

## **Let's Hope We're Not Living in a Simulation**

Eric Schwitzgebel  
Department of Philosophy  
University of California, Riverside  
Riverside, CA 92521-0201  
USA

March 8, 2023

## Let's Hope We're Not Living in a Simulation

Abstract: According to the simulation hypothesis, we might be artificial intelligences living in a virtual reality. Advocates of this hypothesis, such as Chalmers, Bostrom, and Steinhart, tend to argue that the skeptical consequences aren't as severe as they might appear. In *Reality+*, Chalmers acknowledges that although he can't be certain that the simulation we inhabit, if we inhabit a simulation, is larger than city-sized and has a long past, simplicity considerations speak against those possibilities. I argue, in contrast, that cost considerations might easily outweigh considerations of simplicity, favoring simulations that are catastrophically small or brief – small or brief enough that a substantial proportion of our everyday beliefs would be false or lack reference in virtue of the nonexistence of things or events whose existence we ordinarily take for granted. More generally, we can't justifiably have high confidence that if we live in a simulation it's a large and stable one. Furthermore, if we live in a simulation, we are likely at the mercy of ethically abhorrent gods, which makes our deaths and suffering morally worse than they would be if there were no such gods. There are reasons both epistemic and axiological to hope that we aren't living in a simulation.

Keywords: David Chalmers, simulation hypothesis, skepticism, theodicy, virtual reality

Word Count: about 7000 words

## Let's Hope We're Not Living in a Simulation

### 1. *The Simulation Hypothesis.*

There's a new creation story going around. In the beginning, someone booted a computer. All the objects we see are configurations of that computer. We ourselves are also computer processes. We are artificial intelligences in a virtual reality – a simulation. It's a fun idea, worth taking seriously, but we should very much hope that we are not living in a simulation.

Nick Bostrom (2003) presents the standard argument for taking the simulation hypothesis seriously. It's not unreasonable to think that consciousness might eventually be possible in computer systems. If so, engineers in advanced technological civilizations might someday create artificially intelligent conscious entities – “sims” – living entirely in simulated environments. These engineers might create vastly many sims, for entertainment or science. The number of sims in the cosmos might greatly exceed the number of biologically embodied people living in non-simulated reality. If so, then we ourselves might well be among the sims. This idea has captured the imagination of technophiles around the globe. Elon Musk, for instance, estimates the chance that we're living in the “base level” of reality as “one in billions” (Musk 2016).

In *Reality+* (Chalmers 2022), David Chalmers also argues that we might be sims. Chalmers and Bostrom (unlike Musk) admit substantial doubt. Maybe no society will master the technology, because artificial consciousness is difficult or impossible or because technological civilizations quickly self-destruct. Or maybe sims exist but only rarely, forbidden by law or too expensive for mass manufacture. Or maybe sims can generally discern their simulated status, so that even if there are many sims, we can rest assured that we aren't among them. The simulation hypothesis is far from guaranteed. However, Chalmers and Bostrom suggest that it *might* be

true. It's not wholly implausible. Bostrom estimates about a one-in-three chance that we are sims. Chalmers estimates "at least 25 percent or so" (p. 101).

I'm inclined to think these credences are too high. For most of the entities in the cosmos to be sims without realizing it, several things would need to align, concerning the nature of consciousness, the nature of the cosmos, the trajectory and motivation of technological civilizations, and the indiscernibility of simulated worlds. Still, I've argued elsewhere that it's reasonable to have at least about a 0.1% credence that we are sims (Schwitzgebel 2017).

Nothing seems to justify nosebleed heights of near-certainty that we aren't sims – credences of 99.99% or more. Consciousness *might* be possible in artificially intelligent systems living in computer-based virtual realities; the cosmos *might* contain many sims of this sort; and if so, we *might* not be able to rule out the possibility that we are among them. The various cosmological and philosophical arguments pro and con admit of substantial reasonable doubt. If you feel certain that no A.I. system could ever be conscious, or that no technological civilization would ever create conscious sims, or that if we were sims we would surely know it – well, I admire your philosophical confidence! But, historically, claims of certainty about substantive, non-obvious philosophical theses have a poor track record. The "dismal induction" favors doubt.

Suppose, then, that we accept that there's a non-trivial chance that we're living in a simulated reality. How should we react?

