adam acquires a previously uninhabited island and later refuses a shipwrecked Zelda permission to come ashore, as a result of which she remains struggling at sea (and presumably drowns). In the Paternalist Caging example, instead of drowning, Zelda becomes caught offshore in a cage Adam has constructed for catching large sea mammals, and, rather than releasing her, Adam keeps her in the cage and feeds her regularly. In the Knuckle-Scraper Barrier example, Zelda falls asleep on some unowned ground, whereupon a gang of oafish louts encircles her and, using their bodies and arms as barriers, refuses to let her out of the circle (accusing her of assault if she touches them in order to climb over or break through). In the Disabling Property Barrier example, instead of a human barrier, Adam constructs a plastic shield over and around the unowned plot of ground upon which Zelda sleeps, accusing her of trespassing upon his property when she awakens and tries to escape by breaking through the plastic. And in the (similarly named) Disabling Property Barriers example, seem to suggest an Aristotelian-Thomistic conception of natural function, and though this by no means troubles me, it might not be what Mack himself has in mind (nor, of course, is it something every philosopher is going to sympathize with). Mack’s view nevertheless seems to require something like this conception. And something like it —enough like it to do the job Mack needs to be done, anyway—is arguably to be found in Larry Wright’s well- known reconstruction, in modern Darwinian terms, of the traditional notion of natural function. See Larry Wright, “Functions,” Philosophical Review 82, no. 2 (1973): 139–68. Adam, instead of enclosing Zelda in a plastic barrier, encloses in plastic barriers every external object that Zelda would otherwise be able to use — thus, in effect, enclosing her in a larger, all-encompassing plastic barrier of a more eccentric shape. In all of these cases, Mack says, although Zelda’s formal rights of self-ownership have not been violated—no one has invaded the area enclosed by the surface of her skin —her rights over her self-owned powers, and in particular her ability to exercise those powers, have nevertheless been nullified. But a plausible self-ownership- based theory surely cannot allow for this. It cannot, for instance, allow the innocent Zelda justly to be imprisoned in any of the ways described! If Mack is right, then it seems we have, in the SOP, grounds for holding that a water-hole monopolist would indeed be committing an injustice against anyone he refuses water to, or to whom he charges exorbitant prices for access. The injustice would be a straightforward violation of a person’s rights to self-ownership, a case of nullifying a person’s self- owned powers in a way analogous to Adam’s or the knuckle-scrapers’ nullification of Zelda’s self-owned powers. It would not be an injustice in initial acquisition, however. The water-hole monopolist still owns the water hole as much as he ever did; he just cannot use it in a way that violates other individuals’ self-ownership rights (either by drowning them in it or by nullifying their self-owned powers by denying them access to it when there is no alternative way for them to gain access to the water necessary for the use of their self-owned powers). Is Mack right? The hard-liner might dig in his heels and insist that none of Mack’s examples amount to self-ownership-violating injustices; instead, they are merely subtle but straightforward property rights violations or cases of moral failings of various other sorts (cruelty, selfishness, etc.). The Adam’s Island case, for starters, is roughly analogous to the example of the water-hole monopolist, so that it arguably cannot give any non-question- begging support to the SOP, if the SOP is then supposed to show that the water-hole example involves an injustice. The Disabling Property Barriers case might also be viewed as unable to provide any non-question-begging support, since Adam’s encasing everything in plastic might plausibly be interpreted as his acquiring everything, in which case we are back to a water-hole-type monopoly example. The Knuckle-Scraper Barrier and Dis- abling Property Barrier examples might be explained by saying that in falling asleep on the unowned plot of land, Zelda in effect has come (at least temporarily) to acquire it, and (by virtue of walking) to acquire also the path she took to get to it, so that the knuckle-scrapers and Adam violate her property rights (not her self-ownership rights) in not allowing her to escape. The Paternalist Caging example can perhaps be explained by arguing that in building the cage, Adam has acquired the water route leading to it, so that in swimming this route (and thus getting caught in the cage) Zelda has violated his property rights and, therefore, can justly be caged. Accordingly, the hard-liner might insist, we can explain all of these examples in a hard-line way and thus avoid commitment to the SOP. Such a hard-line response would be ingenious (well, maybe), but still, I think, ultimately doomed to failure. Can the Paternalist Caging example, to start with, plausibly be explained away in the manner that I have suggested? Does Adam commit no injustice against Zelda even if he never lets her out? It will not do to write this off merely as a case of excessive punishment (explaining the injustice of which would presumably not require commitment to the SOP). For suppose Adam says, after a mere five minutes of confinement, “I’m no longer punishing you; you’ve paid your debt and are free to go, as far as I’m concerned. But I’m not going to bother exerting the effort to let you out. I never forced you to get in the cage, after all —you did it on your own —and you have no right to the use of my self-owned cage-opening powers to fix your mistake! So teleport out, if you can. Or get someone else —if you can find someone —to let you out.” Adam would be neither violating Zelda’s rights to external property nor excessively punishing her in this case; nor would he be invasively vio- lating her self-ownership rights. But wouldn’t he still be committing an injustice, however noninvasively? Don’t we need something like the SOP to explain why this is so? The barrier examples, for their part, do not require Zelda’s walking and falling asleep on virgin territory, which thus (arguably) becomes her prop- erty. We can, to appeal to the sort of science-fiction scenario beloved of philosophers, imagine instead a bizarre chance disruption of the structure of space-time that teleports Zelda into Adam’s plastic shell or into the midst of the knuckle-scrapers. There is no question now of their violating her property rights; yet don’t they still commit an injustice by nullifying her self-owned powers in refusing to allow her to exit? Consider a parallel example concerning property ownership itself. If your prized $50,000 copy of Captain America Comics number 1, due to another rupture in space-time or just to a particularly strong wind that blows it out of your hands and through my window, suddenly appears on the floor of my living room, do I have the right to refuse to bring it back out to you or to allow you to come in and get it? Suppose I attempt to justify my refusal by saying, “I won’t touch it, and you’re free to have it back if you can arrange another space-time rupture or gust of wind. But I refuse to exert my self-owned powers to bring it out to you, or to allow you on my property to get it. I never asked for it to appear in my living room, after all!” Would anyone accept this justification? Doesn’t your property right in the comic book require me to give it back to you? The hard-liner might suggest that this example transports the SOP advocate out of the frying pan and into the fire. For if the SOP is true, wouldn’t we also have to commit ourselves to a “property-ownership proviso” (POP) that requires us not to nullify anyone’s ability to use his external private property in a way consistent with its “world-interactive powers”? If I build a miniature submarine in my garage, and you have the only swimming pool within one thousand miles, must you allow me the use of your pool lest you nullify my ability to use the sub? If (to take an example of Cohen’s cited by Mack) I own a corkscrew, must I be provided with wine bottles to open lest the corkscrew sadly fail to fulfill its full potential?34 Mack’s response to this line of thought seems basically to amount to a bit of backpedaling on the claim that his proviso really follows from the notion of self-ownership per se —so as to avoid the conclusion that a (rather unlibertarian and presumably redistributionist) POP would also, in par- allel fashion, follow from the concept of property ownership. His response seems, instead, to emphasize the idea that the considerations favoring self-ownership also favor, via an independent line of reasoning, the SOP.35 In my view, however, a better response would be one that took note of some relevant disanalogies between property in oneself and property in external things. Note first that the self-owned world-interactive powers, the possible use of which the SOP is intended to guarantee, are possessed by a living being who is undergoing development, which involves passing through various stages; therefore, these powers are ones that flourish with use and atrophy or even disappear with disuse.36 To nullify these powers even for a limited time, then, is (very often at least) not merely temporarily to inconvenience their owner, but, rather, to bring about a permanent reduc- tion or even disablement of these powers. By contrast, a submarine (or a corkscrew) retains its powers even when left indefinitely in a garage (or a drawer). This difference in the effect that nullification has on self-owned powers versus extra-personal property plausibly justifies a difference in our judgments concerning the acceptability, from the point of view of justice, of such nullification in the two cases; that is, it justifies adoption of the SOP but not of the POP.37 Second, there is an element of choice (and in particular, of voluntary acquisition) where extra-personal property is concerned that is morally relevant here. One’s self-owned powers, along with the SOP-guaranteed right to the non-nullification of those powers, are not something one chooses or acquires; one just has them —indeed, to a great degree one just is the constellation of those powers, abilities, etc.—and owns them fully. By contrast, extra-personal property is something one chooses to acquire or not to acquire, and as we have seen, one always acquires property rights in various degrees, from partial to full ownership—and this would include the rights guaranteed by a POP. If one chooses to acquire a corkscrew under conditions where wine bottles are unavailable, or are even likely at some point to become unavailable, one can hardly blame others if one finds oneself bottle-less. To fail to acquire POP-like rights regarding the corkscrew (by, say, contracting with someone else to provide one with wine bottles in perpetuity) is not the same thing as to have those rights and then have them violated. Someone who buys a corkscrew and then finds that he cannot use it is like the person who acquires only partial property rights in a water hole that others have already acquired partial use rights over. He cannot complain that his co-owners have violated his rights; he never acquired those other rights in the first place. Similarly, the corkscrew owner cannot complain that he has no bottles to open; he never acquired the right to those bottles, only to the corkscrew. If full ownership of a corkscrew requires POP-like rights over it, then all that follows is that corkscrew owners who lack bottles are not full owners of their corkscrews.

#### Thus, self-ownership justifies the appropriation of property – our freedom necessitates being able to set and pursue external things as our ends, including exercising our rights on property. Restricting this arbitrarily limits our freedom which is unjust.

