# TESTING THE SIMULATION HYPOTHESIS: THE ANNIHILATION OF THE UNIVERSE AND CAMPBELL'S CONSCIOUSNESS-BASED ALTERNATIVE

Eliott Edge
(November 14th, 2024)

> "I had a question… a question you're not supposed to ask,
> which gave me an answer… you're not supposed to know."
>
> –*Westworld* [1]

**The** visionary horror writer H.P. Lovecraft was obsessed with the perfection of a literary trope throughout his artistic career: that certain forbidden knowledge of either a scientific or occult nature should turn against his protagonists, or even mankind as a whole, and lead them inexorably to annihilation or madness. In a lonely, first-floor apartment on 169 Clinton Street—"a run-down neighbourhood, with a dubious clientele, and infested with mice"[2]—situated at the border of Red Hook, Brooklyn, and not long after separating from his wife, Lovecraft outlined what would become his most renowned work, "The Call of Cthulhu." Its preamble so succinctly captured that aura of the damned and the forbidden that Lovecraft so endlessly pursued:

> We live on a placid island of ignorance in the midst of black seas of infinity, and it was not meant that we should voyage far. The sciences, each straining in its own direction, have hitherto harmed us little; but some day the piecing together of dissociated knowledge will open up such terrifying vistas of reality, and of our frightful position therein, that we shall either go mad from the revelation or flee from the light into the peace and safety of a new dark age.[3]

Within the literary cosmology of Howard Phillips Lovecraft, our universe holds many unsettling secrets and some of these secrets humanity should simply never come to know.

[1] *Westworld*, "The Original," directed by Jonathan Nolan, season 1, episode 1, (Warner Bros., HBO, October 2, 2016.)
[2] S. T. Joshi, *A Dreamer and a Visionary: H.P. Lovecraft in His Time* (Oxford: Oxford University Press, 2001), 216.
[3] H. P. Lovecraft, "The Call of Cthulhu," *Weird Tales* 11, no. 2 (1928), Popular Fiction Publishing Company.

The end of the human race is not as terribly difficult to imagine as we should altogether hope. Baked into nearly every world religion is some form or another of *eschaton*—the End Times. Today, popular media like *Marvel* and *DC* films play with depictions of the end of the world so frequently that it hardly feels like a cinematic threat to the moviegoer. Practically speaking, our species could easily wipe itself out with a man-made plague. Most life on this planet could come to an end thanks to nuclear weapons. Pump enough greenhouse gasses into the atmosphere and earth could transform into something closer to Venus—a lifeless crag enshrouded by a global cloud layer of sulfuric acid.

The end of our species or our planet is a strikingly casual thought. We all know this to be true. But the end of the universe? That, we normally imagine, is certainly not within the reach of mere mortals. No—a feat of that cosmic magnitude could only be accomplished by something closer to the status of Deity. Or so we imagine…

While humanity's ingenuity when it comes to unlocking nature's secrets and converting those revelations into a new kind of weapon—or a new thoughtless maximizer of waste and profit—stands without question, quite *rarely* is the position maintained that scientific discoveries in and of themselves simply *should not occur*. Today, we do see this attitude when it comes to the development of unregulated AI. Of course, to err on the side of caution is a reasonable position when it comes to uncontrolled digital superintelligence.[4,5] Similarly, genetic engineering has the potential to unleash novel forms of viral and bacterial life, the likes of which we may be wholly unprepared to deal with.[6] Applied to human DNA, genetic engineering could permanently alter the human germline in damaging, unforeseeable ways.[7] Outside of a few extreme circumstances, a mere handful of examples, human beings generally do not maintain the position that something shouldn't be explored scientifically.

While critics and ethicists rightfully concern themselves with the ways scientific knowhow might be used to cause harm, either purposefully or through negligence—some philosophers argue that the *very pursuit* of simulation theory in a meaningful way could have nothing short of extremely dire—indeed, wholly cosmic—consequences. Just as far-reaching as the consequences are for the obsessive heroes in Lovecraft's tales, so too does simulation theory sometimes carry with it the warning that to undertake this line of inquiry seriously could undo more than just our species but, quite possibly, the universe itself. To argue such a point is a tall order (perhaps the tallest order imaginable), but the work we will look at is published in academic journals, and therefore they are certainly worth entertaining—if not merely for entertainment's sake.

In this essay, we will look at some of the fears that simulation theorists argue we should be wary of and attempt to discern if there is any validity to them. We will see that the overarching fear—again, *of the end of the universe*—only holds sway depending on *which version of simulation theory is being considered* as well as the myriad assumptions that come with each iteration.

---

[4] Roman V. Yampolskiy, "Taxonomy of Pathways to Dangerous Artificial Intelligence," in *Proceedings of the 13th AAAI Conference on Artificial Intelligence*, (2016) https://cdn.aaai.org/ocs/ws/ws0156/12566-57418-1-PB.pdf

[5] Daron Acemoglu, "Harms of AI," Working Paper no. 29247 (Cambridge, MA: National Bureau of Economic Research, 2021), https://doi.org/10.3386/w29247.