Chalmers seems largely unconcerned. He writes, "being in an artificial universe seems not worse than being in a universe created by a god" (p. 328). A central theme of *Reality+* is that in a large, stable simulation, most of the ordinary objects we care about really exist and have most of the types of properties we think they have. He also argues that if we live in a simulation, it's unlikely to be a small one – just one city, say (p. 442-444). While Chalmers acknowledges

that we can't *rule out* various skeptical hypotheses with certainty (e.g., p. 461), his general tenor is upbeat and mostly anti-skeptical. For example, the concluding section of *Reality+* is titled "No escape from reality" (p. 458).

In general, defenders of the simulation hypothesis display little anxiety about catastrophic or radically skeptical consequences. Bostrom acknowledges that humanity faces an "existential risk" that the simulation will shut down – but he says that the "lion's share" of existential risk derives instead from our own technological activities (2002, p. 20). Eric Steinhart (2014) embeds the simulation hypothesis in an optimistic religious cosmology. With a detached smile, Musk expresses the hope that we do live in a simulation, since that would mean that technological progress has not been cut short by calamity.

I am not so sanguine. I will argue that we can't justifiably have high confidence (say 90% or more) that if we live in a simulation it's a large and stable one rather than one that's unstable or catastrophically small. Furthermore, if we live in a simulation, we are likely at the mercy of ethically abhorrent gods. There are reasons both epistemic and axiological to hope that we aren't living in a simulation.

Before continuing, let me clarify one thing I will not be arguing. I will not argue that if we live in a simulation, some horrible general metaphysical or epistemic consequence *necessarily* follows, such as that mountains aren't real or that everything is an illusion (for critiques more along those lines, see Avnur forthcoming; Fallis 2023). I agree with Chalmers that if we live in a large, stable simulation, and if we have done so for a long time and will continue to do so for a long time, then everything around us is as real and non-illusory as we could reasonably want. In a large, stable simulated reality we'd have real conversations, real achievements, and real suffering. We'd fall in and out of love. We'd hear beautiful music and

climb beautiful mountains (“mountains” enough, anyway) and have nations, careers, groovy dance parties, and wonderfully fluffy pillows. Who cares, really, about fundamental metaphysics? What does it matter whether, underneath it all, are quarks, leptons, and bosons or instead bits in a computer program? What matters is that reality reliably supports my conscious experiences, and yours, and all of our friends, and all of our ancestors, and all of our descendants, and the patterns of physical or seemingly-physical interaction with the world that we’ve come to expect.

My worry is much less abstract. My worry is that if we live in a simulation, we have excellent reason to doubt that there will be a tomorrow. We exist, perhaps briefly, at the whim of powerful entities whose kindness we have little reason to expect.

## *2. The Size Question.*

Here’s an example of a large virtual reality: The entire observable universe exists – all 93 billion light-years of it, with every star and planet simulated in microscopic detail throughout hundreds of billions of galaxies – and it has existed since the Big Bang, and it will exist for at least another trillion years.

Here’s an example of a small virtual reality: You are in a simulation that started five minutes ago and which will last another five minutes before deletion. Nothing is being simulated except for, briefly, you and your immediately perceivable environment.

Between these extremes lies a wide range of possibilities, some of which are small enough to be epistemically catastrophic – small enough that a substantial proportion of our everyday beliefs would be false or lack reference in virtue of the nonexistence of things or events whose existence we ordinarily take for granted. An epistemically catastrophic virtual reality

might be geographically small, for example if only one city exists; or it might be temporally small, for example if it was created ten years ago; or it might have a small population, for example if most of the seeming-people are really just mock-up sprites without real conscious experiences (Helton 2021). A brief or city-sized or low-population reality wouldn't be epistemically catastrophic for inhabitants who know such facts about their reality, but that's not our position. We normally assume that we had childhoods and that we live on a planet with billions of people. If such assumptions are false, we are wrong about many things of importance.

Consider, then, the following question: How confident ought we to be that if we inhabit a virtual reality the reality is large enough to be epistemically non-catastrophic – that the world contains more or less all of the things we care about, plus a reasonably deep past, plus a reasonably long future, and billions of people? Call this the *Size Question*. We can divide answers to the Size Question into optimistic and pessimistic. The optimist holds that we ought to be confident that if we are sims, we don't live in a catastrophically small simulation. The pessimist denies this. I endorse pessimism. Whatever credence we attach to the simulation hypothesis, we ought to attach a substantial subportion of that credence (10%? 50%? 90%) to catastrophic smallness. To the extent we take the simulation hypothesis seriously, giving it a real, non-trivial weight in our thinking, we should give similar (within one order of magnitude) weight to the possibility that much of what we care about doesn't, or didn't, or won't exist.

