

The Problem of Evil and the Simulation Hypothesis

CHRISTOPHE FACAL*

I argue that the Problem of Evil diminishes the likelihood of the simulation hypothesis being true. In section 1, I present the simulation hypothesis. In section 2, I present the Problem of Evil and how it can be addressed to the simulation hypothesis. In section 3, I strengthen my claim by defending the idea that, on the grounds of the expansion of the moral circle, we can infer that potential simulators would most probably view simulated beings as moral subjects. Section 4 is dedicated to counterarguments and their responses, and section 5 contains a conclusion where I clarify the probabilistic nature of my argument.

Keywords: *Problem of Evil, Simulation Hypothesis, Moral circle expansion, simulation theology*

Introduction

I argue that the Problem of Evil diminishes the likelihood of the simulation hypothesis being true. In section 1, I will present the simulation hypothesis. In section 2, I will present the Problem of Evil and how it can be addressed to the simulation hypothesis. In section 3, I will strengthen my claim by defending the idea that, on the grounds of the expansion of the moral circle, we can infer potential simulators would most probably view simulated beings as moral subjects. Section 4 is dedicated to counterarguments and their responses, and section 5 contains a conclusion where I clarify the probabilistic nature of my argument.

1. The Simulation Hypothesis

The simulation hypothesis was first formulated in a 2003 article by Nick Bostrom, and most recently by David Chalmers. I will present both versions of the simulation hypothesis.

* Christophe Facal, PhD Candidate in Philosophy, McGill University, 845 rue Sherbrooke Ouest, H3A 0G4, Canada, christophe.facal@mail.mcgill.ca.

1.1 Bostrom's version of the simulation hypothesis

Bostrom's version of the simulation hypothesis takes the form of a triple disjunction. In his original paper, Bostrom specifically talks about *ancestor* simulations, but these are technicalities we do not need to bother ourselves with for our present purpose. Here is thus a simplified version of Bostrom's argument:

1. Either every intelligent civilization disappears before reaching the technological stage required to run large-scale computer simulations.
2. Or, almost every intelligent civilization refuses to run such simulations even though they have acquired the technological means to do so.
3. Or, we might be living in a computer simulation.

His argument ought to be understood in the following way: at least one of the three disjuncts necessarily must obtain. If scenario 1 or 2 obtains, scenario 3 cannot obtain. But, if scenarios 1 and 2 do not obtain, then scenario 3 must obtain. If scenario 1 does not obtain, this means some civilizations will reach the technological stage required to run large-scale computer simulations. If scenario 2 does not obtain, these civilizations will create some simulations. Therefore, we must grant the possibility that our world might be that simulation. The question is: do we have any reason to believe scenarios 1 or 2 will obtain? Let's examine both the first and second disjunct in turn and see if we have any reason to believe they do obtain, blocking the possibility of disjunct 3 obtaining.

1.1.1 Is it likely civilizations end before creating large-scale simulations?

For this first scenario to obtain, all possible civilizations must end before reaching such a technologically advanced state. This implies an event that would necessarily wipe out all possible intelligent civilizations before they reach this stage. Even if it were the case that *our* civilization could disappear before reaching this stage, in order to say *all and any possible* civilization would necessarily suffer the same fate, we would have to identify some sort of Great Filter, a Great Wall or a Civilizational Barrier which necessarily prevents any possible civilization from reaching such a

technological stage. But no such thing has ever been identified, nor do we seem to have any reason to believe such a thing would exist.¹

1.1.2 Is it likely all possible civilizations choose not to run simulations?

For the second disjunct to obtain, every intelligent civilization that would reach the technological state required to run large-scale simulations would have to converge towards the decision not to run such simulations. Moreover, those civilizations would need to be able to enforce this prohibition efficiently, without anyone deciding to run simulations by themselves. Again, this scenario is not impossible, but is it probable? If we can deduce anything from our current fondness for simulations, which we constantly use nowadays for a diverse set of uses, whether they be for medical purposes, meteorological or logistical planning, research, decision-making or entertainment, it is quite unlikely that intelligent and technologically advanced civilizations will decide not to run simulations. Given their usefulness, advanced civilizations will most likely not prohibit simulations. But even if they do prohibit simulation, there is no guarantee that this prohibition can be effectively enforced.

Bostrom's conclusions thus naturally follow: if we have no conclusive reason to say scenarios 1 and 2 would obtain, then scenario 3 remains an open possibility, and we consequently have to grant the possibility that we might be living in a simulation. How likely is this? Bostrom does not say how much credence we should lend to each disjunct, but the credence one ascribes to the simulation hypothesis should be proportional to the credence we lend to each of the three disjuncts.² In other words, Bostrom's argument establishes that if we grant some likelihood to the possibility of simulations being created one day (of disjuncts 1 and 2 not

¹ On existential risks, Nick Bostrom, "Existential Risks. Analyzing Human Extinction Scenarios and Related Hazards," *Journal of Evolution and Technology* (2002): 16, says the following: "But the balance of evidence is such that it would appear unreasonable not to assign a substantial probability to the hypothesis that an existential disaster will do us in. My subjective opinion is that setting this probability lower than 25% would be misguided, and the best estimate may be considerably higher." Toby Ord, *The Precipice: Existential Risk and the Future of Humanity*, (London: Bloomsbury Publishing, 2020), 480, who wrote the most comprehensive book on the probability of existential risks, gives a probability of 1/6 per century that humanity goes extinct. But again, these existential risks concern our civilization, and are not exportable to any possible civilization.

² I thank an anonymous reviewer of RES for this insightful comment.

obtaining), we must grant proportional likelihood to the possibility that we might be living in a simulation (disjunct 3 obtaining).

1.2 Chalmer' version of the simulation hypothesis

Chalmers' version of the simulation hypothesis runs along a different line. Chalmers builds a statistical argument which emphasizes the proportion of non-simulated beings (*nonsim*) versus simulated beings (*sim*):

Any nonsim population that survives long enough will eventually be able to create (let's say) at least a thousand sim populations, each with about as many sims as the original nonsim population. It's arguable that at least one in ten human-level nonsim populations will eventually do this. If one in ten nonsim populations creates at least a thousand sim populations each, this will result in at least a hundred sim populations per nonsim population. If this is right, then throughout the cosmos, sims will outnumber nonsims by at least a hundred to one.³

Therefore, it is more likely that we are simulated beings. This version of the argument is more straightforward. It is likely an intelligent civilization will one day be running enormous number of simulations, populated by a large number of simulated beings. If any intelligent civilization does reach that stage, it's not a stretch to think simulated beings (*sims*) will vastly outnumber non-simulated beings (*nonsims*): the fact that there are so much more *sims* than *nonsims* commits us to acknowledge we have more chances of being *sims* than *nonsims*. Therefore, there are non-negligible chances we live in a simulation.

1.3 Objections to the simulation hypothesis

We will now briefly review the possible objections to the simulation hypothesis and highlight how they fall short of the mark.

Chalmers places the objections to the simulation hypothesis in two categories: sim-blockers and nonsim signs. Objections of the "sim-blockers" type point to the impossibility of our world being simulated (they *block* the simulation hypothesis). They are, generally speaking, things which would exist in the (hypothetical) universe of our simulators that would prevent them from building simulations. Nonsim signs objection are more of a probabilistic nature: they are things in our universe which seem to point towards the fact that we do not live in a simulation. Sim blockers objections are of the following kind:

³ David Chalmers, *Reality* (New York City: Norton, 2022), 81.

Simulation is impossible, simulation is too difficult, nonsims populations will all die before creating sims, nonsims will choose not to make simulations.⁴ (...) And nonsim sign objections are of the following kind: “Sims can’t be conscious! (...) Simulators will avoid creating sims (...). Sims won’t have minds like ours (...) Sims won’t experience large universes.”⁵

We will here study three objections. The first is a nonsim sign: there are some things in our world that cannot be simulated on a computer, namely consciousness. The second objection is a sim blocker: it is impossible for computers to reach a computational power sufficient to simulate our universe. The third objection is a nonsim sign: our universe is too vast and complex to be a simulation.

1.3.1 Some objects of our world cannot be simulated

This first objection contends that some objects of our world cannot be simulated on a computer. The possibility of simulated objects is easily satisfiable for most objects: it seems pretty easy to simulate a rock or a chair. But things get trickier when we come to objects such as our consciousness. We have failed, up until now, to create an artificial consciousness, and many think this might be impossible due to the nature of consciousness itself.