[6] National Academies of Sciences, Engineering, and Medicine et al., *Biodefense in the Age of Synthetic Biology* (Washington, DC: National Academies Press, 2018), https://doi.org/10.17226/24890.

[7] K. E. Ormond, D. P. Mortlock, D. T. Scholes et al., "Human Germline Genome Editing," *American Journal of Human Genetics* 101 (2017): 167–76, https://doi.org/10.1016/j.ajhg.2017.06.012.

# THE SIMULATION ARGUMENT

According to the thinkers we will explore herein, living in a simulated world comes with an assumed specter: that our universe could suddenly be switched off. It takes very little force from our fingertips to turn off our laptops, smartphones, and gaming consoles. Surely, if our universe is a simulation, it must be just as easy to turn off as our own computers. But why would this happen, and what can *we* do, if anything, as would-be sims to avoid such a catastrophe?

The possibility of our universe being wiped out was briefly mentioned by the creator of the Simulation Argument Dr. Nick Bostrom in his seminal 2003 paper, "Are you Living in a Computer Simulation?"[8] Bostrom, the former Director of The Future of Humanity Institute, Professor of Philosophy at Oxford, and featured on many lists of the world's most prominent thinkers, takes simulated worlds and global cataclysms quite seriously. He echoed this fear again in his later paper "Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards" where he explores the major threats to the survival of our species. Here, Bostrom flatly asserts, "[W]e suffer the risk that the simulation may be shut down at any time. A decision to terminate our simulation may be prompted by our actions or by exogenous factors."[9] Amazingly, Bostrom places the extinction of our species via 'simulation shutdown' at the unexpectedly high rank of third most probable—just below nuclear holocaust.

Many articles, essays, and interviews featuring Bostrom and other thinkers have been published airing these and similar cataclysmic fears. The majority of 'simulation shutdown' anxieties and arguments revolve around an important assumption: Bostrom's *ancestor simulation hypothesis*. An ancestor simulation is a simulation made by an advanced species either to study and research the history and choices of the past (hence ancestor), or for entertainment. This scenario, first outlined by Bostrom, is called the Simulation Argument and it is a form of simulation theory, or *simulism*, that most philosophers engage with—partially because it is the most well-known. According to Bostrom's Argument, one of the following three conditions must be true:

(1) Human beings will never become 'posthuman'; and in fact, all intelligent species die out *before* they become 'posthuman'-like; that is to say super technologically advanced.
(2) Neither ourselves nor other super advanced species will ever run ancestry simulations.
(3) We are almost certainly in a simulation.

The logic may seem unusual, but it is actually quite sound. The whole argument is based on Bayes' Rule: the likelihood of something being true is based on prior conditionals (or in this case, the prior assumptions) related to that outcome.[10]

To begin, "posthuman" is used as a catch-all term for a highly technologically advanced species; but, despite the name, this species is not necessarily 'human' *per se*. It doesn't really matter. Posthumans are, more importantly, just a super technologically advanced species.

---

[8] Nick Bostrom, "Are You Living in a Computer Simulation?" *Philosophical Quarterly* 53, no. 211 (2003): 243–55.
[9] Nick Bostrom, "Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards," *Journal of Evolution and Technology* 9 (March 2002) (first version: 2001), 7. https://nickbostrom.com/existential/risks.pdf
[10] James M. Joyce, "Bayes' Theorem," *The Stanford Encyclopedia of Philosophy*, Fall 2021 ed, edited by Edward N. Zalta, accessed February 29, 2024, https://plato.stanford.edu/archives/fall2021/entries/bayes-theorem/.

Now, the first conditional: '*an advanced species always dies out*' assumes that they are very similar to human beings here and now; they trifle with widely destructive weapons and practices; they are as shortsighted and warlike as we are. This is of course a massive assumption, as we have no other species even close to human beings at our current stage of development to carefully compare ourselves to. Therefore, we can safely assume that this is probably unlikely. Eventually, somewhere at sometime some civilization probably did (or probably will) get its act together and thrive. In principle, there should not be a reason for this to be impossible.[11,12]

Second conditional: '*advanced species never run ancestry simulations*' even though they are technologically capable of doing so. In other words, they can build a super advanced simulation, and they have an idea of how useful it could be (studying existential risks, or alternative lines of technological and ideological development, etc), but for some reason they choose not to. Bostrom assumes that this is probably unlikely as well; there is no obvious reason for us to assume that this never ever happens.

The final condition: '*we are almost certainly in a simulation*' falls out of this logic seamlessly, because *if the first two assumptions are false* (which Bostrom argues they probably are), and posthumans *do* in fact survive long enough to run ancestry simulations (*which they probably would*), then we are more than likely in a simulated universe ourselves.

Think about it: *if* posthumans exist somewhere *and* they run simulations, then there are probably more than one—unless posthumans are exceedingly rare. That being the case then, there are probably *many* of them—that is, many posthumans running many simulations. So if all that is true, or, at least, *likely,* then we are probably in a simulated universe ourselves.