Feser 3, (Edward Feser, 1-1-2005, accessed on 12-15-2021, Cambridge University Press, "THERE IS NO SUCH THING AS AN UNJUST INITIAL ACQUISITION | Social Philosophy and Policy | Cambridge Core", Edward C. Feser is an American philosopher. He is an Associate Professor of Philosophy at Pasadena City College in Pasadena, California. [https://www.cambridge.org/core/journals/social-philosophy-and-policy/article/abs/there-is-no-such-thing-as-an-unjust-initial-acquisition/5C744D6D5C525E711EC75F75BF7109D1)[brackets](https://www.cambridge.org/core/journals/social-philosophy-and-policy/article/abs/there-is-no-such-thing-as-an-unjust-initial-acquisition/5C744D6D5C525E711EC75F75BF7109D1)%5bbrackets) for gen lang]//phs st

V. Some Implications If what I have argued so far is correct, then the way is opened to the following revised case for strongly libertarian Lockean-Nozickian prop-erty rights: We are self-owners, having full property rights to our body parts, powers, talents, energies, etc. As self-owners, we also have a right, given the SOP, not to have our self-owned powers nullified —we have the right, that is, to act within the extra-personal world and thus to acquire rights to extra-personal objects that the use of our self-owned powers requires.39 This might involve the buying or leasing of certain rights or bundles of rights and, correspondingly, the acquiring of lesser or greater degrees of ownership of parts of the external world, but as long as one is able to exercise one’s powers to some degree and is not rendered incapable of acting within that world, the SOP is satisfied. In any case, such rights can only be traded after they are first established by initial acquisition. In initially acquiring a resource, an agent does no one an injustice (it was unowned, after all). Furthermore, [they] has mixed [their] labor with the resource, significantly altering it and/or bringing it under his control, and is himself solely responsible for whatever value or utility the resource has come to have. Thus, [they] has a presumptive right to it, and, if his control and/or alteration (and thus acquisition) of it is (more or less) complete, his own- ership is accordingly (more or less) full. The system of strong private property rights that follows from the acts of initial acquisition performed by countless such agents results, as a matter of empirical fact, in a market economy that inevitably and dramatically increases the number of resources available for use by individuals, and these benefited individuals include those who come along long after initial acquisition has taken place. (Indeed, it especially includes these latecomers, given that they were able to avoid the hard work of being the first to “tame the land” and draw out the value of raw materials.)40 The SOP is thus, in fact, rarely, if ever, violated. The upshot is that a system of Lockean-Nozickian private property rights is morally justified, with a strong presumption against tampering with exist- ing property titles in general. In any case, there is a strong presumption against any general egalitarian redistribution of wealth, and no case what- soever to be made for such redistribution from the general theory of prop- erty just sketched, purged as it is of the Lockean proviso, with all the egalitarian mischief-making the proviso has made possible.

## Case

#### Have a high threshold for 1AR extrapolations for impacts. Their card is under warranted and causally asserts claims without delineated warrants—don’t fill in gaps for them. Reject new 1ar extinction evidence—we based our 1nc strategy off of bad impact evidence. Study indicts and answers to our impacts/transition solves but discourages sandbagging good ev until after the 1nc.

#### Nuclear war now spurs political will for disarmament without causing extinction.

Deudney 18 [Associate Professor of Political Science at Johns Hopkins University. 03/15/2018. “The Great Debate.” The Oxford Handbook of International Security. www.oxfordhandbooks.com, doi:10.1093/oxfordhb/9780198777854.013.22] Recut Justin

Although nuclear war is the oldest of these technogenic threats to civilization and human survival, and although important steps to restraint, particularly at the end of the Cold War, have been achieved, the nuclear world is increasingly changing in major ways, and in almost entirely dangerous directions. The third “bombs away” phase of the great debate on the nuclear-political question is more consequentially divided than in the first two phases. Even more ominously, most of the momentum lies with the forces that are pulling states toward nuclear-use, and with the radical actors bent on inflicting catastrophic damage on the leading states in the international system, particularly the United States. In contrast, the arms control project, although intellectually vibrant, is largely in retreat on the world political stage. The arms control settlement of the Cold War is unraveling, and the world public is more divided and distracted than ever. With the recent election of President Donald Trump, the United States, which has played such a dominant role in nuclear politics since its scientists invented these fiendish engines, now has an impulsive and uninformed leader, boding ill for nuclear restraint and effective crisis management. Given current trends, it is prudent to assume that sooner or later, and probably sooner, nuclear weapons will again be the used in war. But this bad news may contain a “silver lining” of good news. Unlike a general nuclear war that might have occurred during the Cold War, such a nuclear event now would probably not mark the end of civilization (or of humanity), due to the great reductions in nuclear forces achieved at the end of the Cold War. Furthermore, politics on “the day after” could have immense potential for positive change. The survivors would not be likely to envy the dead, but would surely have a greatly renewed resolution for “never again.” Such an event, completely unpredictable in its particulars, would unambiguously put the nuclear-political question back at the top of the world political agenda. It would unmistakeably remind leading states of their vulnerability It might also trigger more robust efforts to achieve the global regulation of nuclear capability. Like the bombings of Hiroshima and Nagasaki that did so much to catalyze the elevated concern for nuclear security in the early Cold War, and like the experience “at the brink” in the Cuban Missile Crisis of 1962, the now bubbling nuclear caldron holds the possibility of inaugurating a major period of institutional innovation and adjustment toward a fully “bombs away” future.

#### That’s good – war later is worse.

Turchin & Denkenberger 18 [Alexey Turchin & David Denkenberger. Turchin is a researcher at the Science for Life Extension Foundation; Denkenberger is with the Global Catastrophic Risk Institute (GCRI) @ Tennessee State University, Alliance to Feed the Earth in Disasters (ALLFED). 09/2018. “Global Catastrophic and Existential Risks Communication Scale.” Futures, vol. 102, pp. 27–38.]

2. “Civilizational collapse risks” As most human societies are fairly complex, a true civilizational collapse would require a drastic reduction in human population, and the break-down of connections between surviving populations. Survivors would have to rebuild civilization from scratch, likely losing much technological abilities and knowledge in the process. Hanson (2008) estimated that the minimal human population able to survive is around 100 people. Like X risks, there is little agreement on what is required for civilizational collapse. Clearly, different types and levels of the civilizational collapse are possible (Diamond, 2005) (Meadows, Randers, & Meadows, 2004). For instance, one definition of the collapse of civilization involves, collapse of long distance trade, widespread conflict, and loss of government (Coates, 2009). How such collapses relate to existential risk needs more research.

3. “Human extinction risks” are risks that all humans die, and no future generations (in the extended sense mentioned above) will ever exist.

4. “All life on Earth ends risks” involve the extinction of all life on earth. As this includes H. sapiens, such risks are at the very least on a par with human extinction, but are likely worse as the loss of biodiversity is higher, and (without life arising a second time) no other civilizations, human or otherwise, would be possible on Earth.

5. “Astronomical scale risks” include the demise of all civilizations in the affectable universe. This of course includes human extinction, and all life on Earth, and so again are at the very least on a par, and very likely much worse outcomes, than those two.

6. “S-risks” include collective infinite suffering (Daniel, 2017). These differ from extinction risks insofar as extinction leads to a lack of existence, whereas this concerns ongoing existence in undesirable circumstances. These also vary in scale and intensity, but are generally out of scope of this work.

Even with a focus squarely on X Risk, global catastrophic risks and civilizational collapse are critically important. This is because there is at least some likelihood that global catastrophic risks increase the probability of human extinction risks—and the more extreme end of civilizational collapses surely would. Before shifting to a discussion of probability appropriate to X risk, we’ll discuss some reasons to link these kinds of risk.

First, global risks may have a fat tail—that is a low probability of high consequences—and the existence of such fat tails strongly depend on the intrinsic uncertainty of global systems (Ćirković, 2012) (Baum, 2015), (Wiener, 2016) (Sandberg & Landry, 2015). This is especially true for risks associated with future world wars, which may include not only nuclear weapons, but weapons incorporating synthetic biology and nanotechnology, different AI technologies, as well as Doomsday blackmail weapons (Kahn, 1959). Another case are the risks associated with climate change, where runaway global warming is a likely fat tail (Obata & Shibata, 2012a), (Goldblatt & Watson, 2012).

Second, global catastrophes could be part of double catastrophe (Baum, Maher, & Haqq-Misra, 2013) or start a chain of catastrophes (Tonn & and MacGregor, 2009), and in this issue (Karieva, 2018). Even if a single catastrophic risk is insufficient to wipe us out, an unhappy coincidence of such events could be sufficient, or under the wrong conditions could trigger a collapse leading to human extinction. Further, global catastrophe could weaken our ability to prepare for other risks.

Luke Oman has estimated the risks of human extinction because of nuclear winter: “The probability I would estimate for the global human population of zero resulting from the 150 Tg of black carbon scenario in our 2007 paper would be in the range of 1 in 10,000 to 1 in 100,000” (Robock, Oman, & Stenchikov, 2007), (Shulman, 2012).