How should we assess whether optimists or pessimists are closer to right about the Size Question? I see three families of approach.

One approach professes radical ignorance. We have essentially no idea how to evaluate the Size Question. We have no idea what would motivate the creation of an Earth-like simulation; or how many such simulations exist; or how expensive they are to create; or what

hazards they are subject to; or their typical size duration, or features; or how base-level reality gives rise to our virtual reality. Our ordinary concepts might apply so poorly that even using the term “simulation” imports too much of our limited cognitive apparatus into something fundamentally unknowable. The matter lies so far beyond our ordinary understanding and our ordinary sources of evidence that we can’t even begin to assess them. (This might be a “Kantian humility” more radical than Chalmers’ own version of it on p. 419-422; also Kant 1781/1787/1998; Langton 1998; Schwitzgebel 2019a, forthcoming.)

Radical epistemic humility is not an unreasonable stance. However, it appears to justify pessimism rather than optimism. Either we assign some intermediate credence to our living in a small reality conditional upon our being sims, or we refuse to settle on any specific credence. Either way, radical ignorance of this sort is incompatible with the optimist’s confidence.

A second approach appeals to Moorean certainty or Wittgensteinian “hinge epistemology” (Moore 1925, 1959; Wittgenstein 1951/1969; Coliva 2010, 2015). On this approach, we are justified in treating some select propositions as fixed points in our reasoning. Among these propositions might be “here is a hand” and “billions of people exist”. If we are justified in treating enough such propositions as fixed points, then we are justified in believing that our reality is not catastrophically small.

This approach, strictly implemented, can conveniently solve any skeptical problem. But it ought not be strictly implemented. There are conditions under which counterevidence should lead one to doubt that “here is a hand” – for example, a proprioceptive illusion plus a trick with mirrors. Similarly, there are conditions under which counterevidence should generate doubt that billions of people exist – for example, if you discover that you have lived your life in a series of cleverly arranged stage sets (a la *The Truman Show*) as part of an entertainment for what appears



to be a much-smaller-than-you-thought Earthly population. Although (as Descartes suggested) maybe “I exist” is impossible to rationally doubt, beliefs concerning the size of the planet and the number of people who exist are at least in principle amenable to counterevidence and rational change. Recent advocates of hinge epistemology, such as Coliva (2015), acknowledge that hinges can be revised under the right conditions. The hypothetical discovery that you’re a sim is exactly the sort of situation which ought to generate doubt about such matters. To insist that you know, or can be certain, that if you are living in a simulation it is a large one, without further supporting grounds, simply begs the question. It is not a justifiable defense of optimism.

A third approach, which Chalmers appears to favor (as well as Bostrom and Steinhart), and which I also favor, attempts to assess the evidence that we occupy a large or a small simulation. Chalmers considers small simulations of two types: local and temporary.

Might the simulation contain only the city you’re now inhabiting (p. 442-444)? Stipulate that this city has existed for at least a hundred years, but nothing beyond it exists. Your local friends exist, and you have real conversations with them. The room you are currently in exists, and the building, and the roads – but everything stops at the city edge. Anyone looking beyond sees, presumably, some false screen. If they travel beyond the edge, they disappear from existence; and when they return, they pop back into existence with false memories of having been elsewhere. News from afar is all fake. Government edicts that affect city life – for example, increases in tax rates – are all generated ad hoc by the computer program.

Chalmers suggests that “the most obvious objection” to the local simulation is that it lacks simplicity. It lacks simplicity, presumably, because the computer will somehow need to figure out how to coordinate everyone’s false memories, and how to generate appearances of fields, buildings, and roads beyond the city’s boundaries, and how to create all the fake news,

governmental edicts, and so on. In contrast, Chalmers says, “global simulations just require simulating a few simple laws of nature and letting the simulation unfold” (p. 444).

I doubt we should be so quickly satisfied with a simplicity response. Simplicity is a complex business (Sober 1975; Zellner et al. 2001; Cowling 2013; Schaffer 2015). Furthermore, appeals to simplicity are often indecisive. The world might not be so simple; or it might be simple in some respects but not others. It’s unclear whether a planet-sized simulation is simpler than a city-sized one. Yes, a city-sized simulation will face complexities of the sort Chalmers describes, but a planet-sized simulation will have many more objects, and many more people, and many complex events occurring in regions far beyond the city’s boundaries. Perhaps the planet-sized simulation can employ simpler laws, but simplicity of law is only one dimension of simplicity.