Bostrom makes clear, "The simulation argument doesn't try to show that we are living in a simulation. Instead it tries to show that one of three possibilities is true."[13] This is another reason for its popularity; it doesn't require scientific observation, rather a simple understanding of logic—making it easier to build arguments and ideas around. The Simulation Argument in the simplest sense can be stated thusly: *if* a super advanced species exists *somewhere* in the multitude of possible worlds, *and* this species happens to make advanced simulations of lesser advanced species, *then* we are more likely in a simulated universe than not.

Thinkers and philosophers have examined Bostrom's Simulation Argument and found it to be logically consistent with some exceptions (Franceschi 2008, 2014; Agatonović 2021; Bruiger 2023). Nevertheless, such criticisms have not stopped thinkers from continuing to use Bostrom's work; ultimately finding his premises compelling enough to warrant further consideration.

As Dr. Preston Greene, an American philosopher who teaches at Nanyang Technological University, Singapore, explains Bostrom's Argument using our species as an example, "If we believe that our civilization will one day run many simulations concerning its ancestry, then we should believe that we are probably in an ancestor simulation right now."[14] However, it is critical to bear in mind that Bostrom's Argument is *not* scientific, but rather philosophical. Specifically,

---

[11] Recent computer modeling suggests a limit of approximately 1,000 years for technologically advanced civilizations to avert climate disaster via waste heat production. See: Amedeo Balbi and Manasvi Lingam, "Waste Heat and Habitability: Constraints from Technological Energy Consumption," *arXiv preprint*, (September 2024), https://doi.org/10.48550/arXiv.2409.06737.

[12] Sierra Bouchér, "Alien Civilizations Are Probably Killing Themselves from Climate Change, Bleak Study Suggests," *LiveScience*, October 2, 2024, https://www.livescience.com/space/alien-civilizations-are-probably-killing-themselves-from-climate-change-bleak-study-suggests. "If a species has opted for equilibrium, has learned to live in harmony with its surroundings, that species and its descendants could survive maybe up to a billion years."

[13] *Closer to Truth,* "Could Our Universe Be a Fake?" Episode 110. Hosted by R. L. Kuhn. YouTube, March 2, 2022, [6:44], https://youtu.be/X6lbwcCI8TA?si=JUdw9OaUtPJFZr8G.

[14] Preston Greene, "The Termination Risks of Simulation Science," *Erkenntnis* 85, no. 2 (2020): 489–509. https://doi.org/10.1007/s10670-018-0037-1.

it is an *epistemological* argument: epistemology being the branch of philosophy concerned with how knowledge works, what are its limits, methods, and so on.

So, while keeping in mind that this is all completely hypothetical, and that some thinkers take issue with Bostrom's assumptions, let's nevertheless run with the conclusion of Bostrom's Argument, *that we are almost certainly in an ancestor simulation*, and see why it tends to make the universe such a terribly dangerous place for all life as we know it.

## TWO FORMS OF SIMULATION SHUTDOWN: COSTS AND PROBES

In a *New York Times* 2019 article titled, "Are We Living in a Computer Simulation? Let's Not Find Out," Greene publicized his anxieties over exploring simulation theory.[15] In it, Greene outlines the ancestor simulation hypothesis and considers the potential risks to humanity of pursuing it experimentally. After offering a few examples of possible tests that we could run, Greene writes, "So far, none of these experiments [to test simulation theory] has been conducted, and I hope they never will be. Indeed, I am writing to warn that conducting these experiments could be a catastrophically bad idea—one that could cause the annihilation of our universe." If Greene's prose here were just a touch more *purple*, they would be straight out of Lovecraft.

There are two widely discussed reasons why a simulation shutdown might occur: (1) it becomes too "cost ineffective" (2) we compromise ourselves as valuable sims by revealing the fact that we are in a simulation. In other words, we learn too much about the true nature of reality.

Let's take a look at each.

## COSTS

The ancestor simulation hypothesis assumes that our universe is a virtual reality running on a posthuman's computer. Well, the question naturally arises: what if we wanted to run *our own* simulated universe, even though we might be in one already? These are called *nested simulations*: simulated universes within simulated universes—and it is where all the trouble begins. In terms of *cost effectiveness*—not cost as in money *per se*, but cost as in resources and power. We can imagine that *many* resources are already required to run a universe as detailed as the one we live in. If we tried to run a simulation as detailed as our own (or close to it), it would require *even more* resources to perform *that* feat. The posthumans' computer would have to calculate our simulation, *and then* provide the calculations for another simulated universe on top of our own. This would be like cramming a whole universe inside another universe; or shoving a huge program like a video game inside of another already massive program, and then running them both on top of each other. If our universe is running on a computer somewhere, and that computer has *finite resources* (as all computers do), then the potential for a system crash skyrockets. Eventually, you have to close a program or shut down the computer entirely.

[15]Preston Greene, "Are We Living in a Computer Simulation? Let's Not Find Out," *The New York Times*, August 10, 2019. https://www.nytimes.com/2019/08/10/opinion/sunday/are-we-living-in-a-computer-simulation-lets-not- find-out.html.