Tonn also analyzed chains of events, which could result in human extinction and any global catastrophe may be a start of such chain (Tonn and MacGregor, 2009). Because this, we suggest that any global catastrophe should be regarded as a possible cause of human extinction risks with no less than 0.01 probability. Similarly, scenarios involving civilization collapses also plausibly increase the risk of human extinction. If civilization collapses, recovery may be slowed or stopped for a multitude of reasons. For instance, easily accessible mineral and fossil fuel resources might be no longer available, the future climate may be extreme or unstable, we may not regain sufficient social trust after the catastrophe’s horrors, the catastrophe may affect our genetics, a new endemic disease could prevent high population density, and so on. And of course, the smaller populations associated with civilization collapse are more vulnerable to being wiped out by natural events. We estimate that civilization collapse has a 0.1 probability of becoming an existential catastrophe. In section 4, this discussion will form the basis of our analysis of an X risk’s “severity”, which is the main target of our scale. Before getting there, however, we should first discuss the difficulties of measuring X risks, and related worries regarding probabilities. 3. Difficulties of using probability estimates as the communication tool Plain probability estimates are often used as an instrument to communicate X risks. An example is a claim like “Nuclear war could cause human extinction with probability P”. However, in our view, probability measures are inadequate, both for measuring X risks, and for communicating those risks. This is because of conceptual difficulties (3.1), difficulty in providing meaningful measurements (3.2), the possibility of interaction effects (3.3) and the measurement’s inadequacy for prioritization (3.4) purposes. After presenting these worries, we argue that the magnitude of probabilities is a better option, which we use in our tool (3.5). 3.1 Difficulties in defining X risk probabilities Frequentism applies to X risks only with difficulty. One-off events don’t have a frequency, and multiple events are required for frequentist probabilities to apply. Further, on a frequentist reading, claims concerning X risks cannot be falsified. Again, this is because in order to infer from occurrences to probability, multiple instances are required. Although these conceptual and epistemic difficulties may be analyzed and partly overcome in technical scientific and philosophical literature, they would overcomplicate a communication tool. Also, discussion of X risks sometimes involves weird probabilistic effects. Consider, for example, what (Ćirković, Sandberg, & Bostrom, 2010) call the ‘anthropic shadow’. Because human extinction events entail a lack of humans to observe the event after the fact, we will systematically underestimate the occurrence of such events in an extreme case of survivorship bias (the Doomsday Argument (Tegmark & Bostrom, 2005) is similar). All of this makes the probabilities attached to X risks extremely difficult to interpret, bad news for an intended communication tool, and stimulates obscure anthropic reasoning. In addition, the subtle features involved in applying frequentism to one-off events, would otherwise tamper with our decision making process. 3.2 Data & X Risk There are little hard data concerning global risks from which probabilities could be extracted. The risk of an asteroid impact is fairly well understood, both due to the historical record, and because scientists can observe particular asteroids and calculate their trajectories. Studies of nuclear winter (Denkenberger & Pearce, 2016), volcanic eruptions, and climate change also provide some risk probability estimates, but are less rigorously supported. In all other cases, especially technological risks, there are many (often contradicting) expert opinions, but little hard data. Those probability calculations which have been carried out are based on speculative assumptions, which carry their own uncertainty. In the best case, generally, only the order of magnitude of the catastrophe’s probability can be estimated. Uncertainty in GCRs is so high, that predictions with high precision are likely to be meaningless. For example, surveys could produce such meaningless over-precision. A survey on human extinction probability gave an estimate of 19 percent in the 21st century (Sandberg & Bostrom, 2008). Such measurements are problematic for communication, because probability estimates of global risks often do not include corresponding confidence intervals (Garrick, 2008). For some catastrophic risks, uncertainty is much larger than for others, because of objective difficulties in their measurement, as well as subjective disagreements between various approaches (especially in the case of climate change, resource depletion, population growth and other politicized areas). As we’ll discuss below, one response is to present probabilities as magnitudes. 3.3 Probability density, timing and risks’ interactions Two more issues with using discrete frequentist probabilities for communicating X risks are related to probability density and the interactions between risks. For the purpose of responding to the challenges of X risk, the total probability of an event is less useful than the probability density: we want to know not only the probability but the time in which it is measured. This is crucial if policy makers are to prioritize avoidance efforts. Also, probability estimates of the risks are typically treated separate: interdependence is thus ignored. The total probability of human extinction caused by risk A could strongly depend on the extinction probability caused by risks B and C and also of their timing. (See also double catastrophes discussed by Baum, Maher, & HaqqMisra, 2013 and the integrated risk assessment project (Baum, 2017). Further, probability distributions of different risks can have different forms. Some risks are linear, others are barrier-like, other logistical. Thus, not all risks can be presented by a single numerical estimate. Exponentially growing risks may be the best way to describe new technologies, such as AI and synthetic biology. Such risks cannot be presented by a single annual probability. Finally, the probability estimation of a risk depends on whether human extinction is ultimately inevitable. We assume that if humanity becomes an interstellar civilization existing for millions of years, it will escape any near-term extinction risks; the heat death of the universe may be ultimate end, but some think even that is escapable (Dvorsky, 2015). If near-term extinction is inevitable, it is possible to estimate which risks are more probable to cause human extinction (like actuaries do in estimating different causes of death, based in part on the assumption that human death is inevitable). If near-term human extinction is not inevitable, then there is a probability of survival, which is (1- P(all risks)). Such conditioning requires a general model of the future. If extinction is inevitable, the probability of a given risk is just a probability of one way to extinction compared to other ways. 3.4 Preventability, prioritizing and relation to the smaller risks Using bare probability as a communication tool also ignores many important aspects of risks which are substantial for decision makers. First, a probability estimate does not provide sufficient guidance on how to prioritize prevention efforts. A probability estimate does not say anything about the risk’s relation to other risks, e.g. its urgency. Also, if a risk will take place at a remote time in the future (like the Sun becoming a red giant), there is no reason to spend money on its prevention. Second, a probability estimate does not provide much information about the relation of human extinction risks, and corresponding smaller global catastrophic risks. For example, a nuclear war probability estimate does not disambiguate between chances that it will be a human extinction event, a global catastrophic event, or a regional catastrophe. Third, probability measures do not take preventability into account. Hopefully, measures will be taken to try and reduce X risks, and the risks themselves have individual preventability. Generally speaking, it ought to be made clear when probabilities are conditional on whether prevention is attempted or not, and also on the probability of its success. Probability density, and its relation with cumulative probability could also be tricky, especially as the probability density of most risks is changing in time. 3.5 Use of probability orders of magnitude as a communication tool We recommend using magnitudes of probabilities in communicating about X risk. One way of overcoming many of the difficulties of using probabilities as communication tool described above is to estimate probabilities with fidelity of one or even two orders of magnitude, and do it over large fixed interval of time, that is the next 100 years (as it the furthest time where meaningful prognoses exist). This order of magnitude estimation will smooth many of the uncertainties described above. Further, prevention actions are typically insensitive in to the exact value of probability. For example, if a given asteroid impact probability is 5% or 25%, needed prevention action will be nearly the same. For X risks, we suggest using probability intervals of 2 orders of magnitude. Using such intervals will often provide meaningful differences in probability estimates for individual risks. (However, expert estimates sometimes range from “inevitable” to “impossible”, as in AI risks). Large intervals will also accommodate the possibility of one risk overshadowing another, and other uncertainties which arise from the difficulties of defining and measuring X-risks. This solution is itself inspired by The Torino scale of asteroid danger, which we discuss in more detail below. The Torino scale has five probability intervals, each with a two order of magnitude difference from the next. Further, such intervals can be used to present uncertainty in probability estimation. This uncertainty is often very large for even approximately well-defined asteroid risks. For example, Garrick (Garrick, 2008) estimated that asteroid impacts on the contiguous US with at least 10 000 victims to have expected frequency between once 1: 1900 and 1: 520 000 years with 90 percent confidence. In other words, it used more than 2 orders of magnitude uncertainty. Of course, there is a lot more to be said about the relationship between X risks and probability—however here we restrict ourselves to those issues most crucial for our purpose, that is, designing a communication tool for X risks. 4. Constructing the scale of human extinction risks 4.1. Existing scales for different catastrophic risks In section 2 we established the connection between global catastrophic risks, civilizational collapse risks, human extinction and X risks; we explored the difficulty of the use of probabilities as a communication tool for X risks in section 3; now we can construct the scale to communicate the level of risk of all global catastrophic and X risks. Our scale is inspired by the Torino scale of asteroid danger which was suggested by professor Richard Binzel (Binzel, 1997). As it only measures the energy of impact, it is not restricted to asteroids but applies to many celestial bodies (comets, for instance). It was first created to communicate the level of risk to the public, because professionals and decision makers have access to all underlying data for the hazardous object. The Torino scale combines a 5 level color code and 11 level numbered codes. One of the Torino scale’s features is that it connects the size and the probability using diagonal lines, i.e., an event with a bigger size and smaller probability warrants the same level of attention as smaller but more probable events. However, this approach has some difficulties, as was described by (Cox, 2008). There are several other scales of specific global risks based on similar principles: 1. Volcanic explosivity index, VEI, 0-8, (USGS, 2017) 2. DEFCON (DEFense readiness CONdition, used by the US military to describe five levels of readiness), from 5 to 1. 3. “Rio scale” of the Search for Extra-Terrestrial Intelligence (SETI) – complex scale with three subscales (Almar, 2011). 4. Palermo scale of asteroid risks compares the likelihood of the detected potential impactor with the average risk posed by objects of the same size measured both by energy and frequency (NASA, 2017). 5. San-Marino scale of risks of Messaging to Extra-Terrestrial Intelligence (METI) (Almar, 2007). The only more general scale for several global risks is the Doomsday Clock by the Bulletin of the Atomic Scientists, which shows global risks as minutes before midnight. It is oriented towards risks of a nuclear war and climate change and communicates only emotional impact (The Bulletin of the Atomic Scientists, 2017). 4.2. The goals of the scale How good a scale is depends in part on what it is intended to do: who will use it and how will they use it. There are three main groups of people the scale addresses: Public. Simplicity matters: a simple scale is required, similar to the hurricane Saffir-Simpson scale (Schott et al., 2012). This hurricane ACCEPTED MANUSCRIPT 13 measuring scale has 5 levels which present rather obscure wind readings as corresponding to the expected damage to houses and thus can help the public make decisions about preparedness and evacuation. In the case of X risks, personal preparedness is not very important, but the public make decisions about which prevention projects to directly support (via donations or crowdfunding) or voting for policymakers who support said projects. Simplicity is necessary to communicate the relative importance of different dangers to a wide variety of nonexperts. Policymakers. We intend our scale to help initiate communication of the relative importance of the risks to policymakers. This is particularly important as it appears that policymakers tend to overestimate smaller risks (like asteroid impact risks) and underestimate larger risks (like AI risks) (Bostrom, 2013). Our scale helps to make such comparison possible as it does not depend on the exact nature of the risks. The scale could be applicable to several groups of risks thus allowing comparisons between them, as well as providing a perspective across the whole situation. Expert community. Even a scale of the simplicity we suggest may benefit the expert community. It can act as a basis for comparing different risks by different experts. Given the interdisciplinarity inherent in studying X risk, this common ground is crucial. The scale could facilitate discussion about catastrophes’ probabilities, preventability, prevention costs, interactions, and error margins, as experts from different fields present arguments about the importance of the risks on which they work. Thus it will help to build a common framework for the risk discussions. 4.3. Color codes and classification of the needed actions Tonn and Steifel suggested a six-level classification of actions to prevent X risks (Tonn & Steifel, 2017). They start from “do nothing” and end with “extreme war footing, economy organized around reducing human extinction risk”. We suggest a scale which is coordinated with Tonn and Steifel’s classification of actions (Table 1), that is our colors correspond to the needed level of action. Also, our colors correspond to typical nonquantifiable ways of the risks description: theoretical, small, medium, serious, high and immediate. We also add iconic examples, which are risks where the probability distribution is known with a higher level of certainty, and thus could be used to communicate the risk’s importance by comparison. Such ACCEPTED MANUSCRIPT 14 examples may aid in learning the scale, or be used instead of the scale. For instance, someone could say: “this risk is the same level as asteroid risk”. The iconic risks are marked bold in the scale. Iconic examples are also illustrated with the best-known example of that type of event. For example, the best known supervolcanic eruption was the Toba eruption 74,000 years ago (Robock et al., 2009). The Chicxulub impact 66 million years ago is infamous for being connected with the latest major extinction, associated with the non-avian Dinosaur extinction. The scale presents the total risk of one type of event, without breaking categories down into subrisks. For example, it estimates the total risks of all known and unknown asteroids, but not the risk of any particular asteroid, which is a departure from the Torino scale. Although the scale is presented using probability intervals, it could be used instead of probabilities if they are completely unknown, but other factors, such as those affecting scope and severity, are known. For example, we might want to communicate that AI catastrophe is a very significant risk, but its exact probability estimation is complicated by large uncertainties. Thus we could agree to represent the risk as red despite difficulties of its numerical estimation. Note that the probability interval (when it is known) for “red” is shorter and is only 1 order of magnitude, as it is needed to represent most serious risks and here we need better resolution ability. As it is a communication scale, the scientists using it could come to agreement that a particular risk should be estimated higher or lower in this scale. We don’t want to place too many restrictions on how different aspects of a risk’s severity (like preventability or connection with other risks) should affect risks coding, as it should be established in the practical use of the scale. However, we will note two rules: 1. The purple color is reserved to present extreme urgency of the risk 2. The scale is extrapolated from the smaller than extinction risks and larger than extinction risks in Table 2. (This is based on idea that smaller risks have considerable but unknown probability to become human extinction risks, and also on the fact that policy makers may implement similar measures for smaller and larger risks). 4.4. Extrapolated version of scale which accounts for the risk size In Table 2 we extend the scale to include smaller risks like civilization collapse and global catastrophic risks as well as on “larger” ACCEPTED MANUSCRIPT 15 risks like life extinction and universe destruction, in accordance with our discussion in section 2. This is necessary because: 1) Smaller risks could become larger extinction risks by starting chains of catastrophic events. 2) The public and policymakers will react similarly to human extinction level catastrophe and to a global catastrophe where there will be some survival: both present similar dangers to personal survival, and in both similar prevention actions are needed. [[TABLE 2 OMITTED]] 4.5. Accessing risks with shorter timeframes than 100 years In Table 2 above we assessed the risks for the next 100 years. However, without prevention efforts, some risks could approach a probability of 1 in less time: climate change, for instance. We suggest that the urgency of intervening in such cases may be expressed by increasing their color coding. Moreover, the critical issue is less the timing of risks, but the timing of the prevention measures. Again, although extreme global warming would likely only occur at the end of the 21st century, it is also true that cutting emissions now would ameliorate the situation. We suggest, then, three ranks which incorporate these shorter time-frame risks. Note that the timings relate to implementation of interventions not the timings of the catastrophes. 1) Now. This is when a catastrophe has started, or may start in any moment: The Cuban Missile Crisis is an historical example. We reserve purple to represent it. 2) “Near mode”. Near mode is roughly the next 5 years. Typically current political problems (as in current relations with North Korea) are understood in near mode. Such problems are appropriately explored in terms of planning and trend expectations. Hanson showed that people are very realistic in “Near mode”, but become speculative and less moral in “Far mode” thinking (Hanson, 2010). Near mode may require one color code increase. 3) “Next 2-3 decades”. Many futurists predict a Technological Singularity between 2030-2050: that is around 10-30 years from now (Vinge, 1993), (Kurzweil, 2006). As this mode coincides with an adult’s working life, it may also be called “in personal life time”. In this mode people may expect to personally suffer from a catastrophe, or be personally responsible for incorrect predictions. MIRI recently increased its estimation of the probability that AGI will appear around 2035 (MIRI, 2017), pushing AGI into “next 2-3 decades” mode. There is a consideration against increasing the color code too much for near-term risks, as that may lead to myopia regarding longterm risks of human extinction. There will always be smaller but more urgent risks, and although these ought to be dealt with, some resources ought to be put towards understanding and mitigating the longer term. ACCEPTED MANUSCRIPT 19 Having said this, in high impact emergency situations, short term overwhelming efforts may help to prevent impending global catastrophe. Examples include the Cuban missile crisis and fighting the recent Ebola pandemic in Western Africa. Such short-term efforts do not necessarily constrain our long-term efforts towards preventing other risks. Thus, short term global catastrophic and larger risks may get a purple rating. 4.6. Detailed explanation of risk assessment principles in the color coded scale In Table 3, we estimate the main global risks, according to the scale suggested in section 4.4. Table 3. Detailed explanation of the X risks scale Color code Examples of risks White Sun becomes red giant. Although this risk is practically guaranteed, it is very remote indeed. Natural false vacuum decay. Bostrom and Tegmark estimated such events as happening in less than one in 1 billion years, (that is 10-7 in a century) (Tegmark & Bostrom, 2005). Moreover, nothing can be done to prevent it. Green Gamma-ray bursts. Earth threatening gamma-ray bursts are extremely rare, and in most cases they will result only in a crop failure due to UV increases. However, a close gamma-ray burst may produce a deadly muon shower which may kill everything up to 3 km in depth (A. Dar, Laor, & N.J, 1997). However, such events could happen less than once in a billion years (10-7 in a century) (Cirković & Vukotića, 2016). Such an event will probably kill all multicellular life on Earth. Dar estimates risks of major extinction events from gamma ray bursts as 1 in 100 mln years (A. Dar, 2001). Asteroid impacts. No dangerous asteroids have been thus far identified, and the background level of global catastrophic impacts is around 1 in a million years (10- 4 in a century). Extinction-level impact probability is 10-6 per century. There are several prevention options involving deflecting comets/asteroids. Also, food security could be purchased cheaply (Denkenberger, 2015). However, some uncertainty exists. Some periods involve intense comet bombardment, and if we are in such a time investment in telescopes should be larger (Rampino & Caldeira, 2015). High energy accelerator experiments creating false vacuum decay/black hole/strangelet. Vacuum decay seems to have extremely low probability, far below 10-8 currently. One obvious reason for expecting such events to have very low probability is that similar events happen quite often, and haven’t destroyed everything as yet (Kent, 2004). However, we give this event a higher estimation for two reasons. First, as accelerators become more capable such events might become more likely. Second, the risks are at an astronomical scale: it could affect other civilizations in the universe. Other types of accelerator catastrophes, like mini-black hole or strangelet creation, would only kill Earth life. However, these are more likely, with one estimate being <2E-8 risk from a single facility (the Relativistic Heavy Ion Collider) (Arnon Dar, De Rújula, & Heinz, 1999), which should be coded white. There many unknowns about dangerous experiments (Sandberg & Landry, 2015). Overall, these risks should be monitored, so green is advisable. Yellow Supervolcanic eruption. Given historical patterns, the likelihood of living in a century containing a super volcanic eruption is approximately 10-3 (Denkenberger, 2014). However, the chance of human extinction resulting is ACCEPTED MANUSCRIPT 21 significantly lower than this. If such an eruption produces global crop failure, it could end current civilization. Conventional wisdom is that there is nothing that could be done to prevent a super volcano from erupting, but some possible preventive measures have been suggested (Denkenberger, this issue). We estimate supervolcanic risks to be higher than asteroid impacts because of the historical record, as they likely nearly finished us off 74 000 ago (Robock et al., 2009). Natural pandemic. A natural pandemic is fairly likely to kill 1% (to an order of magnitude) of the global population during this century, as the Spanish flu did. However, such a pandemic is very unlikely to cause total extinction because lethality is under 100% and some populations are isolated. Between all natural pandemics, emerging pandemic flus have a shorter timespan and need much more attention. Bird flu has a mortality above 0.5 (WHO, 2017) and could produce widespread chaos and possible civilizational collapse if human-to-human transmission starts. Therefore, we estimate 10% probability this century of 10% mortality. Global warming triggering global catastrophe. According to the IPCC anthropogenic global warming may affect billions of people by the end of the 21st century (Parry, 2007), causing heat waves, crop failures and mass migration. Those events, and downstream consequences such as conflicts, could conceivably kill 1 billion people. However, this would only occur for tail risk scenarios which have order of magnitude 1% probability. Having said this, several experts think that methane release from permafrost and similar positive feedback loops may result in runaway global warming with much larger consequences (Obata & Shibata, 2012).