However, adherents of the simulation hypothesis think consciousness could be simulated. If consciousness depends not on the infrastructure from which it emerges but on the organization of said infrastructure, any system, whether carbon-based and biological or silicone-based and digital, could theoretically be conscious. This view is called reductionist functionalism. It argues for the substrate independence of consciousness, which is only determined by the causal organization of a system, notwithstanding what this system is made of (*functionalist aspect*), and the mind can be reduced to the physical process happening inside the brain (*reductivist aspect*). If a biological brain produces consciousness, it is thanks to how the neurons connect through electric impulses. If a computer were to imitate the internal organization of the brain, the computer, or the software run on said computer, could also be conscious. Even if the reductivist component is removed, a non-reductivist functionalist still holds that a computer could simulate consciousness. They will argue that

⁴ *Ibidem*, 86.

⁵ *Ibidem*, 92–5.

conscious states are not the same thing as the physical states upon which they occur, but they are nonetheless correlated with them to the extent that replicating these physical states with a different infrastructure than a human brain would result in a consciousness identical to that of a biological brain.⁶ In other words, the physical processes inside the mind cannot wholly explain the mind and all its conscious experiences. Therefore, there is no reductionism, but they are nonetheless strongly correlated with them, to the extent that replicating, on another medium, the physical processes found inside the brain, would produce consciousness. If functionalism is adopted, then the problem of simulating consciousness does not seem impossible to overcome, and the objection that it is impossible to simulate a consciousness does not hold.

1.3.2 Computers couldn't possibly be powerful enough

This objection contends that simulating our universe would necessitate computational power way beyond any possible computer. Again, this objection seems misguided.

Granted, we currently do not have any computer powerful enough to run simulations of complete universes populated by conscious beings. However, there seems to be no reason to place an upper bound to the increase of our computational power. Currently, the most powerful supercomputer, OLCF-5, can perform around 10^{18} operations per second, or 1.102 quintillion operations per second. While we do not know precisely what our brains performance in operations per second, the best estimates put it in a similar range⁷: our most powerful computers are likely already powerful enough to match a human brain in terms of computational power. IBM recently announced their first quantum computer. In a few years, the numbers discussed above will probably already be outdated. Quantum computation seems to be on the brink of a vast and rapid expansion.⁸ If such computers are indeed developed, then placing a low upper bound on computational power seems a doubtful move. Of course,

⁶ David Chalmers, *The Conscious Mind. In Search of a Fundamental Theory* (Oxford University Press, 1996).

⁷ Nick Bostrom and Anders Sandberg, "Whole Brain Emulation: A Roadmap," Technical Report #2008-3, Future of Humanity Institute, Oxford University, <https://www.fhi.ox.ac.uk/brain-emulation-roadmap-report.pdf>, accessed December 03, 2022.

⁸ Paul Smith Goodson, "IBM Announces New 400 Qubit Quantum Processor Plus Plans For A Quantum-Centric Supercomputer," *Forbes*, November 9, 2022, <https://www.forbes.com/sites/moorinsights/2022/11/09/ibm-announces-new-400-qubit-quantum->

there are upper limits to computational power.⁹ These limits are generally thought to be imposed by the physical limits of our universe. Still, it is unknown if they occur *before* the computational power required to run large-scale simulations. The possibility of computers reaching the computational power required to run large-scale simulations thus remains open.

1.3.3 Our universe is too vast to be a simulation

This third objection contends that our universe is extraordinarily vast, and this vastness seems to be a nonsim sign. The observable universe is more or less 90 billion light-years wide, could be filled with a trillion trillion stars, and has an astonishing level of complexity, ranging from subatomic particles to supermassive black holes. Whatever the simulation's purpose, a much less detailed and smaller simulation could do the job, and the vastness of our universe would be unnecessary.

There are two possible responses to this objection. First, the next universe up is so much vaster than ours, that it could potentially have the necessary resources to simulate our universe. Second, the largeness of our universe might be illusory, the far-away parts being simulated only when we are observing them, making our simulation a "shortcut simulation". Just as videogames do not generate the whole world at every moment of play, but only the parts of the world the players have access to at the present moment, the simulation could be fine-grained only when a simulated person, namely us, is observing this particular part of the simulation, and as when a tree falls in the forest and no one is there to hear it, or when no one is observing a far-away celestial body, the simulation would spare the computational power necessary to simulate the sound of the tree or the presence of said celestial body.

In this section, we have presented the two main versions of the simulation hypothesis. This hypothesis is defensible, it is plausible, and it seems to have strong answers to every objection that could be presented against it.¹⁰ This being said, I contend it is possible to formulate an ob-

processor-plus-plans-for-a-quantum-centric-supercomputer/?sh=1694191eaa99, accessed December 03, 2022.

⁹ Anders Sandberg, "The Physics of Information Processing Superobjects: Daily Life Among the Jupiter Brains," *Journal of Evolution and Technology* 5, no. 1 (1999): 1–34.

¹⁰ There are some more objections to the simulation hypothesis, but it is outside the scope of this paper to discuss them all. We do however signal the objection by Michael

jection inspired by the Problem of Evil, which reduces its likelihood of being true.

2. The Problem of Evil

Since the simulation hypothesis was first formulated, it has been examined in light of the Problem of Evil to different effects. I will start by presenting what exactly is the Problem of Evil, and then I will show the many ways in which these two ideas have met.

2.1 The Problem of Evil

The Problem of Evil, in a nutshell, goes as follows: If a Deity is all-powerful, all-knowing, and all-good, why is there evil in the world?

Our world is plagued by evils such as infants dying from atrocious diseases and innocent people suffering from natural disasters. Could a God that is omnipotent, omniscient, and omnibenevolent, design such a world? Logically, no. If God is good, then he cannot want those evils to be. If he is omniscient, then he is aware of those evils. If he is omnipotent, then he could eliminate those evils. But these evils haven't been eliminated. Therefore, it follows that God, as described, does not exist, or at least that the probability of his existence is significantly reduced.

Of course, there are evils that result from human actions (murder, torture, genocide, theft, rape, and so on) and evils that are not the result of human actions (natural disasters, illness, death). The Problem of Evil concerns both kinds of evils, and one can thus talk of two kinds of problems of evils: the problem of moral evil and the problem of natural evil. This being said, one could argue the Problem of Evil is more poignant when the focus is placed on natural evils (we will shortly see why). The 1755 Lisbon earthquake is paradigmatic. On the morning of November

Huemer against the brain-in-a-vat hypothesis, who says it is so unlikely, if we were brains in vats, that our sensory experiences would be as they are now, that we can dismiss the brain-in-a-vat hypothesis; Sean Carroll's resolution objection, who says that if we accept the simulation hypothesis, we also have to accept the fact that it is then more likely than not that there are a chain of simulation (a simulation, in a simulation, and so on), and us being in the bottom, because we don't yet have the capacity to run simulations, would have to mean that our universe would have to be of very low resolution, which is not, thus dismissing the simulation hypothesis; and finally Eric Farias' objection who says that because we have access to the smallest possible scale of physical reality, that being the Planck time, this proves we do not live in a simulation, because our supposed simulators would then have to have access to a broader scale than us, which is here impossible, as the Planck time is the theoretically smallest time physically possible.

1st, 1755, a great earthquake hit Lisbon, causing chaos, subsequent fires and tsunamis, and was responsible for the death of around 12,000 people.¹¹ No human action could be blamed for this event, and the only guilty party seemed to be God. This evil didn't seem to serve any purpose. Its occurrence seemed to be a brute fact of our world and happened simply because our world unfolded in accordance with natural laws, in this instance, tectonic laws. If those evils are woven into the natural laws of our world,¹² then the case can be taken up to the designer of our world. And it is hard to see why God would have created the world this way. Therefore, it seemed that either God was not as traditionally described, or that he was not at all.

Theodicies have tried to explain away the Problem of Evil in many different fashions. Free-will theodicies¹³ may work well for moral evils, but their efficacy in dispelling the dangers posed by the problem of natural evils is less certain.¹⁴ Different types of theodicies have thus been developed to try to explain away the existence of evil in our world, without jeopardizing the belief in the existence of God. Classical ones include diabolical and fall theodicies.

Diabolical theodicies place the responsibility of natural evils on demonic forces. Beings more powerful than us yet inferior to God, like Satan, would be responsible for many of the natural evils of our world. As for fall theodicies, they claim that natural evils are the deserved consequences of

¹¹ Joao F. B. D. Fonseca, "A Reassessment of the Magnitude of the 1755 Lisbon Earthquake," *Bulletin of the Seismological Society of America* 110, no. 1 (2020), <https://doi.org/10.1785/0120190198>.

¹² David Kyle Johnson, "The Failure of Plantinga's Solution to the Logical Problem of Natural Evil," *Philo* 15, no. 2 (2013), <https://doi.org/10.5840/philo20121528>.