As Greene articulates:

> [A]ny computation performed in a simulation must ultimately be supported by computation on the basement level. The computing constraints and the objectives of the simulators on the basement level, therefore, put a limit on the amount of nesting that can occur. It is partly for this reason that ancestor simulations entail a termination risk to those that create them.[16]

Greene argued his fears in depth in his 2018 paper, "Termination Risks of Simulation Science." There, he writes, "If the arguments of this paper are sound, then [developing our own] ancestor-simulation technology is even more dangerous than nanobots."[17] Nanobots are often brought up in technology-driven, worldwide catastrophe scenarios. Nanobots are advanced microscopic robots that could be used in highly delicate operations. Theoretically, we could create a type of nanobot that self-replicates. If left unchecked, these nanobots might hypothetically go out of control and self-replicate until they eat all the biomass on the planet (the so-called "gray goo" scenario).[18] According to Greene, simulation experiments are more dangerous than that conjecture.

The cost efficiency basis for a hypothetical simulation shutdown is rather straightforward: *our universe is shut down because of how much power it costs to run it*—and this power cost is directly connected to *whether or not we create ancestor simulations ourselves.*

What else could result in a dangerously high level of computational resource use? Another reason that our resource-cost could become too great is our creation of a super powerful AI. If we were to create a powerful enough AI, one that required an exponential amount of resources, then we would have the same overload problem all over again.

So, our universe is at the risk of being shut down, generally speaking, if we become too technologically advanced and our computers simply gobble up too many resources for '*base reality*' (where the posthumans and their computers are—also called '*basement reality*') to tolerate.

In essence, this whole argument for cost inefficiency has been a roundabout way of saying: *our universe could be shut down if we approach a posthuman state ourselves and create computers that are too powerful for base reality to tolerate*. Now, on its face, this is something we probably do not need to concern ourselves with right away, because it is so far beyond what we are, *or likely will be*, computationally capable of achieving any time soon. The hardware is just not nearly that advanced yet. So this is a version of universal annihilation that we can safely put aside for the time being.

The next version of simulation shutdown that we will explore in the following section is by far more intellectually tantalizing and imminently available to human beings. This is the version of universal annihilation more in-line with Lovecraft's thinking—that is, the uncovering of forbidden knowledge.

---

[16] Greene, "Termination Risks," 7.
[17] Ibid., 19.
[18] Lawrence Osborne, "The Gray-Goo Problem," *New York Times Magazine*, December 14, 2003. https://web.archive.org/web/20140906235944/http://www.nytimes.com/2003/12/14/magazine/14GRAY.html.

# PROBES

We have looked at the annihilation of the universe as a result of running our own ancestor simulations and other computationally-demanding activities. In those scenarios, we become too resource-heavy to the assumed posthumans' computers running our simulated world. As discussed, we need not be too concerned about such a situation, given that we are unlikely to build computers advanced enough to require such demanding processes any time soon. But what about running an experiment that would prove beyond a shadow of a doubt that our universe is a simulation? This, in theory, would be a *far easier* feat for humanity to pull off. It is also *another* scenario that could hypothetically result in our reality being utterly destroyed.

To begin to appreciate this section, it is crucial to understand what *ontology* means. Ontology is a field in philosophy, namely metaphysics, that deals with the nature of being, existence, and reality. A person's *ontological orientation* is how they think they exist in relation to reality. For example, after we realized that the earth revolved around the sun, we had a different ontological orientation—we believed, and therefore effectively *lived*, in a different worldview and a different universe. Once we realized that we lived on a spherical planet that orbits a star, it gave birth to not only a different way of seeing the universe (many stars with many planets), but it also opened up the possibility of space travel, satellites, and wireless communication. Similarly, if you think your disease is caused by evil spirits rather than bacteria or viruses, then you have a worldview that does not allow for the application of advances in modern medicine. We could say, *How I see myself in relation to the universe dictates my ontological orientation, and by extension, what matters to me and what I believe to be possible and knowable.* In simplest terms, your ontological orientation is your worldview.

In 2015, I published an essay titled "Why It Matters That You Realize You're in a Computer Simulation" for The Institute for Ethics and Emerging Technologies, a think tank co-founded by Bostrom. In it, I describe three types of lifeforms and their ontological orientations in a simulated universe: *Simple, Complex,* and *Savvy.*[19]

(1) **Simple** — they can make decisions and engage meaningfully with their environment.
(2) **Complex** — they record history as well as develop sciences, cultures, artifacts, and arts.
(3) **Savvy** — they are conscious of the fact that they are in a simulated universe.

Right now, human beings are *Complex* under this classification system. To become *Savvy* we would need some kind of irrefutable evidence that we are in a simulation, not just philosophical pondering. This evidence might arrive in the form of a scientific experiment. Such an experiment is called a *simulation probe*; "an experiment designed to determine whether the world is a simulation."[20] It wouldn't matter *what* kind of experiment it was, any would do, as long as it proved the matter to be true. Such an experiment with a positive outcome would result in human beings moving from a Complex ontological orientation to a Savvy one—we would *know,* not merely believe or suspect, but categorically *know* our universe is a simulation.