Orange Full-scale nuclear war. There is roughly 0.02-7% chance per year of accidental full-scale nuclear war between the US and Russia (Barrett, Baum, & Hostetler, 2013). With fairly high probabilities of nuclear winter and civilization collapse given nuclear war, this is order of magnitude 10% this century. We should also take into consideration that despite reductions in nuclear weapons, a new nuclear arms race is possible in the 21st century. Such a race may include more devastating weapons or cheaper manufacturing methods. Nuclear war could include the creation of large cobalt bombs as doomsday weapons or attacks on nuclear power plants. It could also start a chain of events which result in civilization collapse. Nanotechnology risks. Although molecular manufacturing can be achieved without self-replicating machines (Drexler & Phoenix, 2004), technological fascination with biological systems makes it likely that self-replicating machines will be created. Moreover, catastrophic uses of nanotechnology needn’t be due to accident, but also due to the actions of purposeful malignant agents. Therefore, we estimate the chance of runaway self-replicating machines causing “gray goo” and thus human extinction to be one per cent in this century. There could also be extinction risks from weapons produced by safe exponential molecular manufacturing. See also (Turchin, 2016). Artificial pandemic and other risks from synthetic biology. An artificial multipandemic is a situation in which multiple (even hundreds) of individual viruses created through synthetic biology are released simultaneously either by a terrorist state or as a result of the independent activity of biohackers (Turchin, Green, & Dekenbergern, 2017). Because the capacity to create such a multipandemic could arrive as early as within the next ten to thirty years (as all the needed technologies already exist), it could overshadow future risks, like nanotech and AI, so we give it a higher estimate. There are also other possible risks, connected with synthetic biology, which are widely recognized as serious (Bostrom, 2002). Agricultural catastrophe. There is about a one per cent risk per year of a ten per cent global agricultural shortfall occurring due to a large volcanic eruption, a medium asteroid or comet impact, regional nuclear war, abrupt climate change, or extreme weather causing multiple breadbasket failures (Denkenberger 2016). This could lead to 10% mortality. Red AI risks. The risks connected with the possible creation of non-aligned Strong AI are discussed by (Bostrom, 2014), (Yudkowsky, 2008), (Yampolskiy & Fox, 2013) and others. It is widely recognized as the most serious X risk. AI could start an “intelligence explosion wave” through the Universe, which could prevent appearance of the other civilizations before they create their own AI. Purple Something like the Caribbean crisis in the past, but larger size. Currently, there are no known purple risks. If we could be sure that Strong AI will appear in the next 100 years and would probably be negative, it would constitute a purple risk. Another example would be the creation of a Doomsday weapon that could kill our species with global radiation poisoning (much greater ionizing radiation release than all of the current nuclear weapons) (Kahn, 1959). A further example would be a large incoming asteroid being located, or an extinction level pandemic has begun. These situations require quick and urgent effort on all levels.