Chalmers acknowledges that large simulations are likely to be much more costly than local simulations, which he suggests might be a reason for simulators to create only Earth instead of a whole galaxy of stars, if they are mainly interested in our lives (p. 444). Everything beyond the Solar System could easily enough be mock-up patterns of light, misleadingly designed to resemble a universe of stars, planets, nebulas, black holes, dark matter, cosmic background radiation shaped by quantum fluctuations shortly after a Big Bang, and so on. But notice now that we’ve lost the simplicity of simply starting with basic laws: The computer will somehow need to figure out how to generate the right patterns of light. Chalmers’ thinking is, presumably, that the cost tradeoff could make sense. If the simulators are really only interested in us, they might not pay the cost of creating a whole 93 billion light-year wide observable universe if a mock-up will suffice. Grant – as I would not – that simplicity concerns mainly the basic laws rather than the number of objects and the complexity of their interactions. Chalmers

acknowledges that simplicity and cost can trade off. The creators might choose a simulation with a boundary at the edge of the Solar System, even if that isn't the simplest choice.

Analogous reasoning applies to the city. Maybe one city is all our creators want or need. It might be easy enough to create fake boundaries, fake news, and fake memories, all nicely coordinated. Not so long ago, many artificial intelligence researchers thought that novel and fluent conversational prose would be difficult or impossible to generate with early 21<sup>st</sup> century computer technology, but recent large language models have proven surprisingly adept. Similar high-power modeling rules might create seemingly well-structured boundaries and fake memories. And it needn't be perfect. If the city's inhabitants start to notice inconsistencies, maybe the inconsistencies can be repaired post-hoc and inhabitants' memories rewritten. If our creators want a solo city, they might well have the resources to fool the inhabitants well enough. It's an empirical question, which we're really in no position to confidently answer, whether it's easier and more efficient to create a whole planet or a whole observable universe for the sake of a city or whether it's easier and more efficient just to create the one city and somehow deal with the problem of faking the world beyond.

Thus, we cannot justifiably feel confident that if we live in a simulation, it is planet-sized rather than only city-sized. Similar considerations undercut a similar argument by Bostrom (2011) that the most effective way to model intelligent conversation with distant people is to have actual distant people as one's conversational partners.

Chalmers considers a few types of temporary simulation, but let's focus on the possibility that our reality was created on January 1, 2023, with memories, historical documents, old buildings, fossils, etc., already in place. This simulation would be catastrophically small: All of our historical beliefs would be false, we would never have been children, and most of the people

we regard as deceased would never have existed. We would be radically wrong about a huge number of things we ordinarily care about.

Chalmers' argument against the temporary simulation scenario also relies on an appeal to simplicity. He suggests that "the obvious way" to create the right kind of fossil records and so forth would be to run a detailed simulation of the past – in which case we're not in a temporary simulation after all (p. 446).

Again, however, considerations of cost and simplicity might compete. If the simulators are only interested in us as we exist now, they might not want to pay the cost of running a full simulated Earth for millions or billions of years, and they might be able to avoid that cost by generating a plausible enough distribution of historical records, memories, etc. Again, it's an empirical question, whose answer we cannot confidently assess, how the cost of a long-enduring sim balances against the difficulty of seeding a well-organized start date.

As Chalmers notes, some temporary simulations might be easy to create – for example, one person awakening from a nap in a dark room (p. 446). If the person doesn't survive long they won't have much chance to check for flaws and shortcuts. Another simple model might be a few dozen people together in a room listening to a philosophy talk. Do we need a whole history and fossil record in place? Do we need Bangladesh to exist? No one will really be checking; and even if a few of you fire up your phones for sports news or whatever, plausible inputs could probably be faked for the several minutes we are here existing together.

Chalmers acknowledges that he "can't rule... out with certainty" the possibility that he is living in a brief simulation, though he notes that as long as local reality exists – the chair he's sitting in, for example – he avoids a "global skepticism" in which we don't know anything at all (p. 446). Maybe so! But this isn't optimism yet. There's more to optimism than avoiding truly

global skepticism in which we know literally nothing about the world outside. If our whole reality is limited to just us here in this room for a few minutes together, that's epistemic catastrophe enough.