[19] Eliott Edge, "Why It Matters That You Realize You're in a Computer Simulation," *Institute for Ethics and Emerging Technologies* (IEET), November 14, 2015. https://web.archive.org/web/20171230132724/https://ieet.org/index.php/IEET2/more/Edge20151114.
[20] D. Braddon-Mitchell and A. J. Latham, "Ancestor Simulations and the Dangers of Simulation Probes," *Erkenntnis*, 2022, https://doi.org/10.1007/s10670-022-00560-8.

Now, let's take a look at a few of these potentially universe-destroying experiments.

Former NASA and Department of Defense physicist Tom Campbell offers five such potential simulation probes in his paper "On Testing The Simulation Theory."[21] These experiments involve testing the *von Neumann-Wigner interpretation* of quantum mechanics: *consciousness collapses the wavefunction into a particle*.[22] If this turns out to be the case, it gives strong credence to the idea that what we perceive as the universe *is only what we need to perceive at any given time*, much like how a player experiences the world in a first-person video game. In simulation theory, this is called the 'local simulation hypothesis'—the whole universe is *not* being processed *all of the time*. Rather, only what a given conscious observer (a player), along with what the simulation requires to maintain its integrity, is what is processed (i.e. made seemingly "real") at any given time.[23] Everything else is an unprocessed, unrendered, yet-to-be-generated *potential*. In essence, what we experience in the world is just an 'effect' that is waiting to 'pop up' whenever a player 'needs to see it.' This creates a sense of continuity and realness in an otherwise completely simulated reality.[24] In other words, just like the process of *occlusion culling* in a video game—where anything outside the in-game camera's view isn't rendered—in our simulated world, if you're not looking at it, it isn't there.

Melvin Vopson, an Associate Professor of Physics at the University of Portsmouth, has argued that "information is the fifth state of matter." All simulation theories put forth that the universe is finite, rather than infinite, and that *information* is the fundamental stuff of the universe. Vopson's work focuses on mass and its relation to energy and information. His "proposed mass-energy-information (M/E/I) equivalence principle—[suggests] mass can be expressed as energy or information, or vice versa," *and*, "that information bits must have a small mass."[25] Vopson's paper "Experimental protocol for testing the mass–energy–information equivalence principle," "[offers] specific predictions about the mass of information as well as the most probable information content per elementary particle." Vopson's novel experiment "involves a matter–antimatter annihilation process." He continues, "This experiment could, therefore, confirm [...] the existence of information as the fifth state of matter in the universe."[26] Simply put, if you can prove that the universe is made of information—in this case, by using a

---

[21] Tom Campbell, Houman Owhadi, James Sauvageau, and David Watkinson, "On Testing the Simulation Theory," *International Journal of Quantum Foundations* 3 (2017): 78-99. https://doi.org/10.48550/arXiv.1703.00058.

[22] Zvi Schreiber, "The Nine Lives of Schrödinger's Cat: On the Interpretation of Non-Relativistic Quantum Mechanics" (master's thesis, University of London, October 1994), 46. "The rules of quantum mechanics are correct but there is only one system which may be treated with quantum mechanics, namely the entire material world. There exist external observers which cannot be treated within quantum mechanics, namely human (and perhaps animal) minds, which perform measurements on the brain causing wave function collapse." https://doi.org/10.48550/arXiv.quant-ph/9501014.

[23] David J. Chalmers, *Reality+: Virtual Worlds and the Problems of Philosophy* (New York: W. W. Norton & Company, 2022), 31. "The local simulation hypothesis says that the simulation simulates only a part of the universe in detail. It might simulate just me, or just New York, or just Earth and everyone on it, or just the Milky Way galaxy." Chalmers discusses the issues in some detail: see the 'local simulations' index entry, 512. However, under Campbell, **only** whatever a player perceives is ever rendered and made "real."

[24] Bostrom, "Living in a Computer Simulation?," 5. "If the environment is included in the simulation, this will require additional computing power – how much depends on the scope and granularity of the simulation. Simulating the entire universe down to the quantum level is obviously infeasible, unless radically new physics is discovered. But in order to get a realistic simulation of human experience, much less is needed – only whatever is required to ensure that the simulated humans, interacting in normal human ways with their simulated environment, don't notice any irregularities. The microscopic structure of the inside of the Earth can be safely omitted. Distant astronomical objects can have highly compressed representations: verisimilitude need extend to the narrow band of properties that we can observe from our planet or solar system spacecraft. On the surface of Earth, macroscopic objects in inhabited areas may need to be continuously simulated, but microscopic phenomena could likely be filled in ad hoc. What you see through an electron microscope needs to look unsuspicious, but you usually have no way of confirming its coherence with unobserved parts of the microscopic world."

[25] Melvin M. Vopson, "How to Test If We're Living in a Computer Simulation," *The Conversation*, November 21, 2022. https://theconversation.com/how-to-test-if-were-living-in-a-computer-simulation-194929.