#### Evil AI causes infinite torture

**Turchin & Denkenberger 18** {Turchin is a researcher at the Science for Life Extension Foundation; Denkenberger is with the Global Catastrophic Risk Institute (GCRI) @ Tennessee State University, Alliance to Feed the Earth in Disasters (ALLFED). 5-3-2018. “Classification of Global Catastrophic Risks Connected with Artificial Intelligence.”}//JM

6.4. AI that is programmed to be evil We could imagine a perfectly aligned AI, which was deliberately programmed to be bad by its creators. For example, a hacker could create an AI with a goal of killing all humans or torturing them. The Foundational Research Institute suggested the notion of s-risks, that is, the risks of extreme future suffering, probably by wrongly aligned AI (Daniel 2017). AI may even upgrade humans to make them feel more suffering, like in the short story “I have no mouth but I must scream” (Ellison 1967). The controversial idea of “Roko’s Basilisk” is that a future AI may torture people who did not do enough to create this malevolent AI. This idea has attracted attention in the media and is an illustration of “acausal” (not connected by causal links) blackmail by future AI (Auerbach 2014). However, this cannot happen unless many people take the proposition seriously.

#### Can’t rebuild industrial civilization.

John Jacobi 17. [Leads an environmentalist research institute and collective, citing Fred Hoyle, British astronomer, formulated the theory of stellar nucleosynthesis, coined the term “big bang,” recipient of the Gold Medal of the Royal Astronomical Society, professor at the Institute of Astronomy, Cambridge University. 05-27-17. “Industrial Civilization Could Not Be Rebuilt.” The Wild Will Project. <https://www.wildwill.net/blog/2017/05/27/industrial-civilization-not-rebuilt/>] Recut Justin

A suggestion, for the sake of thought: If industrial civilization collapsed, it probably could not be rebuilt. Civilization would exist again, of course, but industry appears to be a one-time experiment. The astronomist Fred Hoyle, exaggerating slightly, writes: It has often been said that, if the human species fails to make a go of it here on Earth, some other species will take over the running. In the sense of developing high intelligence this is not correct. We have, or soon will have, exhausted the necessary physical prerequisites so far as this planet is concerned. With coal gone, oil gone, high-grade metallic ores gone, no species however competent can make the long climb from primitive conditions to high-level technology. This is a one-shot affair. If we fail, this planetary system fails so far as intelligence is concerned. The same will be true of other planetary systems. On each of them there will be one chance, and one chance only. Hoyle overstates all the limits we actually have to worry about, but there are enough to affirm his belief that industry is a “one-shot affair.” In other words, if industry collapsed then no matter how quickly scientific knowledge allows societies to progress, technical development will hit a wall because the builders will not have the needed materials. For example, much of the world’s land is not arable, and some of the land in use today is only productive because of industrial technics developed during the agricultural revolution in the 60s, technics heavily dependent on oil. Without the systems that sustain industrial agriculture much current farm land could not be farmed; agricultural civilizations cannot exist there, at least until the soil replenishes, if it replenishes. And some resources required for industrial progress, like coal, simply are not feasibly accessible anymore. Tainter writes: . . . major jumps in population, at around A.D. 1300, 1600, and in the late eighteenth century, each led to intensification in agriculture and industry. As the land in the late Middle Ages was increasingly deforested to provide fuel and agricultural space for a growing population, basic heating, cooking, and manufacturing needs could no longer be met by burning wood. A shift to reliance on coal began, gradually and with apparent reluctance. Coal was definitely a fuel source of secondary desirability, being more costly to obtain and distribute than wood, as well as being dirty and polluting. Coal was more restricted in its spatial distribution than wood, so that a whole new, costly distribution system had to be developed. Mining of coal from the ground was more costly than obtaining a quantity of wood equivalent in heating value, and became even more costly as the 54 most accessible reserves of this fuel were depleted. Mines had to be sunk ever deeper, until groundwater flooding became a serious problem. Today, most easily accessible natural coal reserves are completely depleted. Thus, societies in the wake of our imagined collapse would not be able to develop fast enough to reach the underground coal. As a result of these limits, rebuilding industry would take at least thousands of years — it took 10,000 years the first time around. By the time a civilization reached the point where it could do something about industrial scientific knowledge it probably would not have the knowledge anymore. It would have to develop its sciences and technologies on its own, resulting in patterns of development that would probably look similar to historical patterns. Technology today depends on levels of complexity that must proceed in chronological stages. Solar panels, for example, rely on transportation infrastructure, mining, and a regulated division of labor. And historically the process of developing into a global civilization includes numerous instances of technical regression. The natives of Tasmania, for example, went from a maritime society to one that didn’t fish, build boats, or make bows and arrows. Rebuilding civilization would also be a bad idea. Most, who are exploited by rather than benefit from industry, would probably not view a rebuilding project as desirable. Even today, though citizens of first-world nations live physically comfortable lives, their lives are sustained by the worse off lives of the rest of the world. “Civilization . . . has operated two ways,” Paine writes, “to make one part of society more affluent, and the other more wretched, than would have been the lot of either in a natural state.” Consider the case of two societies in New Zealand, the Maori and the Moriori. Both are now believed to have originated out of the same mainland society. Most stayed and became the Maori we know, and some who became the Moriori people settled on the Chatham Islands in the 16th century. Largely due to a chief named Nunuku-whenua, the Moriori had a strict tradition of solving inter-tribal conflict peacefully and advocating a variant of passive resistance; war, cannibalism, and killing were completely outlawed. They also renounced their parent society’s agricultural mode of subsistence, relying heavily on hunting and gathering, and they controlled their population growth by castrating some male infants, so their impact on the non-human environment around them was minimal. In the meantime, the Maori continued to live agriculturally and developed into a populated, complex, hierarchical, and violent society. Eventually an Australian seal-hunting ship informed the Maori of the Moriori’s existence, and the Maori sailed to the Chathams to explore: . . . over the course of the next few days, they killed hundreds of Moriori, cooked and ate many of the bodies, and enslaved all the others, killing most of them too over the next few years as it suited their whim. A Moriori survivor recalled, “[The Maori] commenced to kill us like sheep . . . [We] were terrified, fled to the bush, concealed ourselves in holes underground, and in any place to escape our enemies. It was of no avail; we were discovered and eaten – men, women, and children indiscriminately.” A Maori conqueror explains, “We took possession . . . in accordance with our customs and we caught all the people. Not one escaped. Some ran away from us, these we killed, and others we killed – but what of that? It was in accordance with our custom.” Furthermore, we can deduce from the ubiquitous slavery in all the so-called “great civilizations” like Rome or Egypt that any attempt to rebuild a similar civilization will involve slavery. And to rebuild industry, something similar to colonization and the Trans-Atlantic Slave Trade would probably have to occur once again. After all, global chattel slavery enabled the industrial revolution by financing it, extracting resources to be accumulated at sites of production, and exporting products through infrastructure that slavery helped sustain. So, if industrial society collapsed, who would be doing the rebuilding? Not anyone most people like. It is hard to get a man to willingly change his traditional way of life; even harder when his new life is going into mines. And though history demonstrates that acts like those of the Maori or slave traders are not beyond man’s will or ability, certainly most in industrial society today would not advocate going through the phases required to reach the industrial stage of development.

#### Empirics and worse disasters disprove their impact.

Eken 17 [Mattias Eken – PhD student in Modern History at the University of St Andrews. “The understandable fear of nuclear weapons doesn’t match reality”. 3/14/17. <https://theconversation.com/the-understandable-fear-of-nuclear-weapons-doesnt-match-reality-73563>] Recut Justin

Nuclear weapons are unambiguously the most destructive weapons on the planet. Pound for pound, they are the most lethal weapons ever created, capable of killing millions. Millions live in fear that these weapons will be used again, with all the potential consequences. However, the destructive power of these weapons **has been vastly exaggerated**, albeit for good reasons. Public fear of nuclear weapons being used in anger, whether by terrorists or nuclear-armed nations, has risen once again in recent years. **This is** in no small part **thanks to the current political climate** between states such as the US and Russia and the various nuclear tests conducted by North Korea. But whenever we talk about nuclear weapons, it’s easy to get carried away with doomsday scenarios and apocalyptic language. As the historian Spencer Weart once argued: “**You say ‘nuclear bomb’ and everybody immediately thinks of the end of the world.**” Yet the means necessary to produce a nuclear bomb, let alone set one off, remain incredibly complex – and while the damage that would be done if someone did in fact detonate one might be very serious indeed, **the chances that it would mean “the end of the world” are vanishingly small**. In his 2013 book Command and Control, the author Eric Schlosser tried to scare us into perpetual fear of nuclear weapons by recounting stories of near misses and accidents involving nuclear weapons. One such event, the 1980 Damascus incident, saw a Titan II intercontinental ballistic missile explode at its remote Arkansas launch facility after a maintenance crew accidentally ruptured its fuel tank. Although the warhead involved in the incident didn’t detonate, Schlosser claims that “if it had, much of Arkansas would be gone”. But that’s not quite the case. The nine-megaton thermonuclear warhead on the **Titan II** missile had a blast radius of 10km, or an area of about 315km². The state of Arkansas spreads over 133,733km², meaning the weapon **would have caused destruction across 0.2% of the state.** That would naturally have been a terrible outcome, but certainly not the catastrophe that Schlosser evokes. Claims exaggerating the effects of nuclear weapons have become commonplace, especially after the September 11 terrorist attacks in 2001. In the early War on Terror years, Richard Lugar, a former US senator and chair of the Senate Foreign Relations Committee, argued that terrorists armed with nuclear weapons pose an existential threat to the Western way of life. What he failed to explain is how. It is by no means certain that a single nuclear detonation **(or even several)** would do away with our current way of life. Indeed, **we’re still here despite having nuked our own planet more than 2,000 times** – a tally expressed beautifully in this video by Japanese artist Isao Hashimoto). While the 1963 Limited Test Ban Treaty forced nuclear tests underground, **around 500 of** all **the nuclear weapons detonated were unleashed in the Earth’s atmosphere**. This includes the world’s largest ever nuclear detonation, the 57-megaton bomb known as **Tsar Bomba**, detonated by the Soviet Union on October 30 1961. Tsar Bomba was more than 3,000 times more powerful than the bomb dropped on Hiroshima. That is immense destructive power – but as one physicist explained, **it’s only “one-thousandth the force of an earthquake, one-thousandth the force of a hurricane”.** The Damascus incident proved how incredibly hard it is to set off a nuclear bomb and the limited effect that would have come from just one warhead detonating. Despite this, some scientists have controversially argued that an even limited all-out nuclear war might lead to a so-called nuclear winter, since the smoke and debris created by very large bombs could block out the sun’s rays for a considerable amount of time. To inflict such ecological societal annihilation with weapons alone, we would have to detonate hundreds if not thousands of thermonuclear devices in a short time. Even in such extreme conditions, the area actually devastated by the bombs would be limited: for example, **2,000 one-megaton explosions with a destructive radius of five miles each would directly destroy less than 5% of the territory of the US**. Of course, if the effects of nuclear weapons have been greatly exaggerated, there is a very good reason: since these weapons are indeed extremely dangerous, any posturing and exaggerating which intensifies our fear of them makes us less likely to use them. But it’s important, however, to understand why people have come to fear these weapons the way we do. After all, nuclear weapons are here to stay; they can’t be “un-invented”. If we want to live with them and mitigate the very real risks they pose, we must be honest about what those risks really are. Overegging them to frighten ourselves more than we need to keeps nobody safe.