Chalmers concludes by acknowledging that he can't be certain that a particular friend of his exists or that Australia exists; but he also says that the "conspiracies" required to fool him on such matters would have to be complex, and it might be possible to rule them out on those grounds (p. 461). The regularities of our experience are a good guide, he says, to the approximate structure of the world, which limits the strength of skeptical arguments.

I find myself somewhat unsure how to map Chalmers' view onto pessimism versus optimism regarding the Size Question. In general, Chalmers acknowledges sources of skeptical doubt, admitting uncertainty. On the other hand, he argues that considerations of simplicity tend to speak against various skeptical scenarios and that there's reason to expect conformity between the structures of our experience and the structures of the world. Chalmers' remarks are compatible with optimism on the Size Question if his thought is that considerations of simplicity justify *high confidence* that we're not in a catastrophically small simulation, even though we can't be completely *certain*.

In any case, I want to be clear and direct on this point. If we adopt an evidential approach to the Size Question, the evidence is indecisive. We can mine Chalmers' remarks for some considerations in favor of thinking that the simulation would not be catastrophically small, but those considerations – mainly the appeal to simplicity – are indecisive and opposed by the general observation that, to the extent we can guess about such things at all, it seems reasonable to suppose that large simulations will be higher cost.

Let me briefly mention two other arguments for optimism on the Size Question. Sometimes I've heard the suggestion that whoever is running the simulation will be morally advanced in virtue of being technologically advanced. Maybe technologically advanced societies usually destroy themselves if they lack strict ethical codes. Then, the argument continues, we might expect the simulators to have ethical principles that would forbid creating entities like us in catastrophically small sims.

It's an interesting argument, but it hardly justifies high confidence. Technological and moral progress might not generally go hand in hand. Societal preservation might be better facilitated by norms concerning the treatment of ingroup members than by ethical treatment of outgroups. An advanced technological society might flourish best by keeping its own members happy and well-coordinated, in part by permitting them distracting entertainments like the creation of temporary play worlds. Also, in the next section, I'll give some reasons to think that if we're in a simulation, our simulators are distinctly unbenevolent.

Chalmers (see also Bostrom 2011) has observed that even if large simulations are less common than small ones, large simulations will contain many more people, with the possible consequence that any individual person is likely to be in a large simulation (p. 139-140). Suppose, for example, that the relevant technological society creates a million simulations of a single person alone in a room and only one simulation containing a planet of a billion people. On some plausible self-location principles, it is a thousand times more likely that you are in the large sim than in one of the million small ones. After all, of the 1,001,000,000 simulated people who exist, most are in the large simulation.

I'm inclined to accept the self-location principle involved. But it doesn't yield the optimistic conclusion unless we are warranted in thinking that the ratios are favorable. It's

unclear what would justify confidence that the ratios are favorable. Large simulations might require considerably more resources than small simulations. A simulation of a billion people might require a billion times the resources. Of course, if there are economies of scale, it might require considerably less. On the other hand, if there are complexities of combinatorial interaction, it might require considerably more. Whatever motives our creators might have for simulating us will need to be weighed against these costs. Simulator scientists might prefer many small low-cost simulations whose outcomes can be systematically compared.

Entertainment seekers might prefer sims that are only large enough to be entertaining at relatively low cost. Consider the simpler simulations that we early 21<sup>st</sup> century humans run. Rarely do we simulate worlds with billions of individually modeled inhabitants. We run lots of small simulated scenarios as entertaining games and scientific projects, as well as climate models that model people only as aggregates. Now whoever is running this simulation, if we are in a simulation, might have motives far beyond our comprehension; but that adds only more uncertainty to our reasoning. We cannot justify high confidence in a favorable ratio of large to small simulations, even if we grant that a ratio far below 1:1 could count as favorable.

To recap: We cannot justifiably feel confident that if we are sims the reality we occupy is large enough to contain most of the objects, events, and people that we care about. If we are sims, we could quite easily occupy a small or brief reality. To treat the existence of billions of other people as a fixed framework assumption is unjustifiable dogmatism once one allows the possibility that we are living in a simulation, since allowing that possibility is exactly the sort of thing that should call into doubt our assumptions about the size of the world. To treat the question as radically unanswerable is perhaps reasonable, but that should undercut any confidence in speculations about what size our simulation would be if we are in one. To appeal

to plausibility considerations concerning the size of simulations, such as what would be simplest, or most cost effective, or socially allowable, or best in accord with likely simulators' motives, throws us again into doubt, since the considerations don't decisively align on one side.