¹³ Alvin Plantinga's Free Will Defense goes like this: "A world containing creatures who are significantly free (and freely perform more good than evil actions) is more valuable, all else being equal, than a world containing no free creatures at all. Now God can create free creatures, but He can't cause or determine them to do only what is right. For if He does so, then they aren't significantly free after all; they do not do what is right freely. To create creatures capable of moral good, therefore, He must create creatures capable of moral evil; and He can't give these creatures the freedom to perform evil and at the same time prevent them from doing so. As it turned out, sadly enough, some of the free creatures God created went wrong in the exercise of their freedom; this is the source of moral evil. The fact that free creatures sometimes go wrong, however, counts neither against God's omnipotence nor against His goodness; for He could have forestalled the occurrence of moral evil only by removing the possibility of moral good." See Alvin Plantinga, *The Nature of Necessity* (New York: Oxford University Press, 1974), 166–67.

¹⁴ *Ibidem*. See also James Sterba, *Is a Good God Logically Possible?* (Palgrave Macmillan, 2019).

human actions which caused the biblical Fall. Fall theodicies do not need to accept every detail from the biblical account, but generally explain natural evils by “human beings[’s] falling away from right relationship with God, whatever the exact details of that Fall”.¹⁵ However popular these kinds of theodicies were in the past,¹⁶ they do not today garner as much sympathy today. As Dainton points out, “the progress of natural sciences over the past three centuries has been such that there is very little room left for supernatural interventions in the natural order”,¹⁷ and if we have to choose between, say, what geology tells us about the causes of earthquakes and some alternative explanation according to which earthquakes would be caused by the action of malevolent forces, it is more reasonable to accept the former. However commonsensical this critique might seem, not everyone thinks it succeeds in discrediting diabolical theodicies.¹⁸

Modern theodicies employ a different route and generally try to lessen the dangers posed by the problem of Evil by attacking its logical structure. Two major examples are Alvin Plantinga and Peter Van Inwagen.

Plantinga’s attempt at deactivating the Problem of Evil (apart from the one mentioned in footnote 13) consisted in arguing that if we have certain knowledge of both the existence of God and the existence of Evil, the compatibility of those two phenomena need not be proven, for their mere coincidental existence is already a testimony to their compatibility.

Van Inwagen’s strategy consists in “stories”, or defenses, which explain the coexistence of both Evil and God in a way that is not surprising on the theistic hypotheses. In other words, Inwagen’s defense consists in a story which we have no reason to believe is false, which is true for all we know, and which presents a kind of God that is compatible with the evils we see. Since we have no reason to believe his story to be less probable than a God inconsistent with the “patterns of suffering” present in our world, then the problem of Evil loses some of its appeal.

¹⁵ Dustin Crummett, “The Real Advantages of The Simulation Solution to The Problem of Natural Evil,” *Religious studies* 57 (2021): 627, doi:10.1017/S0034412519000726.

¹⁶ *Ibidem*.

¹⁷ Barry Dainton, “Natural Evil: The Simulation Solution,” *Religious studies* 56, no. 2 (2018): 1–22.

¹⁸ Crummett, “The Real Advantages,” 623.

2.2 A way to relate the problem of Evil to the simulation hypothesis.

Dainton, Crummett and Johnson¹⁹ have recently argued for what they call the “simulation solution”. They argue that from a theistic perspective, the simulation hypothesis might be an answer to the Problem of Evil.²⁰ Indeed, if we live in a world designed by simulators who have free will, then the so-called “natural evils” of our world are, in fact, the “moral evils” of our simulators. From then on, a free-will theodicy could explain why they can inflict such suffering on us, the simulated beings, and the idea of an omniscient, omnipotent and omnibenevolent God is safe from the criticism of the Problem of Evil. This might be, as Johnson puts it, “an unpalatable position”²¹ for theists, but if there are evils in our world, and God cannot create worlds with evils, then it seems reasonable to think that God did not create our world, thus hinting at non-divine creators. This solution is characterized by Crummett as a *subsumption theodicy*: “[it] attempts to subsume apparently natural evils under the category of moral evils by showing that apparently natural evils are in fact the result of creaturely wrongdoing.”²²

This is the main way in which the simulation hypothesis and the Problem of Evil have crossed paths: the simulation hypothesis offers a solution to the Problem of Evil. I believe there is a second, more fruitful way of linking the simulation hypothesis and the Problem of Evil.

¹⁹ David Kyle Johnson, “Natural Evil and the Simulation Hypothesis,” *Philo* 14, no. 2 (2011): 161–75. The relevant papers for Dainton and Crummett have already been referenced in footnotes 14 and 16.

²⁰ Bostrom, “Are You Living in a Computer Simulation?,” *Philosophical Quarterly* Vol. 53, No. 211, (2003): 243–55 (First version: 2001), also argues that if we do live in a simulation, the problem of Evil stops being problematic. The fact that we live in a simulation could offer us the following solution: all our memories of pain and suffering are illusions our simulators have implanted in our minds. What Bostrom suggests is that if the simulation hypothesis is true, it opens up the door to a new solution to the problem of evil: “evils (...) just don’t really exist, and they could in fact be illusory memories the simulators have implanted in our minds”. As he concedes himself, this only holds if you are not suffering at the present moment. This idea has not been discussed further in the literature, perhaps because it is indeed difficult to defend: first, painful feelings can be summoned almost at will by reflecting on past painful experiences, thus making the suffering present and not merely a memory, thus establishing the existence of suffering in our world beyond the mere illusion of suffering, and second, it is hard to see why our simulators would give us these illusory memories of pain.

²¹ Johnson, “Natural Evil,” 162.

²² Crummett, “The Real Advantages,” 620.

2.3 How the problem of evil and the simulation hypothesis should interact

Dainton, Crummett and Johnson might be taking the problem backwards. The Problem of Evil turns on a self-evident observation, that there is suffering in the world, whereas the simulation hypothesis isn't as certain as the existence of evil. Because we have more reason to believe that "evil exists" than we have to believe that "we are in a simulation", we would have to grant more credence to our belief in the existence of evil than in our belief of the simulation hypothesis (because we have good empirical reasons to believe there are evils in our world, and comparatively less evidence to believe we live in a simulation). Therefore, if we are to study the interaction between those two ideas, we ought to consider the existence of pointless suffering to be true and to evaluate the simulation hypothesis in light of this. This does not mean that Dainton, Crummett and Johnson are *wrong*, but that they have only explored one side of the coin. They have examined what happens to the Problem of Evil when we take the simulation hypothesis seriously. I will now explore what happens to the simulation hypothesis when we take the Problem of Evil seriously.

In the case of the Problem of Evil as addressed to God, it resulted from this critique that God either did not exist, or that he could not have the characteristics we traditionally attributed to him. I believe it is not necessary for the Problem of Evil to be addressed to a deity. It only ought to be addressed to a world-making figure. If the world they created contains pointless suffering, then they can be held accountable for the creation of such suffering. Furthermore, if the fact that the world they created contains suffering contradicts with some characteristics we can reasonably expect them to have, then we can conclude they do not exist, or that the likelihood of their existence is reduced. If the Problem of Evil can thus be addressed to any world maker, we can address it to the simulators postulated by the simulation hypothesis.

2.4 The Problem of Evil as addressed to the simulation hypothesis.

If we assume we live in a simulation, it seems reasonable to believe the simulators have both vast power over the simulation (near omnipotency), and vast knowledge over it (near omniscience). Adherents of the simulation hypothesis say so themselves!

As Chalmers points out, simulations tend to give the simulator tremendous flexibility and power as to how to structure the simulation and intervene in it. If a simulator has built a simulation from its foundations, developing the entire software running the simulation, then by tampering with its own creation, the simulator can do pretty much anything and everything in the simulation.

As he also points out, simulators will most likely know pretty much everything that is happening inside the simulation. A complete monitoring is possible. What could possibly be the purpose of an inscrutable simulation? No matter what the goal of a simulation is, this goal, whether it be research, entertainment, or decision-making, is much harder to accomplish if the simulation is unreadable (if it is impossible for the simulator to know what is happening inside the simulation). If unreadableness defeats the purpose of any simulation, it's safe to assume a high degree of readability for any simulation.

We can therefore assume that simulators, were they to exist, would have, to a high degree, control and knowledge of their simulation of our universe.²³ The central question of the Problem of Evil naturally follows: given the fact that the simulators are omnipotent and omniscient in regard to their simulation, why are there evils in our world? If they know of these evils and are able to prevent them, why don't they do so?

If two out of the three "omni" characteristics traditionally attributed to God are met (at least to some extent) by the simulators, the evils of our world could only be explained by the lack of the third characteristic: goodness. Therefore, if we live in a simulation, we have to assume the simulators are not good enough to prevent the natural evils of our world. It would not be a stretch, in this case, to call them cruel, and would provide us a simple and straightforward explanation of the Problem of Evil as addressed to the simulation hypothesis: if the simulators know of the evils of our world and are able to stop them, why don't they do so? Because they are not all-good. Or, put simply, they just don't care enough about our suffering to prevent it.