[26] Melvin M. Vopson, "Experimental Protocol for Testing the Mass–Energy–Information Equivalence Principle," *AIP Advances* 12, no. 3 (March 1, 2022): 035311. https://doi.org/10.1063/5.0087175.

matter-vs-antimatter annihilation technique—then you are moving in the direction of proving that we live in a fundamentally information-based world—much like a computer.

"Constraints on the Universe as a Numerical Simulation" by Dr. Silas R. Beane (Professor of Physics, University of Washington), Dr. Zohreh Davoudi (Associate Professor of Physics, University of Maryland, College Park), and Dr. Martin J. Savage (Professor of Physics, Institute for Nuclear Theory, University of Washington) argues that searching for anomalies ("glitches") in cosmic rays might reveal that we are in a simulated universe.[27] This paper assumes that if our reality is a simulation, perhaps space is laid out in a grid, with the boxes of that grid being infinitesimally small in size. How forces behave along these proposed grid lines may offer us clues. The idea is to exploit the high energy particles that are emitted by cosmic rays: particles that are moving at nearly the speed of light. "If these particles are unable to exceed a specific amount of energy this would prove the existence of a grid," explains Dr. Hartmut Wittig.[28] Prove there is a grid at the very lowest possible level of reality, and you end up supporting the idea of something along the lines of how pixels behave on a computer screen. Herein, space is made of discrete little boxes, not a continuous (unbroken), infinite field.

This proposal is particularly interesting because we already know that light and matter *seem continuous* (a light beam, or a stream of water for instance), however we now realize that they are actually made of discrete little bits (photons and $H_2O$ molecules). Yet, it is still a mystery whether or not *space* itself is discrete or continuous. The discovery of a grid, like a checkerboard, would add another piece of evidence arguing for a digital universe.

Whether we are looking at cosmic rays, or how matter and antimatter blow each other apart, or how conscious observers interact with physical phenomena, physicists are well at work proposing exciting new ways of testing simulation theory; and therefore, potentially, destroying the whole universe in the process.

Who knew it could be so easy?

Now, for some simulation theorists, if you perform an experiment that reveals our universe is a simulation, *it is not the experiment itself that destroys the universe*; it is our *new understanding,* our new ontological reorientation, that destroys the universe. Simply put, "The beings inside the simulation start to realize that they are in a simulation, so it is not a simulation anymore and cannot be used as such."[29] Greene comments: "If we were to prove that we live inside a simulation, this could cause our creators to terminate the simulation—to destroy our world."[30] Even talking about it or discussing it is *potentially* dangerous because it could very well lead to such an experiment happening.

Greene underscores the stakes of trying to find out whether or not we live in a simulation with superb urgency: by comparing it to playing a game of *Russian roulette*.

> Imagine that you are uncertain whether the gun you are holding is loaded with a bullet. One way to test is to play a single round of Russian roulette. If the gun does not discharge, then this creates weak evidence that the gun is unloaded. Call this the "boring" result. Call the result in which the

[27] Silas R. Beane, Zohreh Davoudi, and Martin J. Savage, "Constraints on the Universe as a Numerical Simulation," *European Physical Journal A* 50, no. 9 (October 2012). https://doi.org/10.1140/epja/i2014-14148-0.
[28] somethingwithscience, "Constraints on the Universe as a Numerical Simulation - Fast Forward Science 2015," YouTube video, July 31, 2015, [03:40], https://youtu.be/68hb7aGeCzg.
[29]Alexey Turchin, Michael Batin, David C. Denkenberger, and Roman V. Yampolskiy. "Simulation Typology and Termination Risks," *ArXiv* (2019), 16, "Pure philosophizing about being in a simulation would likely not be enough to cause termination, as such theorizing is likely inevitable. However, observing some undeniable glitches would destroy the purpose of the simulation as an illusion of reality, unless the simulation is rewound and corrected." https://doi.org/10.48550/arXiv.1905.05792.
[30] Greene, "Living in a Computer Simulation? Let's Not Find Out."

> gun discharges the "shocking" result. The shocking result proves that the gun was loaded, but, in some important sense, it does not matter, because you are dead. Since the shocking result has negative value, it is irrational to use Russian roulette as a means of investigating unless you assign an overriding positive value to the boring result.
>
> [...] Given the negligible value of the boring result, if experimental [simulation] probes have non-negligible positive expected value it must be because of the value of the shocking result. Therefore, we should focus our attention on determining the expected value of the shocking result.[31]

Greene uses this wonderfully morbid analogy to argue that if finding out we live in a simulation results in a shutdown, it simply isn't worth doing. He states, "This is my point: The results of the proposed experiments will be interesting only when they are dangerous. While there would be considerable value in learning that we live in a computer simulation, the cost involved—incurring the risk of terminating our universe—would be many times greater."[32]

What about the acquisition of knowledge? After all, the whole pursuit of science and philosophy is to accumulate knowledge about reality. But Greene is unconvinced—he does not believe that simple fact-finding is a worthy enough cause in this case; "the benefit of knowledge alone may not be a sufficient justification for experimental simulation investigations."[33]

According to Greene, we need an *overwhelming positive value* for proving that we live in a virtual reality. "Given, therefore, that extrapolation from current trends in computing and simulation technologies point to termination as the possible result of a successful simulation probe, how should we think of their expected value? The most obvious benefit is that of adding to our knowledge of reality. Would that benefit outweigh the potential cost?"[34] Without an overwhelming positive value we should not do anything to screw around with the simulation. No probes. No tests. No building our own ancestor simulations. Put the gun down. The fate of the universe is at stake.