Thus, the most epistemically reasonable position is doubt. Whatever credence you assign to the simulation hypothesis, you should assign a substantial subportion of that credence to the possibility that you, or we, live in a small simulation in which much of what you take for granted about the world is false.

Will Earth exist in 2025? If the ground level reality is what we ordinarily take it to be – fundamental physical particles of the sort detected by ordinary physics – high credence appears to be warranted. Our planet is under no imminent threat. If, in contrast, we are sims living in a virtual reality, a variety of risks impair that forecast. Maybe Earth has never existed; it's just us here in this room. Or maybe Earth exists, but only as a brief experiment. Maybe the simulators *want* Earth to exist in 2025, but the computer crashes. Maybe they lose funding. Maybe the program goes buggy. Maybe the simulators' juvenile offspring erase the program to make room for something else, or maybe they spawn a black hole nearby for the fun of watching us react to the disaster. Earth becomes subject not just to harassment by unlikely rogue asteroids but all kinds of risks that aren't ordinarily within the scope of our thinking and whose probability is difficult to assess.

I haven't discussed instability, but instability is a closely related issue. A large but unstable simulation would also be epistemically catastrophic – a simulation in which the laws of nature, for example, will undergo a radical change tomorrow.

### *3. The Pathetic, Cruel, or Indifferent Gods.*



Stipulate that we know, somehow, that we live in a large, stable virtual reality. Planet Earth exists, has existed for a long time, and will exist long into the future, with billions of people leading lives of approximately the sort they think they're leading, even if the fundamental metaphysics is surprising. Would this be bad? Would this be worse than living in the base level of reality? Chalmers suggests that in some respects it might be a *little* worse (ch. 17). For example, the world would lack the full, rich history that we normally assume it has, and maybe that history matters to us. However, if the simulation is complete enough – if ordinary things really are more or less how we think they are – then our reality has most of what gives life value: real emotional states, real interactions with other people, real ethical decisions, real accomplishments, real engagement with interesting and challenging environments. So there's not much reason, Chalmers argues, to hope that we aren't living in a simulation, as long as we know that it's appropriately large and stable.

However, despite having a chapter on theology (Chapter 7), Chalmers does not, I think, sufficiently explore the unfortunate theological consequences of the simulation hypothesis. We have ethical or axiological grounds for hoping we aren't sims, in addition to the epistemic and prudential grounds already discussed.

If we live in a simulation, there's a creator – some entity with sophisticated mentality and advanced technological capacities. Otherwise our artificial reality isn't an *artifice* – something intentionally created. Although maybe some non-mental swamp plant belched forth our world without intending to do so, as a side-effect of some digestive process, that's not the simulation hypothesis.

Posit, then, a simulator who launched our reality. This simulator, or others of its kind, or others not quite of its kind, also designed our reality. Perhaps this simulator has the power to

delete our reality or to interfere with it, creating “miracles” that violate what we regard as the ordinary laws of nature. This simulator exists outside of our spatial manifold, uncontained by our spatial dimensions, presumably capable of existing even if our whole reality ceases. The simulator’s time might differ radically from our time. Maybe things can run much faster for us, or much slower for us; or maybe the simulator can pause us without our realizing it. Maybe the simulator can rewind us to a save point, tweak a few things and then restart, in a sense changing the past. Maybe the simulator can copy our whole world. Maybe they can change the laws of nature. All of this might be true even if at the base level of reality, the simulator is unimpressive – some foolish adolescent gamer living in their parents’ basement.

If Zeus is a god, then our simulator is a god. If the simulator has even a small portion of these powers, that puts Zeus to shame. Even just the power to choose to launch or not launch our world, combined with existence outside of our spatial manifold, is arguably sufficient for the simulator to be a proper referent of the term “god”. At first, Chalmers seems to agree (p. 128), though later he demurs on grounds that he would not want to worship the simulator (p. 144). I’m inclined to think worship-worthiness is not a necessary criterion of godhead, so I’ll continue to call the simulator a god, but I take this to be a terminological point. (See also Schwitzgebel 2019b, ch. 21-22.) “God” turns out, perhaps surprisingly, to be a relational term: an entity can be a god relative to one reality and a non-god relative to another reality. If we accept that our creators are gods, then the simulation hypothesis is a theistic hypothesis, despite its origins in naturalistic theories of technology and its dissociation from traditional religions.