But is that plausible? I will argue it is unlikely they are malevolent. I will argue that, were they to exist, we would have good reason to think

²³ Ultimately, the exact degree of power and knowledge the simulators would have over the simulation is impossible to know precisely. One could claim that I overestimate it, though I believe that it is safe to assume knowledge large enough to know that simulated consciousnesses suffer, and power large enough to prevent this suffering.

they would be benevolent beings, who would, therefore, not fill their simulation with suffering. If that is indeed the case, and they are therefore benevolent, near-omnipotent, and near-omniscient, given the fact that our world is filled with suffering, we can conclude they are most likely not the creators of our world, or in other words, that we do not live in a simulation. My final task is thus to argue for the fact that were they to exist, they would, in addition to being omniscient and omnipotent (about the simulation), also be benevolent beings. I will argue for this using the idea of the expansion of the moral circle.

Before I move on to the third section, allow me a short methodological note. My argument is probabilistic in nature. I contend that given the existence of evil in our world, we either have to accept that our simulators are cruel, or we have to accept that we do not live in a simulation. I think it is not probable they are this cruel, and this, therefore, lowers the probability of the simulation hypothesis being true. At this point, adherents of the simulation hypothesis would have to explain why simulators are this cruel and why they allow evils to exist. If they want to claim we might be in a simulation, they have to give an explanation for the existence of evils and the cruelty of the simulators. But first, a principle of charity would have us restrain from believing simulators would be this cruel, without having good reason for believing so. The burden of proof is on adherents of the simulation hypothesis. However, not only do we not have any reason to believe simulators would be this cruel, I contend there are good reasons to believe they would not be cruel, but instead, quite moral beings.

3. The Expansion of the Moral Circle

I argue it is unlikely simulators would be malevolent beings because they would most likely have a wide and expansive moral circle, meaning they would treat conscious and sentient beings like us benevolently.

Much of my argument rests on this idea. To formulate it in other words, the benevolence of potential simulators towards simulated beings can be inferred from the expansion of the moral circle: as time goes by, a given group will consider more and more kinds of beings to be moral subjects.

I will first give a brief historical overview of this idea, from its birth to its modern version. I will then explain why we can assume our simulators would most likely have undergone an expansion of their moral circle and

would consider us to be beings worthy of moral consideration. This case will be approached through the tools of evolutionary ethics. I will finally provide a “proof of concept” and show that such an expansion has indeed taken place in our world. In a formalized way, my general argument for this section goes like this:

Premise 1: Simulations can only be made by civilizations that are collaborative and cooperative.

Premise 2: These kinds of civilization will, therefore, be benevolent and most likely have undergone an expansion of their moral circle, likely to the point of including simulated but conscious and sentient beings in their moral circle.

Premise 3: Civilizations of benevolent simulators will most likely not inflict widespread suffering on simulated beings.

Conclusion: Since there is widespread suffering in our world, something simulators would most probably not allow, chances are we do not live in a simulation.

The first premise seems a given. The second premise is the crucial one, and I will defend it at length in section 3.2.

3.1 Brief historical overview of the idea

The idea of there being an expansion of the moral circle has a long history. In the 2nd century, Stoic philosopher Hierocles already characterized our concern and duty for others as a series of concentric circles: at the center of these circles, standing as our most important object of concern and whom we have the strongest duty towards, is ourselves. As the circles gradually become wider and move away from the center, objects of lesser concern and lesser moral obligations are included. In closer circles, there are kins. Then comes extended family, our local community, other cities, our country, and so on, all the way to the entire human race.

Closer to us and more relevant to our current study, the 19th-century Irish philosopher W. E. H. Lecky wrote in his “History of European morals” that,

The question of morals must always be a question of proportion or of degree. At one time the benevolent affections embrace merely the family, soon the circle expanding includes first a class, then a nation, then a coalition of nations, then all humanity, and finally, its influence is felt in the dealings of man with the animal world. In each of these

stages a standard is formed, different from that of the preceding stage, but in each case the same tendency is recognized as virtue.²⁴

Lecky here mentions that this circle is not fixed but can actually expand to gradually encompass more and more beings. This idea was studied in greater depth by Peter Singer in his 1981 “The Expanding Circle”. In this work, Singer studies the impacts of sociobiology on ethics and contends that even though it is true we have natural altruistic tendencies towards our biological kin, this isn’t enough to justify a takeover of ethics by biology. Contrary to what E. O. Wilson thought, evolutionary biology alone does not exhaust the domain of ethics, which remains firmly in the hands of philosophers. In order to extend this circle of concern, Singer contends, we must rely on ethical thinking. Leaving aside the description of this phenomenon and instead claiming that this phenomenon should be harvested and put to use in order to ensure the welfare of more and more beings, he says that

The circle of altruism has broadened from the family and tribe to the nation and race, and we are beginning to recognize that our obligations extend to all human beings. The process should not stop there (...). The only justifiable stopping place for the expansion of altruism is the point at which all whose welfare can be affected by our actions are included within the circle of altruism. This means that all beings with the capacity to feel pleasure or pain should be included; we can improve their welfare by increasing their pleasures and diminishing their pains.²⁵

Forty years on, the expansion of the moral circle recently came back to the forefront of the philosophical scene thanks to the Effective Altruism (EA) movement, which considers that restricting our moral concern for persons alive today is a grave blind spot: future people should be accounted for and, given the fact there will be tremendously many of them, thinking about our impact on them should be a moral priority. Jacy Anthis and Eze Paez thus think the expansion of the moral circle should cover future people. They define the expansion of the moral circle as such: “a community’s moral circle has expanded with respect to some previous time t if, and only if, a number of entities which used to be given less than full

²⁴ William Lecky, *History of European Morals. From Augustus to Charlemagne* (New York, London: D. Appleton and Company, 1869), 107.

²⁵ Peter Singer, *The Expanding Circle* (Princeton University Press, 2011), 124.

moral consideration at t are now given more moral consideration.”²⁶ Of course, this expansion usually occurs in terms of sets and not individuals. Groups are included in the moral circle.

Now, for my argument to take off, I need to establish that our simulators (were they to exist) would have undergone such an expansion, meaning they would most likely be benevolent and moral beings.

3.2 The moral circle of our simulators: a case from evolutionary ethics

The case I wish to build here goes as follows: as cooperative species evolve, the moral consideration they bestow upon their cooperative partners expands with their circle of cooperation.

More specifically, I argue that as a species becomes a cooperative and social one, its members must develop altruistic behaviour to help facilitate this cooperation.

Any cooperative species must necessarily develop altruistic behaviour towards their cooperation partners to ensure that cooperation runs smoothly. Cooperation among a group of malevolent beings would be impossible. In order to prevent cooperation from breaking down, collaborative species necessarily had to develop, throughout their evolutionary history, altruistic behaviour. Such altruistic behaviours can be extremely varied and need not be the same for every species, but it intuitively seems that they must be present in one form or another. In the case of humans, Joshua Greene contends that “built in our moral brains are automated psychological programs that enable and facilitate cooperation”.²⁷ Such psychological programs include concern for others, direct reciprocity, commitment to threats and promises, reputation, assortment, and indirect reciprocity. Indeed, we seem to have a natural aversion to harming one another, and a natural tendency to help one another (concern for others). We cooperate when we know we can reap the benefits of cooperation (direct reciprocity), we punish uncooperative partners (commitment to threats and promises), we value reputation a great deal when choosing cooperative partners (reputation), readily cooperate with members of our group (assortment), and trust that uncooperative partners will be punished by others too (indirect reciprocity). As Greene explains, “All

²⁶ Jacy Anthis and Eze Paez, “Moral circle expansion: A promising strategy to impact the far future,” *Futures* 130 (2021): 2.

²⁷ Joshua Greene, *Emotion, Reason, and the Gap Between Us and Them* (London: Atlantic Books, 2015), 62–3.

of this psychological machinery is perfectly designed to promote cooperation among otherwise selfish individuals”.²⁸

Therefore, it seems that, as Green explains,

Each of us occupies the center of a set of concentric social circles. Immediately surrounding us are our closest relatives and friends, bounded by a larger circle of more distant relatives and acquaintances. Beyond our circles of kith and kin are strangers to whom we are related through our memberships in groups of various kinds and sizes: village, clan, tribe, ethnic group; neighborhood, city, state, region, country; church, denomination, religion.²⁹

Therefore, any cooperative species tends to act benevolently towards their cooperative partners, and as a species becomes more and more collaborative, this tendency is extended to ever more distant fellows, to the point of encompassing non-collaborative partners. This tendency is again, empirically confirmed in highly collaborative species, and Michael Tomasello gives us a possible principled reason for this, which he calls the “Interdependence Hypothesis”. As cooperative groups grow in size, they eventually split into smaller size groups “still unified at the tribal level”.³⁰ When this split occurs, the tendency to treat benevolently in-group members is extended to new out-group members.