If we cannot come up with a *very strong reason* indeed to test any aspect of simulation theory, we should not try it at all.

## AUXILIARY ANNIHILATIONS AND THEIR DISCONTENTS

Although the two annihilation scenarios outlined above are the most commonly discussed in the literature, *there are even more*. In "Simulation Typology and Termination Risks" Alexey Turchin, Michael Batin (both members of the Foundation Science for Life Extension), Dr. David Denkenberger (Global Catastrophic Risk Institute), and Dr. Roman Yampolskiy (a Senior Member of the Institute of Electrical and Electronics Engineers and the Artificial General Intelligence Society) offer a slew of other possibilities. Their paper "explore[s] what is the most probable type of simulation in which humanity lives (if any) and how this affects simulation termination risks."[35] Their conjectures are as inventive as they are amusing.

---

[31] Greene, "Termination Risks of Simulation Science," 22-23.
[32] Greene, "Living in a Computer Simulation? Let's Not Find Out."
[33] Greene, "Termination Risks of Simulation Science," 25.
[34] Ibid., 24.
[35] Turchin et al., "Simulation Typology and Termination Risks," 1.

My personal favorite is also the most paranoid: the *Always Game Over* scenario. What if we, the sims, fulfill whatever it is we are intended or expected to do here? We "beat the game" so to speak. That would also potentially result in a shutdown. "The simulation finally solves the unknown-to-humanity task, and there is no need to run the simulation any more."[36] No need to continue running this simulation if we accomplish our mission. You turn the game off once you have beaten it, right? The problem here is, of course, we have no idea what "beating the game" looks like. Ergo, we have no idea how to avoid "winning" and thus bring the simulation to its purposeful conclusion. But it gets even more complicated! If we somehow *knew* what we were supposed to do to beat the game—and thus incur a shutdown—and then *purposefully avoided it*, that could *also* result in a shutdown because *we would be trying to willfully dodge what we as sims were made to do in the first place!* The posthumans would naturally see this and decide we've compromised our own purpose for existing. Game Over again. I love this version because it is the ultimate no-win, double-bind scenario. You're damned if you do and you're damned if you don't.

According to the authors, another shutdown could happen after human beings cross some kind of "threshold." They write: "The threshold could not only be AGI [Artificial General Intelligence] creation, but some moral threshold, such as a large population, intense suffering, or sentient computers. Alternatively, the threshold could be technological, such as self-replicating nanotechnology or genetically engineered babies or philosophical, like the capability of thinking about being in a simulation clearly."[37] These situations could make such an immoral mess of the simulation that the posthumans running it decide to shut it down. Here, yet again, we do something we are not supposed to do—but, we can never be quite sure of what that prohibition might be. The posthumans "rage quit" the simulation.

The simulation could also be shut down due to an accumulation of glitches, which render the simulation unusable. "The beings inside the simulation start to realize that they are in a simulation, so it is not a simulation anymore and cannot be used as such. Pure philosophizing about being in a simulation would likely not be enough to cause termination, as such theorizing is likely inevitable. However, observing some undeniable glitches would destroy the purpose of the simulation as an illusion of reality, unless the simulation is rewound and corrected."[38] Too many glitches, oddities, incongruencies—too much chaos—and, once again, we have a shutdown.

One of their starkest hypotheses is that our universe "[is a] simulation designed to test different theories about the end of the world."[39] In this truly dark science fiction-like scenario, the whole point of our simulated universe is to see how we destroy ourselves. This could be of certain value to the hypothetical posthumans running such a simulation so they avoid such a fate themselves. Herein, we exist solely as research subjects to see just how we become the authors of our own doom for the benefit of a wholly other group. In the words of Agent Smith towards the climax of *The Matrix Revolutions,* herein "The purpose of life is to end."[40]

Then there is simply Accidental Shutdown, the *Whoops* scenario—the posthumans turn off our simulated universe entirely by accident. Someone trips over a surge protector. They forget to update their computer, or pay the electric bill. *Zip*. No more universe for no good reason at all.[41]

---

[36] Ibid., 16.
[37] Ibid., 17.
[38] Ibid., 16.
[39] Ibid., 16.
[40] *The Matrix: Revolutions,* directed by Lana Wachoski and Lilly Wachowski. (Warner Bros., 2003.)
[41] Turchin et al., "Simulation Typology and Termination Risks," 16.