We can now do some “natural theology”. That is, we can examine the world around us with the aim of making educated guesses about the properties of God or the gods. One striking fact regularly noticed by natural theologians is this: The world contains plenty of apparently

unnecessary suffering (e.g., tooth decay) and moral evil (e.g., the Holocaust). This appears to force a theological choice: Either God *cannot* prevent all the bad things that happen, or God *prefers not* to prevent all the bad things that happen (classic treatments include Bayle 1697/1965; Hume 1776/1998; Stump 2010).

Consider *cannot* first. God closed her/his/its/their seventeen bug-faceted eyes and pressed the launch button, hoping for the best. “Go, little world!” Then God let things alone, knowing intervention isn’t possible, maybe toodling off to other tasks, never looking back; or maybe God watches in impotent horror as Genghis Kahn, Hitler, and Stalin kill their millions, as children die of painful cancer, as earthquakes, plagues, and starvation ravage us, as billions of people suffer badly designed bodies prone to pointless sinus infections and back trouble. Maybe God tries to help but proves ineffectual. Call this the *pathetic God* possibility. God couldn’t even have given Hitler a heart attack? That single, minor, virtually undetectable action, which wouldn’t have struck any observer as distractingly miraculous, could plausibly have prevented an enormous moral atrocity. Our creator has so little control over such important features of their creation? That’s sad, pitiful, a terrible design mistake. If a world will have billions of genuinely conscious, suffering people, it needs a good user interface.

More likely – to the extent we can assess likelihood – if the gods have the power to create a simulated world, they also have the power to make adjustments. They just prefer not to. Maybe God is an angry adolescent who loves watching us kill each other in wars, with our cute little guns and military costumes. Maybe God is a scientist who regards us as lab animals whose suffering is irrelevant as long as the hypotheses of interest can be adequately tested. Maybe God is an artist whose audience is awed by the cruel display. “Don’t you dare touch Hitler!” God

says to the leering viewers. “The perfection of my artistic vision requires the Holocaust.” The audience departs in tears, and God racks up the attendance fees and grant money.

Of course, a rich theological tradition suggests responses. Following Leibniz (1710/1968), we might posit that this is the best of all possible worlds, or at least the best of all technologically feasible worlds. No simulator could have made things better. Give Hitler a heart attack and something even worse than the Holocaust would have happened. God dare not save that innocent three-year-old from the slow, painful death of Tay-Sachs disease, by curing the disease or at least giving a quicker, less painful death, otherwise... well, something. Following Stump (2010), maybe God presents every person only the suffering required to give them the best chance to flourish: Every child who starves to death is suffering only because starvation unto death is their best chance to achieve the desires of their heart.

It is perhaps not strictly impossible that Leibniz or Stump are correct. As Hume emphasizes, if one knew in advance for certain that an omnipotent, benevolent God exists, maybe you could justifiably accept a view of this sort. Nor is it wholly irrational at least to *hope* that every bad thing is ultimately for the best. But an even-handed natural theology, which examines the world empirically to discover the features of God, hardly suggests that this is the best of all possible worlds or that every human agony serves a great soul-building purpose.

Maybe somehow, for some reason, the gods faced a choice between either creating our world with all its evil and suffering or not creating our world at all? Overall, it’s better (the argument goes, and I agree) that our world exists than that it does not. There’s plenty of good too. They might, then, be benevolent after all.

But this is wishful thinking against the evidence. There’s a reason theologians construct rococo rationalizations of our needless suffering, falling into desperately implausible solutions

like Leibniz's and Stump's. To the extent we can evaluate likelihood at all, the power to design and create suggests the power to intervene and improve. Even casual examination suggests fixes that should be easy for civilizations of such immense power. The absence of such fixes suggests that our creators are cruel or at best grossly irresponsible.

If so, our world is ethically and axiologically worse than a world in which evil and suffering arise from the indifferent forces of nature. It is ethically and axiologically worse because there is an evil or reprehensibly indifferent entity overseeing it all. It is worse in the same way that it's worse for a child to die due to murder or neglect than for that same child to die by unforeseeable accident. Because our creators are responsible for our existence and for our relatively happy or miserable state, they owe us benevolence. To all appearances, they are abhorrently unbenevolent. What kind of IRB approves of *this* experiment?

This connects to the Size Question. If we are at the mercy of a moral monster, all the more reason to worry that the whole thing might stop tomorrow or be changed radically to introduce entertaining catastrophes. Alternatively, if God is pathetically impotent, things might go haywire for reasons beyond God's control – similarly if God is deceased or somehow excusably ignorant. The utility bill might go unpaid. The computer might be repossessed.