However, one wishes to precisely formulate this idea, it seems that treating benevolently collaborative partners is a necessity, and also a tendency which we can reasonably expect to be extended to, at some point, non-collaborative partners. Therefore, since we can assume our simulators would necessarily be a collaborative and cooperative group, we can expect them to be benevolent, moral beings, who have a wide and expansive moral circle.

3.3 Proof of concept from our world

One could be tempted to reject the idea presented in the previous subsection on the grounds that in our world, we can see that this process has never taken place: our world is violent, ripped apart by wars, and altruism does not seem to be the hallmark of humanity, supposedly a collaborative species. This would be a misguided criticism, since the case we have built

²⁸ *Ibidem*, 62.

²⁹ *Ibidem*, 49.

³⁰ Michael Tomasello, *A Natural History of Human Morality* (Harvard University Press, 2016), 5.

in the previous subsection seems to be empirically confirmed if we retrace the expansion of the moral circle of the human species. It is a historical fact that our moral circle expands. In the last 300 years or so,

groups that were previously denied equivalent moral standing—including propertyless men, women, ethnic and religious minorities, homosexuals, and slaves—have become moral equals deserving of similar legal treatment. Slavery and the slave trade were abolished; democracy spread; women throughout the world have been granted the right to vote; and workers enjoyed an ever-expanding set of social and economic protections.³¹

The moral circle can be understood as a metaphorical boundary that separates who or what is considered morally relevant and who or what is not. Historically, many societies have started with a narrow moral circle that only includes members of their own tribe, community, or nation. As societies evolve and progress, the moral circle tends to expand to include a broader range of entities. The 19th century witnessed the emergence of animal welfare movements advocating for the humane treatment of animals. In the mid-20th century, amidst civil rights and feminist movements, the moral circle expanded further. African-Americans, women, and other marginalized groups began to be acknowledged as full-fledged citizens worthy of moral consideration. The latter half of the 20th century and the early 21st century brought a growing recognition of environmental ethics. As ecological awareness spread, ecosystems, plants, and non-human species were integrated into the moral circle.

Michael Huemer, building on the very extensive work done by Steven Pinker in his 2011 book “The Better Angels of Our Nature”, shows that the decrease in war and murder, the decrease of torture and capital execution, the eradication of slavery, the decrease in racism and sexism, and the processes of decolonization and democratization, all show that societies tend to give more and more people rights (and thus moral consideration). In most Western countries, even non-human beings are making moral gains: animals already have some rights, and the rights of non-sentient beings, such as robots or potential artificial intelligences are already being discussed with seriousness. Chalmers goes all the way to defend the moral worth of simulated beings:

³¹ Manvir Singh and Moshe Hoffman, “Individualistic moral principles and the expansion of the moral circle,” *Ethics and Psychology* (2022): 1. <https://doi.org/10.31235/osf.io/pqem7>.

If consciousness is all that matters for moral status [he thinks it is the case], simulated humans have the same moral status as ordinary humans,³² (...) Is it ok to create a simulation involving a million conscious beings all experiencing agony for their whole existence? I'd say it clearly is not.³³

It is useful here to back my claim about the expansion of the moral circle with some data. Some measures have been devised to precisely measure the scale at which individuals or societies grant moral status to other beings, like the "Moral Expansiveness Scale", but I think it is here more warranted to use different proxies to measure the size of our moral circle through time.

Among the candidates for good proxies, we can think of rights, war and violence, political systems and political empowerment. I take these proxies to be self-explanatory. If we consider a certain kind of being to be a moral subject, we are more likely to grant them rights, less likely to treat them violently, and more likely to grant them political freedom and power. Therefore, if there indeed has been an expansion of the moral circle, we should expect to see, throughout history, an increase in rights and political freedom and a decline in war and violence. It turns out this is indeed the case. Here are but a few telling examples.

Violence, both through warfare and outside of it, has drastically declined during human history.³⁴

³² Chalmers, *Reality*, 345.

³³ *Ibidem*, 346.

³⁴ Steven Pinker, *The Better Angels of Our Nature. Why Violence has Declined* (New York: Penguin Books, 2012) and Bill Gates, "Better Angels of Our Nature in Graphs and Numbers," *GatesNotes*, Sept 2, 2012, <https://www.gatesnotes.com/about-bill-gates/better-angels-of-our-nature-in-graphs-and-numbers>, accessed December 05, 2022.

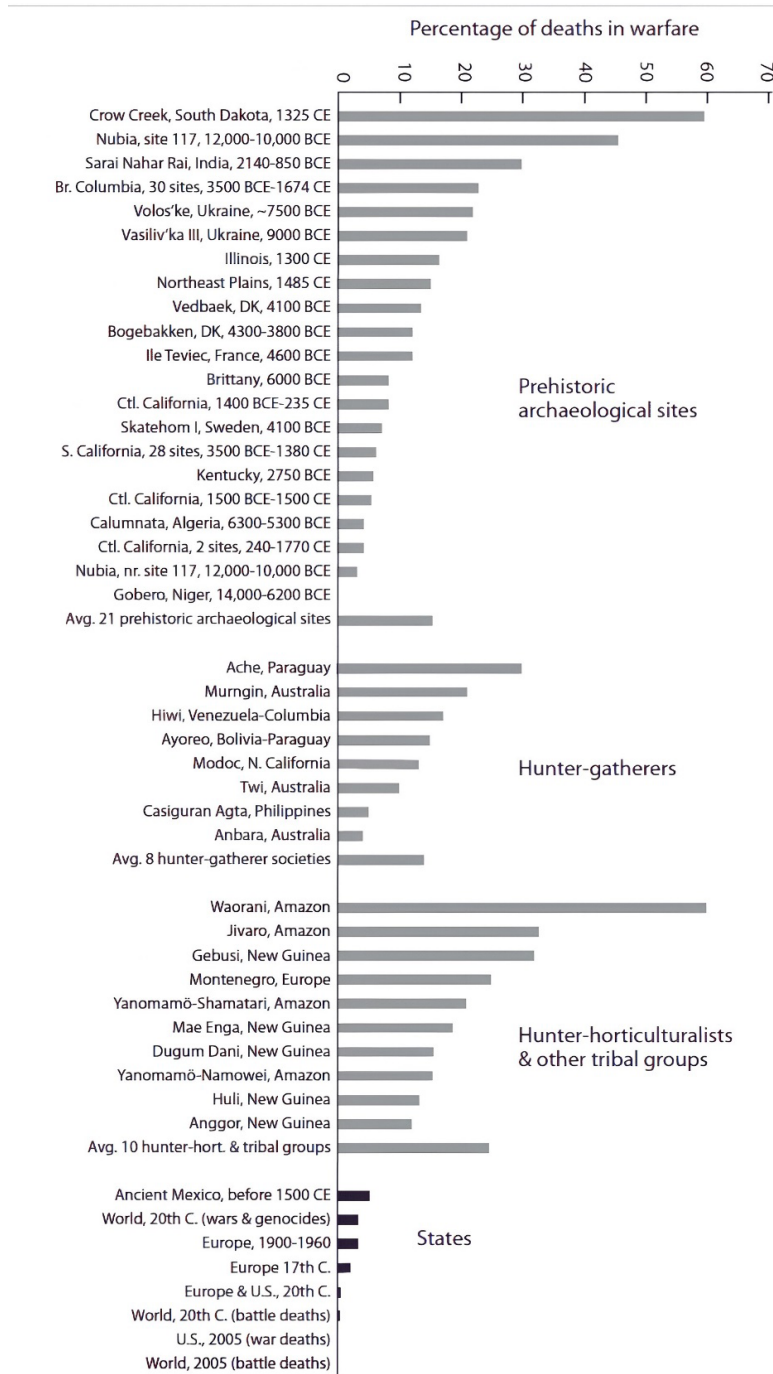


Figure 1. As we can see from this graphic, the percentage of deaths through warfare has significantly declined through history, indicating that violence, at least in the form of violent war, has also significantly declined.³⁵

³⁵ Michael Shermer, "The Enemy Within: A Review of «The Violence Paradox»" *Skeptic*, November 19, 2019, https://www.skeptic.com/reading_room/enemy-within-review-of-the-violence-paradox-nova-pbs/, accessed May 10, 2024.