#### Err negative on impact weighing – their evidence is unwarranted pessimism – updated models.

Rodriguez 20 [Luisa Rodriguez is research fellow at the Forethought Foundation for Global Priorities Research. Previously, she researched nuclear war at Rethink Priorities and as a visiting researcher at the Future of Humanity Institute, "What is the likelihood that civilizational collapse would directly lead to human extinction (within decades)? - EA Forum", 24th Dec 2020, <https://forum.effectivealtruism.org/posts/GsjmufaebreiaivF7/what-is-the-likelihood-that-civilizational-collapse-would#Concrete_example__A_large_nuclear_war_that_causes_a_nuclear_winter//imp>]

Case 2: 90% population loss, infrastructure damage, and extreme climate change (e.g. nuclear war that caused nuclear winter) In a scenario in which a catastrophe causes the deaths of 90% of the population (800 million survivors), major infrastructure damage, and climate change — for example, a severe, global nuclear war that caused a nuclear winter — I believe the question of whether humans would be able to meet their basic needs becomes more difficult.[14] The questions I consider for this scenario are: What is the likelihood that survivors are able to continue to survive using traditional forms of agriculture, given a catastrophe that causes severe infrastructure damage and climate change? What is the likelihood that radiation causes extinction? What is the likelihood that humanity would survive in the event of conflict immediately following the catastrophe? What is the likelihood that survivors are able to continue to survive using traditional forms of agriculture? Time spent on this section: 2–3 hours Types of sources: Academic literature, non-academic reports, and expert interviews Expert judgment: Several experts, including ALLFED director David Denkenberger, have affirmed this conclusion — they do not expect humanity to dip below the minimum viable population even in relatively extreme sun-blocking scenarios. Literature review: The nature of all of the catastrophes we know of that would cause extreme global cooling (e.g. nuclear winter, asteroid impacts) **would have unevenly distributed impacts** — causing extreme global cooling in some parts of the world, but more moderate cooling in others. For example, in the case of a nuclear war between the US and Russia, nuclear winter models suggest that the most **severe climate effects would be limited** to the Northern Hemisphere, where temperatures would fall by 10–30 degrees C. But in the Southern Hemisphere, and especially at the equator, those effects would be much less severe: between 5–10 degrees Celsius. With heterogeneous impacts like this, it’s likely that agriculture would still be possible in some regions — especially in New Zealand and Australia, and possibly in South America and Central Africa.[15] To be clear, I’m describing a very grim scenario, in which basically everyone in the Northern Hemisphere — and in many parts of the Southern Hemisphere — would be unable to grow food using standard agricultural techniques. Given this, I expect there would be mass starvation and violent competition and conflict until a new equilibrium was reached, one where the remaining survivors didn’t exceed the Earth’s carrying capacity. While I expect this would be a truly terrible period of widespread suffering, I believe this equilibrium would be reached long before the population got anywhere near the minimum viable population. My best guess is the population would fall to hundreds of thousands to tens of millions, but not much lower. While I haven’t looked into this much, I feel fairly convinced that hundreds of thousands or **millions** of people **could survive** using traditional approaches to agriculture in parts of the world with more moderate climate effects (and basic mitigation strategies, like switching to crop types that are more resilient to temperature and precipitation fluctuations). And as with Case 1, at least some of the survivors in a Case 2 scenario would probably be able to survive the immediate aftermath of a catastrophe that caused civilizational collapse by exploiting food and other supplies in stores and larger stockpiles. This would give survivors some buffer time to learn additional skills required to survive once those supplies run out (e.g. fishing) or develop the techniques necessary to produce food using methods that don’t rely on climate factors like warm temperatures and regular precipitation. BOTEC: The longer the buffer time, the more likely humanity would be to subsequently survive. But there are a number of different considerations (relative to Case 1) that affect the calculus of just how long such a grace period would be in the context of a catastrophic event like a nuclear war that killed 90% of people and caused a nuclear winter. So I’ve done a similar exercise to the one above where I try to account for some of those differences. Note: As above, the following BOTEC relies on particularly poor sources, makes a bunch of dubious assumptions (discussed more below), and I’m not confident I’ve thought of all of the most important supplies. It should be considered very rough. TABLE5 See table note here.[16] Bottom line: I think it’s extremely likely that these supplies would last somewhere between around a year and a decade or more. I expect it would be closer to the lower end, given that competition and violence could lead to the depletion of supplies more quickly than if the population were reduced to a smaller number by the catastrophe directly. All this in mind, I think it is very likely that the survivors would be able to learn enough during the grace period to be able to feed and shelter themselves ~indefinitely. What is the likelihood that radiation causes extinction? Time spent on this section: 2–3 hours Types of sources: Academic papers, Wikipedia, and interviews with experts Literature review: In the aftermath of a nuclear war, radioactive fallout from the nuclear detonations would have long-lasting health impacts. In **the most extreme** nuclear war **scenario**s considered by academics (a nuclear war between the US and Russia and their allies, using 10,000 megatons (MT) of nuclear bombs), approximately 30% of the geographic area in the Northern Hemisphere would have enough fallout to be lethal to any adult in the area (Ehrlich et al., 1983). The current US and Russian nuclear arsenals don’t currently have that kind of megatonnage (they currently have closer to 2,500 MT). If we naively assume that radiation scales linearly, we might expect a modern day US-Russia nuclear war to contaminate up to 7.5% of the land area of the Northern Hemisphere. This may not sound like much, but consider that 95% of the world’s population lives on just 10% of its land area — meaning that 7.5% of land area could be home to millions or even billions of people. What’s more, tens to hundreds of millions more might be exposed to enough radiation to be more susceptible to cancer for the rest of their lives. On top of this, there are currently around 440 civilian nuclear power reactors scattered around the world, and likely tens or hundreds more military reactors. These have fail-safes and automatic shut down measures that are designed to ensure that all of the nuclear material in these reactors would be safely contained in the event of a global catastrophe that meant people stopped attending to them. Concretely, these fail safes make sure that water continues to be circulated around the nuclear fuel to ensure it doesn’t get so hot it causes a meltdown — i.e., an event where the nuclear core partially or completely melts, which might allow the nuclear fuel to breach its multiple layers of containment and leak out into the environment. If fuel did reach the environment, the radioactive fallout could spread across continents, creating exposure levels ranging from immediately fatal (in areas ranging from tens to thousands of square kilometers) to non-lethal but causing potential higher rates of cancer and infertility. But some of these fail-safes could plausibly fail during a catastrophe that caused infrastructure damage (or afterward, if any components of the fail system degraded). For example, some nuclear reactors rely on backup generators to power the pumps that keep water circulating in the core of the reactor. If those backup generators eventually all broke down, the reactor might melt down. I currently don’t have a good sense of how likely these failures would be. Newer nuclear reactors rely on more robust safety systems, with parts that wouldn’t break down as easily. And all nuclear reactor safety systems are designed to account for infrastructure damage caused by earthquakes and other physical shocks. But in a large-scale nuclear war, it seems very plausible that at least some nuclear reactors would melt down. My best guess is that this wouldn’t happen at a large scale, but even if it did, some areas would likely be far enough away from reactors to be spared the radioactive contamination. For example, Australia has just one nuclear reactor. Even if that reactor were to melt down, much of Australia would likely remain uncontaminated (Australia is just under 3 million square miles, and the Chernobyl meltdown is estimated to have contaminated under 60,000 square miles; and only a much smaller fraction of that area was sufficiently contaminated as to be lethal to humans). Bottom line: While radioactive fallout from nuclear detonations and power plant meltdowns would increase the death toll in the years following the collapse, I expect it **wouldn’t be** widespread enough to be immediately **fatal to everyone**, nor would it cause fertility rates or life expectancy to decrease enough to threaten extinction. And at the very least, **some** areas **are sufficiently far away as to be** relatively **safe** from radioactive fallout. What is the likelihood that humanity would survive in the event of conflict immediately following the catastrophe? Time spent on this section: 1–2 hours Types of sources: Academic literature, expert interviews, and speculation Historical base rate: In Case 2, it seems slightly more plausible to me that violence would lead to human extinction than in Case 1, but still fairly unlikely. I don’t think human extinction could be caused by a conflict fought with conventional weapons; **there would** just **be** **too many survivors (~800 million)** to be killed in conventional warfare (compare this to WWI and WWII, during which ~20 million and ~75 million people were killed, respectively). Weapons of mass destruction: My best guess is that the only way violence in the wake of a Case 2 civilizational collapse could directly lead to human extinction is if one group of **survivors** had access to and deployed weapons of mass destruction. This seems unlikely to me, first because it seems hard to imagine a group of survivors incapable of recovering critical infrastructure — and barely capable of meeting even their basic needs — would be able to successfully deploy weapons of mass destruction (though I’m not very confident about this). Second, it’s hard to imagine a scenario where the use of weapons of mass destruction kills millions of survivors, spread all over the world, without modern technologies like transportation. For example, with potentially many survivor groups, it seems hard to imagine how nuclear detonations would kill ~everyone despite the fact that the groups would likely be spread out all over the world, potentially in small bands that can’t each be individually targeted. Similarly, it’s hard to imagine how a pathogen could spread ~everywhere when survivors would likely have greatly reduced mobility (the latter isn’t obviously impossible, but it at least seems exceedingly difficult to me). There’s one counterargument I find somewhat persuasive, which is that it seems possible that all of the survivors might be confined to a relatively small area (for example, if only a small fraction of the Earth’s land area is habitable), making them more vulnerable to a single, large attack. If this were the case, it’s easier for me to imagine that the use of weapons of mass destruction could kill all of the remaining survivors. This would presumably mean the aggressors would be killing themselves, which makes it seem even less likely to me. But we’ve seen humans come dangerously close to threatening their own survival before, often because human aggressors aren’t always good at predicting how cascading effects could threaten their survival as well. A random example to make this concrete: If all of the survivors of a nuclear war were confined to Australia, which might be less impacted by a nuclear winter, one group might choose to use nuclear weapons against another group, not realizing that the radioactive fallout or further climate change could make Australia uninhabitable, even for them. Bottom line: I expect the survivors in Case 2 would not deploy weapons of mass destruction against their competitors, as it would likely pose a pretty big risk to the aggressor as well as the target. But I’m uncertain about this — humans have come close to making similarly self-destructive choices before. Thankfully, even if one group did use weapons of mass destruction against their competitors, I still think it’s very unlikely that their use would cause human extinction. This is because except in a few very specific and very strange scenarios, I expect the survivors would be too geographically distributed and disconnected to be wiped out by a single act of aggression. I therefore expect the result would be a much higher death toll, but not extinction. Concrete example: A large nuclear war that causes a nuclear winter So what, concretely, do I think would happen in the event of a catastrophe like a nuclear war that led to the death of 90% of the population, and caused severe infrastructure damage and significant global cooling? I expect that, in addition to the billions of people killed in the initial catastrophe, hundreds of millions or more would likely die in the famines and violent competition that followed. But my best guess is that hundreds of thousands to hundreds of millions of the survivors of the initial catastrophe would survive this violent period. I think it’s extremely likely these survivors would be able to support themselves using leftover food stocks and supplies, before eventually working out how to feed themselves through traditional agriculture and fishing and/or modified agriculture (using methods that don’t rely on climate factors like warm temperatures and regular precipitation). **All of the catastrophes** we know of **that would lead to extreme cooling** would only do so **for** 1–**10 years, and agriculture would become possible again once the climate began to return to normal**. At that point, it seems even more likely that the surviving humans would be able to meet their own basic needs by returning to traditional forms of agriculture. My key uncertainties are around whether I’m putting too much weight on the idea that humans would figure out how to subsist without traditional agriculture just because it’s technically possible, and whether conflict could lead to extinction through channels I haven’t foreseen. Another toy calculation suggests that these **uncertainties** probably **aren’t troubling enough to change my bottom line**. Note: I again assume each group’s fate is independent of the fates of other groups. I actually think this is a pretty reasonable assumption in this case. I expect that the **survivors** of a catastrophe like a severe nuclear war **would end up somewhat spread out** (at least across the Southern Hemisphere), as doing so would create less competition for resources within a smaller area (I discuss this more later). The farther apart the surviving groups are, the less likely they are to be affected by the same shocks (natural disasters, disease outbreaks, conflict). Additionally, in the event of a catastrophe like a nuclear war, transportation, communication, and other technologies that facilitate contact between geographically distributed groups would be enormously limited. This would further limit the extent to which each group’s fate ended up relating to another’s. There would be other sources of variation between groups that made their fates less correlated: Some groups might be made up mostly of farmers, while others will be made up of lawyers, some groups will tend toward cooperation, while others toward conflict, plus pure randomness (e.g. some groups might have a high proportion of survivors with genetic immunity to a particular disease). But there are also factors that point in the other direction — factors that suggest the surviving groups would be at least somewhat correlated. For example, nuclear winter climate conditions, while nonuniform, would nonetheless impact all surviving groups. Similarly, more severe natural disasters might affect large regions, meaning that at least all of the survivor groups at the regional level might end up experiencing very similar challenges to survival simultaneously. Likewise, there might be things about "human nature" that would be shared amongst all survivors. For example, it’s possible that all of the survivors, having witnessed the initial catastrophe, would have similar psychological experiences — like shock, stress, and social distrust, among others — that would make it more difficult to survive and cooperate. As above, the higher the true correlation between survivor groups, the more my toy calculations will cause me to underestimate the probability that all of the survivor groups would be wiped out. TABLE6 With 800 million survivors, the degree of pessimism you have to have about their ability to survive to end up believing that no groups would survive indefinitely is actually kind of extreme. The exact beliefs you’d have to have would depend on whether survivors were concentrated into a few big groups, or distributed in many smaller ones. Specifically: Even if you thought any given group of 100, 1,000, or 10,000 survivors had a 99% chance of being wiped out, it would still be virtually guaranteed that at least one group would survive. If you thought there was a 99% chance that any one of 800 groups of 100,000 people would be wiped out, there would still only be a 1 in 3,000 chance of extinction. The probability of extinction is higher (45%) if you believe that larger groups of 10 million would also have a 99% chance of being wiped out. But, again, to hold that view, you’d have to think that out of a group of 10 million people (again, bigger than the largest US city), not even a few hundred of those people would overcome the obstacles of the post-collapse environment (how to fish, how to farm despite global cooling, avoiding being killed by a hurricane or drought). I do not find this view very plausible. Similarly, the probability of extinction is very high indeed if you think that any given group of 100 million survivors has a 99% chance of being wiped out. Again, to believe extinction risk was that high, you’d have to think that there would be a 99% chance that none of the 100 million people would work out how to survive (for reference, only 14 countries have a population of 100 million or higher). Given all of this, my subjective judgment is that **it’s very unlikely that this scenario would more or less directly lead to human extinction.**