Or flip the moral reasoning around: If we're convinced that God must be merciful, that's grounds to suspect that all the suffering is fake. Really, the Holocaust never happened! No child has ever died of neglect or cancer. Maybe it's just us here in this room right now, having not too unpleasant an experience (I hope). But that, of course, tosses us back into a pessimistic answer to the Size Question (Helton 2021, p. 237).

Let's hope we're not living in a simulation. If we are sims, our existence and the existence of many of the people and things we care about depend on contingencies difficult to

assess and beyond our control, and all the badness of the world appears to reflect either the gods' intentional cruelty or their callous disregard. A large, stable rock is a more dependable and less axiologically troubling fundamental ground for reality.

## References

- Avnur, Yuval (forthcoming). Reality+: Virtual worlds and the problems of philosophy. *Philosophy*.
- Bayle, Pierre (1697/1965). Paulicians. In *Historical and critical dictionary*, trans. R. H. Popkin. Bobbs-Merrill.
- Bostrom, Nick (2002). Existential risks: Analyzing human extinction scenarios and related hazards. Reprinted from *Journal of Evolution and Technology*, 9. <https://nickbostrom.com/existential/risks.pdf> [accessed Mar. 7, 2023].
- Bostrom, Nick (2003). Are we living in a computer simulation? *Philosophical Quarterly*, 53, 243-255.
- Bostrom, Nick (2011). Bostrom's response to my discussion of the simulation argument. Blog post at The Splintered Mind (Sep. 2).
- Coliva, Annalisa (2010). *Moore and Wittgenstein*. Palgrave.
- \*Coliva, Annalisa (2015). *Extended rationality*. Palgrave.
- Cowling, Sam (2013). Ideological parsimony. *Synthese*, 190, 3889-3908.
- Fallis, Don (2023). *If we're living in a simulation, we're probably massively deluded*. Manuscript [accessed Mar. 8, 2023].
- Helton, Grace (2021). Epistemological solipsism as a route to external world skepticism. *Philosophical Perspectives*, 35, 229-250.
- Hume, David (1776/1998). *Dialogues concerning natural religion*, 2nd ed., ed. R. H. Popkin.
- Kant, Immanuel (1781/1787/1998). *Critique of pure reason*, ed. and trans. P. Guyer and A.W. Wood. Cambridge University Press.
- Langton, Rae (1998). *Kantian humility*. Oxford: Oxford.

- \*Leibniz, Gottfried Wilhelm (1710/1968). *Theodicy*, trans E. M. Huggard. Open Court.
- Moore, G. E. (1925). A defence of common sense. In *Contemporary British philosophy*, ed. J.H Muirhead. London: George Allen & Unwin.
- Moore, G. E. (1939). Proof of an external world. *Proceedings of the British Academy*, 25, 273-300.
- Musk, Elon (2016). Is life a video game? *Code Conference 2016*.  
[https://www.youtube.com/watch?v=2KK\\_kzrJPS8&list=PLKof9YSAshgyPqIK-UUYrHfIQaOzFPSL4](https://www.youtube.com/watch?v=2KK_kzrJPS8&list=PLKof9YSAshgyPqIK-UUYrHfIQaOzFPSL4)
- Schaffer, Jonathan (2015). What not to multiply without necessity. *Australasian Journal of Philosophy*, 93, 644-664.
- Schwitzgebel, Eric (2017). 1% Skepticism. *Noûs*, 51, 271-290.
- Schwitzgebel, Eric (2019a). Kant meets cyberpunk. *Disputatio*, 55, 411-435.
- Schwitzgebel, Eric (2019b). *A theory of jerks and other philosophical misadventures*. MIT Press.
- Schwitzgebel, Eric (forthcoming). *The weirdness of the world*. Princeton University Press.
- Sober, Elliott (1975). *Simplicity*. Oxford University Press.
- Steinhart, Eric Charles (2014). *Your digital afterlives*. Palgrave Macmillan.
- Stump, Eleonore (2010). *Wandering in darkness*. Oxford University Press.
- Wittgenstein, Ludwig (1951/1969). On certainty, ed. G. E. M. Anscombe and G. H. von Wright. Harper.
- Zellner, Arnold, Hugo A. Zeuzenkamp, and Michael McAleer, eds. (2001). *Simplicity, inference, and modeling*. Cambridge University Press.