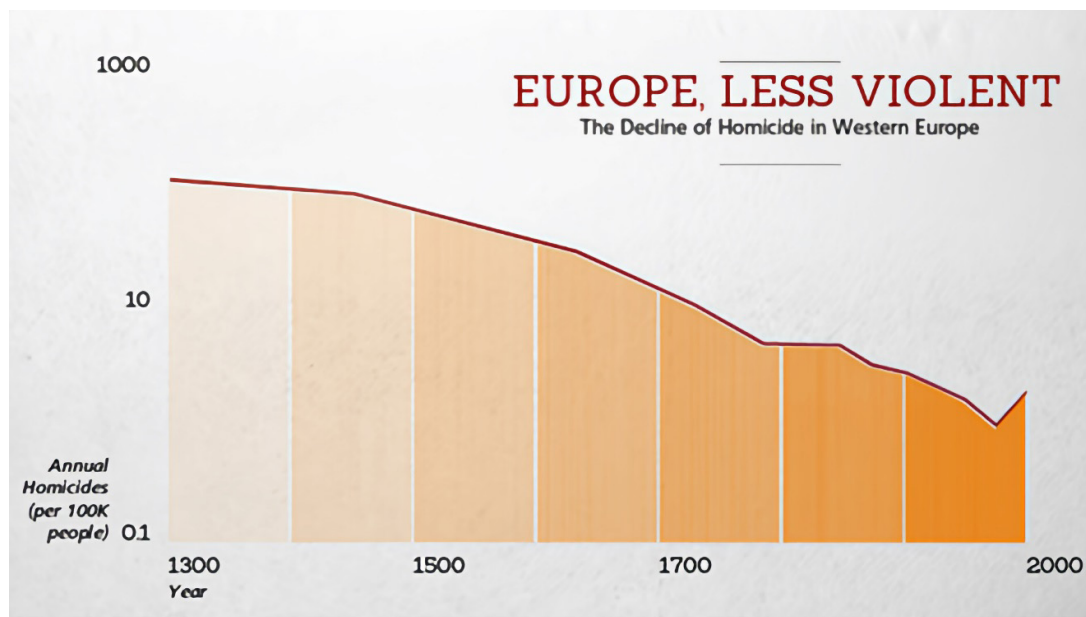


Figure 2. As we can see from this graphic, the number of homicide (per capita) has significantly declined, indicating, like the first graphic, that violence, at least in the form of homicide, has significantly decreased over time.³⁶

Data about the time of the abolition of torture³⁷ also shows a profound change in mentality. The abolition of torture also participates in the moral circle expansion: it is gradually more and more recognized that people have inalienable rights to bodily integrity, and notwithstanding the fact that torture is still sometimes used in warfare, its official abolition widely restricted its use.

³⁶ Bill Gates, "Better Angels of Our Nature in Graphs and Numbers," *GatesNotes*, Sept 2, 2012, <https://www.gatesnotes.com/about-bill-gates/better-angels-of-our-nature-in-graphs-and-numbers>, accessed December 05, 2022.

³⁷ Steven Pinker, "A History of Violence: Edge Master Class 2011," *Edge*, Sept 27, 2011, https://www.edge.org/conversation/steven_pinker-a-history-of-violence-edge-master-class-2011, accessed December 05, 2022.

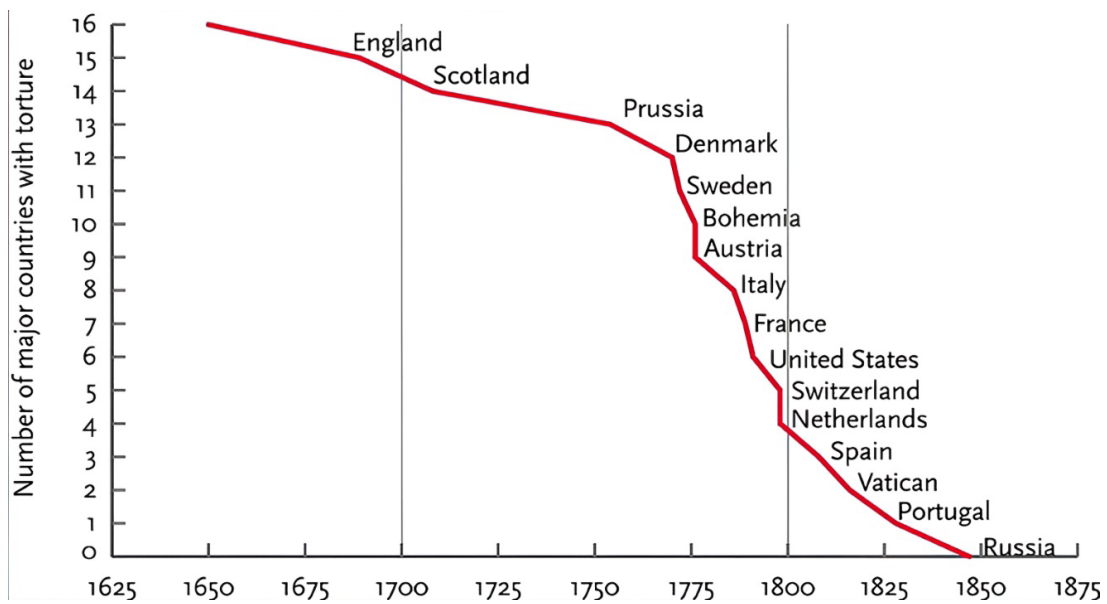


Figure 3. As we can see from this graphic, throughout history, more and more countries have abolished the use of torture, indicating the expanding recognition that torture violates the moral standing of beings who ought to be granted such moral standing.³⁸

The same goes for the abolition of slavery,³⁹ the recognition of women's rights,⁴⁰ and the gradual democratization⁴¹ of political systems around the world. All of these processes can be seen worldwide and have been constant throughout history.

³⁸ Steven Pinker, "A History of Violence: Edge Master Class 2011," *Edge*, Sept 27, 2011, https://www.edge.org/conversation/steven_pinker-a-history-of-violence-edge-master-class-2011, accessed December 05, 2022.

³⁹ Michael Huemer, "A Liberal Realist Answer to Debunking Skeptics: The Empirical Case for Realism," *Philosophical Studies* 173, no.7 (2016): 8.

⁴⁰ *Ibidem*, 8.

⁴¹ *Ibidem*, 9.

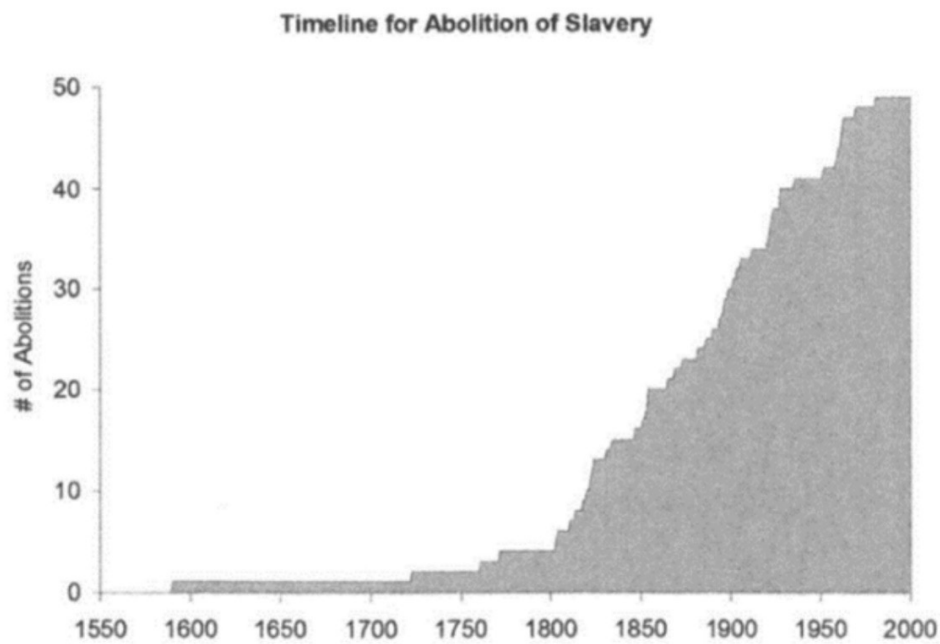


Figure 4. As we can see from this graphic, slavery has been, throughout history, abolished by more and more countries, indicating the expanding recognition that slavery violates the moral standing of beings who ought to be granted such moral standing.⁴²