#### Isolated island populations repopulate.

Turchin and Green 18 [Alexey Turchin – Scientist for the Foundation Science for Life Extension in Moscow, Russia, Founder of Digital Immortality Now, author of several books and articles on the topics of existential risks and life extension. Brian Patrick Green – Director of technology ethics at the Markkula Center for Applied Ethics, teaches AI ethics in the Graduate School of Engineering at Santa Clara University. <MKIM> “Islands as refuges for surviving global catastrophes”. September 2018. DOA: 7/20/19. <https://www.emerald.com/insight/content/doi/10.1108/FS-04-2018-0031/full/html?fullSc=1&mbSc=1&fullSc=1>] Recut Justin

Different types of possible catastrophes suggest different scenarios for how survival could happen on an island. What is important is that the island should have properties which protect against the specific dangers of particular global catastrophic risks. Specifically, different islands will provide protection against different risks, and their natural diversity will contribute to a higher total level of protection: **Quarantined island survives pandemic**. An island could impose effective quarantine if it is sufficiently remote and simultaneously able to protect itself, possibly using military ships and air defense. **Far northern aboriginal people survive an ice age**. Many far northern people have adapted to survive in extremely cold and dangerous environments, and under the right circumstances could potentially survive the return of an ice age. However, their cultures are endangered by globalization. If these people become dependent on the products of modern civilization, such as rifles and motor boats, and lose their native survival skills, then their likelihood of surviving the collapse of the outside world would decrease. Therefore, preservation of their survival skills may be important as a defense against the risks connected with **extreme cooling**. Remote polar island with high mountains survives brief global warming of median surface temperatures, up to 50˚C. There is a theory that the climates of planets similar to the Earth could have several semi-stable temperature levels (Popp et al., 2016). If so, because of climate change, the Earth could transition to a second semi-stable state with a median global temperature of around 330 K, about 60˚C, or about 45˚C above current global mean temperatures. But even in this climate, **some regions of Earth could still be survivable for humans**, such as the Himalayan plateau at elevations above 4,000 m, but below 6,000 (where oxygen deficiency becomes a problem), or on polar islands with mountains (however, global warming affects polar regions more than equatorial regions, and northern island will experience more effects of climate change, including thawing permafrost and possible landslides because of wetter weather). In the tropics, the combination of increased humidity and temperature may increase the wet bulb temperature above 36˚C, especially on islands, where sea moisture is readily available. In such conditions, proper human perspiration becomes impossible (Sherwood and Huber, 2010), and there will likely be increased mortality and morbidity because of tropical diseases. If temperatures later returned to normal – either naturally or through climate engineering – **the rest of the Earth could be repopulated**. ‘‘Swiss Family Robinsons’’ survive on a tropical island, unnoticed by a military robot ‘‘mutiny’’. Most AI researchers ignore medium-term AI risks, which are neither near-term risks, like unemployment, nor remote risks, like AI superintelligence. But a large drone army – if one were produced – could receive a wrong command or be infected by a computer virus, leading it to attack people indiscriminately. Remote islands without robots could provide protection in this case, allowing survival until such a drone army ran out of batteries, fuel, ammunition or other supplies: Primitive tribe survives civilizational collapse. The inhabitants of **North Sentinel Island**, near the Andaman Islands in the Indian Ocean, are hostile and uncontacted. **The Sentinelese survived the 2004 Indian Ocean tsunami apparently unaffected** (Voanews, 2009), and if the rest of humanity disappear, **they might well continue their existence without change.** Tropical Island survives extreme global nuclear winter and glaciation event. Were a **nuclear**, bolide impactor or volcanic “**winter**” scenario to unfold, these islands would remain surrounded by Warm Ocean, and local volcanism or other energy sources might provide heat, energy and food. Such island refuges may have helped life on Earth survive during the **“Snowball Earth”** event in Earth’s distant past (Hoffman et al., 1998). Remote island base for project “Yellow submarine”. Some catastrophic risks such as a gamma ray burst, a global nuclear war with high radiological contamination or multiple pandemics might be best survived **underwater in nuclear submarines** (Turchin and Green, 2017). However, after a catastrophe, the submarine with survivors would eventually need a place to dock, and an island with some prepared amenities would be a reasonable starting point for rebuilding civilization. Bunker on remote island. For risks which include multiple or complex catastrophes, such as a bolide impact, extreme volcanism, tsunamis, multiple pandemics and nuclear war with radiological contamination, **island refuges could be strengthened with bunkers**. Richard Branson survived hurricane Irma on his own island in 2017 by seeking refuge in his concrete wine cellar (Clifford, 2017). Bunkers on islands would have higher survivability compared to those close to population centers, as they will be neither a military target nor as accessible to looters or unintentionally dangerous (e.g. infected) refugees. These bunkers could potentially be connected to water sources by underwater pipes, and passages could provide cooling, access and even oxygen and food sources.