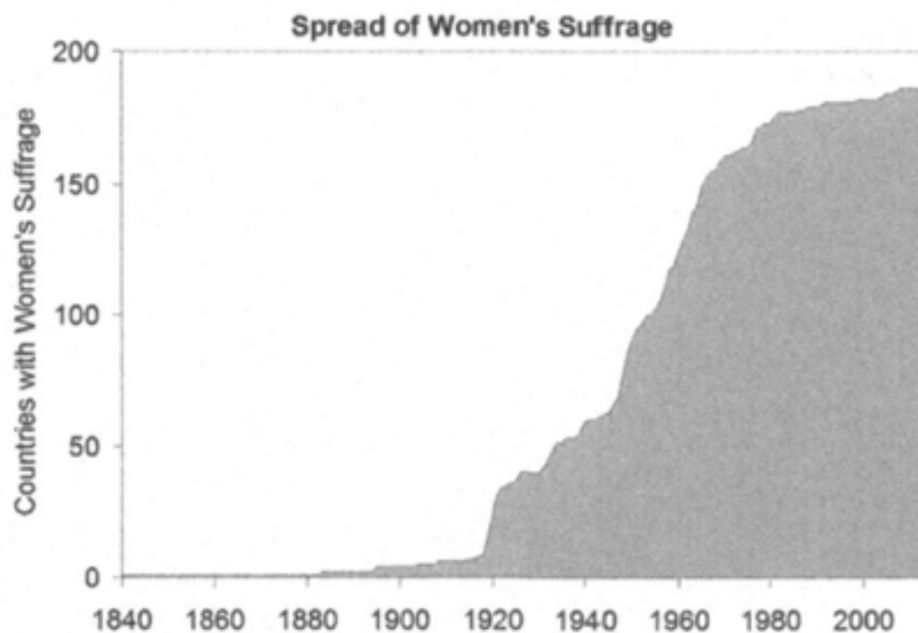


Figure 5. As we can see from this graphic, women's suffrage has been, throughout history, granted by more and more countries, indicating the expanding recognition that

⁴² Michael Huemer, "A Liberal Realist Answer to Debunking Skeptics: The Empirical Case for Realism," *Philosophical Studies* 173, no.7 (2016): 8.

women's suffrage participates to grant women political autonomy and full-fledged moral standing.⁴³

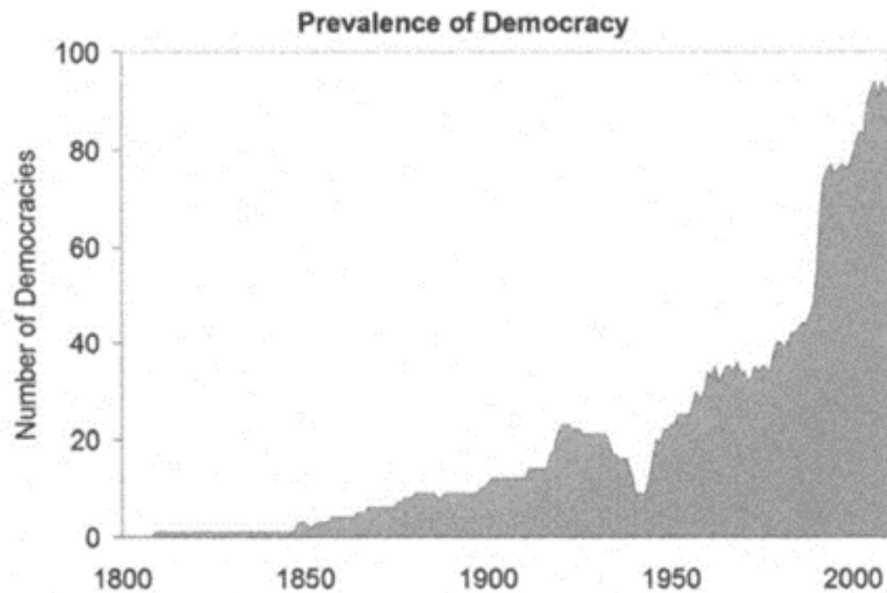


Figure 6. As we can see from this graphic, democracy has been, throughout history, adopted by more and more countries, indicating the expanding recognition that democracy participates to grant citizens political autonomy and full-fledged moral standing.⁴⁴

Finally, the expansion of the moral circle can be verified through a very telling proxy: the concern for minority rights.⁴⁵ It might be the best proxy because the legal gain minorities have obtained are proof that they, who prior to their moral recognition were not considered full-fledged moral subjects, have achieved this status and that there is an expansion of the number of beings who are considered moral subjects.

⁴³ Michael Huemer, "A Liberal Realist Answer to Debunking Skeptics: The Empirical Case for Realism," *Philosophical Studies* 173, no.7 (2016): 8.

⁴⁴ Michael Huemer, "A Liberal Realist Answer to Debunking Skeptics: The Empirical Case for Realism," *Philosophical Studies* 173, no.7 (2016): 8.

⁴⁵ Gates, "Better Angels," 8.

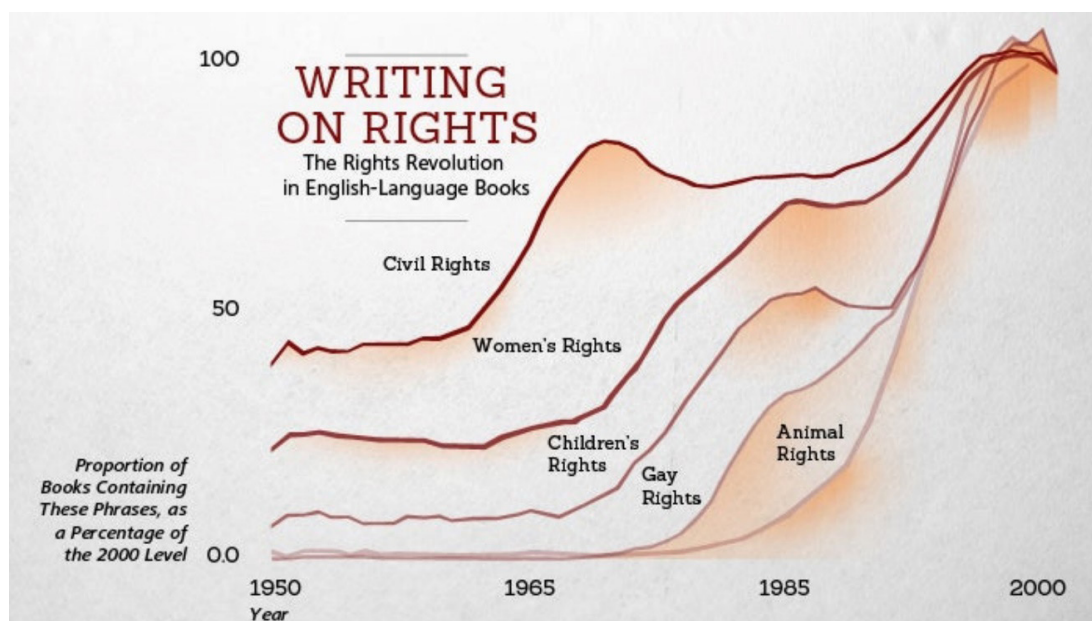


Figure 7. As we can see from this graphic, writing on rights has been, throughout history, more and more prominent, indicating the expanding concern given to rights and, correlatively, to moral standing.⁴⁶

Apart from those examples, we can also see that, throughout history, the human rights index has climbed up, there are more and more independent national human rights institutions; individual liberties and equality before the law have increased, private civil liberties have increased, the physical integrity index has increased, democracies are more common, women's suffrage has increased, women's political empowerment has increased, universal suffrage has increased, equal access to power for all groups has increased and the share of adult citizens who have the right to vote has increased.⁴⁷

What I take myself to have established here is that the expansion of the moral circle is indeed possible, and this lends credence to our claim of section 3.2 that any collaborative species must be, in some sense, benevolent. Of course, setbacks can happen. My claim is not that such setbacks are impossible and that any cooperative species are moral saints. My only claim is that, in broad terms, more and more groups of living beings, whether humans or not, are, through time, granted moral

⁴⁶ Bill Gates, "Better Angels of Our Nature in Graphs and Numbers," GatesNotes, Sept 2, 2012, <https://www.gatesnotes.com/about-bill-gates/better-angels-of-our-nature-in-graphs-and-numbers>, accessed December 05, 2022.

⁴⁷ These can all be confirmed thanks to the website "Our World in Data," <https://ourworldindata.org/>.

consideration by cooperative groups. If this is at least plausible, then it is at least plausible that civilizations more technologically advanced than ourselves, which have had a lot of time to expand their moral circle, will also grant moral status to a wide variety of beings.

3.3 The benevolence of our simulators

As there indeed is, as time goes by, an expansion of the moral circle, and as it is safe to assume any cooperative species must have benevolent tendencies, it is also safe to assume that simulators (who, in order to get to the technological stage required to run simulations have had a long history of cooperation) are likely moral and benevolent beings. This means their moral circle could be just as large as ours, or even larger.⁴⁸ As this process seems widespread among all civilizations, it is safe to assume that every intelligent civilization that evolves over a long period of time would also develop a wide moral circle.