#### They solve any impact from the aff but not ours – empirics.

Turchin and Green 18 [Alexey Turchin – Scientist for the Foundation Science for Life Extension in Moscow, Russia, Founder of Digital Immortality Now, author of several books and articles on the topics of existential risks and life extension. Brian Patrick Green – Director of technology ethics at the Markkula Center for Applied Ethics, teaches AI ethics in the Graduate School of Engineering at Santa Clara University. <MKIM> “Islands as refuges for surviving global catastrophes”. September 2018. DOA: 7/20/19. <https://www.emerald.com/insight/content/doi/10.1108/FS-04-2018-0031/full/html?fullSc=1&mbSc=1&fullSc=1>] Recut Justin

In recent decades, researchers have identified many global risks which could result in the collapse of civilization and/or human extinction. These risks may be divided in three classes: natural, like asteroid risks (Gritzner et al., 2006); low-tech, connected with currently existing technologies, like the risk of nuclear war (Barrett et al., 2013); and futuristic hightech, connected with new expected super technologies, like nanotechnology (Freitas, 2000) and artificial intelligence (AI) (Bostrom, 2014). **Super technological risks are the most dangerous**, as they are expected to be the most powerful and least controllable (Bostrom, 2002). The best way to fight all the types of risks is to prevent (Bostrom, 2013) or mitigate them, **but another option, or plan B, is to adapt to them to survive them**. There are several ideas for how such risks could be survived, including a Mars colony (Musk, 2017), a Moon colony (Shapiro, 2009; Turchin and Denkenberger, 2018), underground bunkers (Jebari, 2014), space bunkers (Torres, 2016) and retrofitted nuclear submarines [which are one of the most cost-effective solutions (Turchin and Green, 2017)]. Planning for surviving these risks, whether by mitigation or adaptation, should be a paramount ethical duty of humankind (Jonas, 1984; Green, 2014, 2016). Several authors (Jebari, 2014; Beckstead, 2015) have analyzed the problem of global risk survival and concluded that most catastrophes are either too small or too large for bunkers or other refuges to be a useful option. But **even a 1 per cent increase in the chance of survival is worth considering,** especially because there are not yet useful working ideas of the magnitude of some larger risks, such as unsafe AI (Bostrom, 2014). Additionally, at the workshop on existential risk to humanity (Gothenburg Chair Programme for Advanced Studies, 2017), Karim Jebari mentioned that such refuges will also be important for cultural transfer and as consolidation points, even if there might be many survivors in other places. Baum has suggested that the gold standard for global risks refuges should be “surface independence” (Baum et al., 2015). Islands only partly satisfy this criterion: they are not connected to the mainland, thus making them discontinuous with the land surface of the Earth, but they are still accessible by air and sea. However, if they were very remote and equipped with underground and/or underwater shelters, they could provide a higher level of protection than surface-independent bunkers on the mainland for certain types of catastrophes. By definition, global risks affect much or all the surface of the Earth, or at least all populated areas. This creates a chance for survival, as **there is a probability that some parts of the Earth will be affected to a lesser extent**. For example, a gamma ray burst (Cirkovi c and Vukotic, 2016 ) that happened away from the equatorial plane would have less of an effect on one of the polar regions. Likewise, extreme global warming (Hanna and Tait, 2015) would be more survivable on mountains at high latitudes, while atmospheric pollution (Mount, 1970) by some toxin or contamination could be less of an issue in the Southern hemisphere because of geography and atmospheric circulation patterns. Yet, most catastrophes which could be survived on temporary space refuges on the Moon or Mars could also be survived on Earth, if there were adequate shelters or refuges, with some notable exceptions. Such exceptions include very large asteroid impacts, a severe and long-term case of multiple pandemics (with many lethal diseases active in the environment) or massive and irreversible global warming. For some preliminary calculations of the usefulness of shelters from global catastrophes see Turchin and Green (2017). Islands have proven to be survival refuges for some species which are extinct in other places, like mammoths, which survived on Wrangel island up to 2000 BC (Vartanyan et al., 1995). **Islands have proven to be effective refuges for humans as well**. For example, the islands of New Caledonia and American Samoa did not have a single death from the 1918 Spanish flu because of their **effective quarantine measures** (Bell et al., 2006). While islands have been extensively discussed as refuges for animals and plants, the topic of using islands as a means for humans to survive global catastrophic risks has not yet been formally explored. This article seeks to remedy this deficiency. Section 2 looks at the requirements for survival on islands; Section 3 looks at the possible role of islands as consolidation centers after a social collapse; Section 4 reviews several islands as possible refuges; Section 5 puts island refuges in the context of other possible types of refuges; Section 6 discusses how to maximize protection by combining islands refuge with subterranean and/or submarine refuges; and Section 7 discusses other places on Earth, similar to islands, where survival might be possible. Islands offer excellent protection against natural and/or **low-tech catastrophes** which are neither too large nor too small. Remoteness, isolation and the diverse conditions found on different islands could be helpful features to aid survival in the face of different types of catastrophes. Islands could provide protection against a human-to-human transmitted biological pandemic; as mentioned in the Introduction, some islands were able to escape the 1918 flu pandemic by implementing effective quarantine measures. Islands may help to survive a **long-term collapse in food production caused by nuclear winter**, agricultural pests and other catastrophes. Islands often have **non-traditional food sources**, such as birds and sea flora and fauna, which may provide independent subsistence for an indefinitely long period. On remote islands, **the extent of radioactive and chemical contamination from catastrophes would likely be smaller**. This is especially true of islands located in the Southern hemisphere close to the Antarctic, as **winds around the pole maintain some isolation from the rest of the atmosphere**. **Constant rains and winds may accelerate the decontamination of some islands** (like Kerguelen). In addition, **sea animals may be relatively less contaminated food sources**. Islands away from the equator could provide protection against some of the direct effects of a gamma ray burst (muons) (Cirkovi c and Vukoti c, 2016 ) if they were in the constant shadow of the Earth, below the horizon of the gamma ray source. In the case of global war or technological collapse, **many islands could become unreachable**. This would protect them against human-borne diseases, pirates, looters and certain autonomous weapon systems such as land-based or short-range drones. Additionally, remote and sparsely populated islands may not be interesting military targets. In case of war, it may be more expensive to reach them than to ignore them, though this depends on the nature of the war. For example, the Germans used remote unpopulated islands in the Arctic (Grossman, 2016) and in the Southern Ocean (Rogge and Frank, 1956) as secret bases during Second World War, and the allies later sent cruisers to Kerguelen to check if Germans were hiding there. It might be too expensive for a hostile AI to seek out and kill small groups of people in remote places, if they do not pose an immediate risk to the AI’s interests. However, over time, the AI’s risk calculation might change.

#### French Kerguelen Islands have unique characteristics conducive to repopulation.

Turchin and Green 18 [Alexey Turchin – Scientist for the Foundation Science for Life Extension in Moscow, Russia, Founder of Digital Immortality Now, author of several books and articles on the topics of existential risks and life extension. Brian Patrick Green – Director of technology ethics at the Markkula Center for Applied Ethics, teaches AI ethics in the Graduate School of Engineering at Santa Clara University. <MKIM> “Islands as refuges for surviving global catastrophes”. September 2018. DOA: 7/20/19. <https://www.emerald.com/insight/content/doi/10.1108/FS-04-2018-0031/full/html?fullSc=1&mbSc=1&fullSc=1>] Recut Justin

One of the most attractive islands for long-term survival of global risks is **the French archipelago of Kerguelen** in the southern Indian Ocean. Kerguelen’s main Grand Terre Island has the following attractive features for long-term survival: It **is very remote from any other constant human settlement**s; for example, is it 3,000 km from the island of Reunion. The Kerguelen Islands lie **outside the main trade lines**, so the probability of a random ship arriving there is low. **The islands are inside the circumpolar Antarctic current**, and they are surrounded by strong winds (the “Roaring Forties” and “Furious Fifties”), which will not accidentally bring any ships from further north. A return trip from Reunion to Kerguelen by ship takes 28 days. **The islands do not have an airport**, so they cannot be reached by air, **and they are too remote for helicopter travel.** While Easter Island is even more remote from other human settlements, it is more populated and more often accessed by ships and planes. **The intense and isolating wind circulation around the South Pole could increase the time required for ash or radioactive clouds** from the northern hemisphere **to reach** the South Polar Region. But the Kerguelen Islands are also not too close to the South Pole: they are at the equivalent latitude as southern Germany; thus, they get quite a bit of sunlight The Kerguelen Islands have a stable but cold climate, with temperatures above freezing most of the time. The main island has **edible vegetation and many edible animals**, including 3,000 sheeps. The island is very large, approximately 7,000 km2 , and **it has many deep gulfs and fjords that could be used as harbors**. The main island has high mountains (over 1,000 m) with **an ice cap which could provide fresh water**. Nearby ice-free mountains hundreds of meters high could provide **protection against tsunamis**. The highest mountain is volcanic, and was active 100,000 years ago (Weis et al., 1998). However, **residual geothermal heat could provide heating and energy for a refuge**. The main island has a continuous population of only about 45 people, who live at a scientific station. Scientists who are selected for long expeditions are **more organized and educated than random people, so they may be better prepared for survival**. Such a scientific base will not be a military target in case of war. There are several other South Ocean islands similar to Kerguelen, like South Georgia, Auckland Island and Macquarie Island (Schalansky, 2010).