In every case, we clearly see that as time goes by, earthly civilizations move towards pacifism and the protection of the rights of more and more living beings. We can thus conclude that if this trend is exportable to any possible collaborative civilization (and we have good reason to think it is, thanks to our case from evolutionary ethics), a civilization that would reach the stage of running large-scale simulations would not inflict needless suffering on the inhabitants of their simulations. Therefore, we can say that whereas we have no reason to believe simulators would be cruel and would inflict unnecessary evil on simulated beings, we have comparatively good reason to believe they would not. From this, it follows that we should give more credence to their benevolence than to their cruelty. If they are not cruel but benevolent, they would not allow for useless evil, and as there are useless evils in our world, it seems to hint at the fact that our world is not, in fact, a simulation.

To summarize, my argument goes as follows. I assume a civilization capable of running universe-scale simulations must be collaborative and cooperative. If they are cooperative, then we have good reasons to believe they are benevolent, moral beings. If this is true, the fact that we live in a

⁴⁸ This might seem like a strange idea to us, but if we went back in time a couple hundred years ago and told people that women, racial minorities, and sexual minorities had the same rights as today's heterosexual white male, they would be quite surprised. If they learned animals and potential future artificial intelligence might soon gain rights, they would be quite simply baffled.

simulation is highly unlikely, given the painful and pointless natural evils of our world. If we formulate the same argument in terms more closely resembling the Problem of Evil, we can say that simulators are local gods who are near omnipotent and near omniscient, and who are likely to be mostly good (or at least, it is more likely they are good than it is likely they are cruel). Given the fact that natural evils exist in our world, it is more likely our world was not created by these beings. Therefore, the possibility that our world is a simulation is significantly reduced.

Is it possible we are living in a simulation? There will always be a possibility, even extremely small, that we do. Is it possible our simulators are downright cruel and enjoying our suffering? There will always be a possibility. Is it possible our simulators have just put the simulation in fast-forward, gone on a coffee break or were bored from monitoring it, and haven't checked the simulation yet to stop it if it went awry? There will always be a possibility. My only point here is that the Problem of Evil greatly diminishes the chances of the simulation hypothesis being true, and the only question that has to be asked is what is more probable: that we live in a simulation run by cruel people, or that we simply do not live in a simulation? There are more reasons to believe we are not in a simulation than there are to believe we live in a simulation run by cruel people, and therefore, it is more warranted to hold the non-simulation scenario for the true one. My argument is thus only probabilistic and aims at diminishing the likelihood of the simulation hypothesis being true.

4. Counterarguments

Even if my claim is only probabilistic, it is still subject to counterarguments.

First, the simulators might not have the power over their simulation I grant them. For example, they might just have set up the initial parameters of the simulation, and then let it run free. This would make sense, because what could be the use of a pre-determined simulation of which they know the outcome? From both an entertainment, research and decision-making point of view, an open-ended simulation would make more sense than a deterministic simulation. Therefore, they wouldn't have the power to intervene in the simulation. But I find this implausible on the grounds that even if simulations are open-ended, they can be equipped with automatic stops in the case of something going wrong. If they are indeed moral beings as I suggest they are, we can bet pretty confidently on the fact that there are, embedded in their simulation, stopping criteria:

if the amount of pain and suffering becomes unbearable, then for ethical purposes, the simulation could be stopped. To this, of course, one could reply that we have no reason to think we have actually passed this stopping criterion.⁴⁹

Second, someone can grant that they are indeed a benevolent species, but we cannot exclude the fact that our simulation might have been created by a single malevolent person acting on their own. Our world might be a simulation run by a single deviant person who does not care for our suffering. But this is unlikely given the resources problem: it is unlikely a single person, acting on their own, has the capacity and resources to simulate our whole universe. I take it that simulating the entirety of our universe is likely to be a large-scale project necessitating the efforts and resources of many actors, and that a single person would probably not have the resources to, alone, simulate an entire universe.

Third, someone could contend that, maybe, they *have* to make us suffer. Say that, in their world, our simulation yields great utility, that the knowledge acquired from the simulation produces a tremendous amount of well-being in their world (for example, they might have learned how to prevent a pandemic). Then, from a utilitarian perspective (which they might hold to be the true moral view), the utility of both worlds (our simulated world and their non-simulated world) combined would outweigh the suffering of our world. In other words, we might be myopic about the amount of well-being in the whole universe. We see only the well-being and ill-being of our world, but maybe the ill-being in our world could be justified by the great amount of well-being it produces in their world. Therefore, they could consider themselves allowed to make us suffer.

I do not have a definitive answer to this counter-argument, but I have two possible replies. First, if the amount of well-being in their world truly compensates for the amount of ill-being in ours, it seems other routes could have been taken to spare us from this suffering. They could have designed the simulation in such a way that the beings that inhabit it, us, wouldn't be sentient, and therefore wouldn't suffer. Second, and this is my simplest, yet, I believe, most effective response: what's more probable? The fact that we live in a simulation where the ill-being of our world is justified by a moral species of simulators who live in a world where their

⁴⁹ I thank an anonymous reviewer for bringing this to my attention.

well-being rests on the simulation of our universe to a point where they are willing to accept to make us suffer in order to be better off, or... that we just don't live in a simulation? As I said, my argument is probabilistic. There are two options in front of us, and we have to choose. One option is that we live in a simulation run by people who make us suffer (for a reason we don't really know). The other option is that we just don't live in a simulation. The latter option seems much more probable.

This answer also holds for all the counterarguments presented here. What I have done throughout this paper is to show that if someone wishes to hold on to the simulation hypothesis, they have to introduce qualifications to their hypothesis, which reduces its likelihood. What I mean by this is the following: yes, our world might be a simulation, but given the fact that we can assume the simulators would be, generally speaking, omniscient, omnipotent, and good, if we want to defend the idea that we live in a simulation, we have to say our simulation was actually created by a single malevolent person, or that our simulation was actually created with the purpose of serving the greater good in their universe. These are all plausible scenarios, but the details included in the hypothesis lower its probability of being true. Therefore, what I have done is to show that in order to hold on to the simulation hypothesis (which one could do), they would need to come up with unlikely scenarios to explain the existence of evil in our world. For example, one could say that the simulation is in "fast-forward": a thousand years in our world is a single second in our simulators' world, and they just haven't had the time yet to shut it off.⁵⁰ Or, the simulation is some form of downloadable game in which anyone can create a world as they please. Again, this is possible, but it is a detail which further reduces the likelihood of the simulation hypothesis being true. In a sense, I am ready to grant most of the counterarguments have bite, but all of them, even if successful, also reduce the likelihood of the initial hypothesis being true.

5. Conclusion

As we see from my last response, the whole discussion around the simulation hypothesis has to be expressed in probabilistic terms, for it is a hypothesis which, notwithstanding what some scientists not too well versed in philosophy think, is untestable. Therefore, if we have to talk in

⁵⁰ I thank an anonymous reviewer for this comment.

probabilistic terms, we can pit scenarios against one another and simply ask: which is more likely?

The core of my argument ran along this exact line: I highlighted the fact that there is pointless suffering in our world, and from this, I inferred things about the simulators who would run such a simulation. What I ask is this: is this scenario, the one where there is an intelligent civilization that runs a simulation filled with pain and suffering and remains unfazed by it, more likely than the more commonsensical one where we just don't live in a simulation?

I have not, nor cannot say, exactly how much my claim reduces the likelihood of the simulation hypothesis being true, just like no proponent of the simulation hypothesis can precisely say how likely it is we live in a simulation. My argument reduces its likelihood, and that is all I can say about it, and all that I wish to establish. Just how significant this reduction of likelihood is, depends on the credence we lend to the initial simulation hypothesis, how much credence we lend to the simulators' benevolence, how much evils are in our world, and many other factors. Since I take their benevolence to be a safe bet, our world to be significantly filled with suffering, and the additional qualifications one needs to add to the initial hypothesis in order to maintain it to be unlikely, I take my argument to represent a significant reduction in likelihood, although this reduction cannot be quantified.

I have shown that we are justified in believing that every intelligent civilization that reaches a highly technological stage and, therefore, has had a long history of cooperation and collaboration is likely to be a moral civilization that would not engage in such pointless torture of beings of their creation. The Problem of Evil, thus addressed to the simulation hypothesis, seems to have the same effect as its standard formulation addressed to God: because the hypothetical creator of our world could not possess the characteristics, we attribute to them (in the case of God, omnipotence, omniscience, and all-goodness, and in the case of the simulators, near omnipotence, near omniscience, and highly moral attitudes), we have to conclude they do not exist. Therefore, we most probably do not live in a simulation.

I would like to thank Keven Bisson for his help in developing an intuition we both had regarding the simulation hypothesis into the idea presented in the paper. I would also like to thank him and Chris Howard for their numerous comments on previous versions of this paper, an



Christophe Facal

anonymous reviewer of RES for their comments, along with the participants to the 2nd annual Duke-UNC graduate philosophy conference for their questions and comments.