The authors conclude their harrowing list by adding, “From all listed above ideas about the termination of the simulation, the idea of the termination because of the computational overload appears to be the most probable, as it inevitably follows from the setup of the SA [Simulation Argument].”[42] In the end, it is cost inefficiency once again.

We have explored all these possibilities not to consider their validity, but to reach an important realization: under the ancestor simulation hypothesis *our universe is radically unstable*, far more unstable than human beings have ever imagined. That being said, and indeed, more to the point, what we are looking at here is not just a myriad of reasons *why* our universe could come to an end—what we are *really* looking at is a tangle of proposals that all inevitably lead to such an outcome. In essence: hypotheticals. Each represents an offshoot of the primary assumption: that our universe is an ancestor simulation created by posthumans. All these doomsday ideas really amount to armchair intellectual exercise. This is all pure speculation.

In fact, we can think of it as something of a parlor game. Ask a few of your closest friends: *if this universe is a simulation, and you assume the position of an outside moderator watching it, what would make you turn it off*? Perhaps a common answer would be why post people turn off a video game when work, school, and other obligations don’t need to be considered; it simply becomes boring.

Ultimately, much of what we just went over can be easily swapped out. Change the language of the simulation hypothesis and ask it instead in theological terms: *if I were God, why would I create the universe? And, why would I destroy the universe?*

Have fun.

## ENTER THE CONTRARIANS

There is no shortage of end-of-the-world scenarios when it comes to the ancestor simulation hypothesis. It really is a ‘pick your favorite flavor’ of universal doom. But, not all thinkers agree with Greene and the others that the simulation theory makes the universe an unbelievably precarious place.

Dr. David Braddon-Mitchell (Professor of Philosophy at the University of Sydney) and Andrew J. Latham’s (Postdoctoral Research Fellow in the Department of Philosophy and History of Ideas at Aarhus University) paper “Ancestor Simulations and the Dangers of Simulation Probes” offers counterpoint after counterpoint *against* the arguments for the destruction of our simulated universe. Based on the paper’s humor and playfulness, the authors seem to have had a refreshingly good time writing it. Indeed, when it comes to testing simulation theory, they are positively gung-ho.

They aim a large portion of their attack on the proposal that gaining simulation awareness thanks to a simulation probe (becoming Savvy) necessarily results in the destruction of the universe, taking the arguments apart piece by piece. To begin with the most pressing question: *shouldn’t the posthumans end the simulation once their sims figure out what’s really going on?* “[P]erhaps they are more likely to be interested in simulating what happens when there is a successful simulation discovery, than in simulated societies that never discover this.”[43] Sure, why

[42] Ibid., 17.
[43] Braddon-Mitchell and Latham, "Dangers of Simulation Probes," 3.

not? A group of sims that *realize* they are in a simulation are bound to react to that in plenty of interesting ways. Similarly, Braddon-Mitchell and Latham flip the whole issue on its head and propose: "In fact, such simulators might even terminate simulations in which the simulation discovery looks like it's not going to be made."[44] In this case, if we are in a simulation, and *we never realize it as such*, maybe we are useless to the posthumans on the other side of the screen!

Some examples that Braddon-Mitchell and Latham come up with regarding the supposed dangers of the destruction of the universe are downright outstanding:

> Preston Greene has suggested that even if the simulators know they are themselves in a simulation, they might nonetheless have interests that would be derailed should their simulation discover that they are in a simulation. This is no doubt true; if they are interested in the results of various ways a mid-21st century pedophilia scandal might go before the world has discovered itself to be a simulation (to use Greene's example) this might get disrupted by the discovery in their simulation that it is a simulation. But there are countless things the simulators might be interested in, and for each interest different developments in the simulation might be a problem. If they were interested in the discovery that the simulated world is a simulation, a pedophilia scandal emerging in their simulated world might disrupt things in a way that masks the usual effects of the simulation discovery![45]

With examples ranging from Russian roulette to pedophilia scandals, all in effort to explain how we might be living in a Matrix, the thinkers at the forefront of simulism are engaged in a wild discourse to say the least.

Now, let's imagine that we realize we are in a simulation. Here the authors make a wonderful point: so what? They write, "We would still be confronted by massive problems that would need to be solved, such as climate change, pandemics, war and so on."[46] This is a critical point—yes, even if we *do* realize we are in a simulated universe, there will still be the human social, economic, and political systems as they exist today. We would still have to deal with the mess we have made of the planet and the inequalities in our own system. (An interesting complement to this issue, as we will see in the following section on Tom Campbell's *consciousness-based simulation theory,* is that our simulation is designed to encourage prosocial evolution). However, again, even if another form of simulism turns out to be true, it might simply amount to "a (very) interesting discovery about the fundamental nature of our world."[47] Possibly no more than that. Like the discovery of the quantum, or heliocentrism before that, we will learn more about the nature of reality but there will still be the overarching issue of human barbarism to deal with.

When it comes to all the worry of universal calamity that is so common in the field, they write, "There is no reason to fear discovering that we are in a simulation, at least insofar as it might lead to our termination. Either it makes no significant difference (as we suspect) to what simulators care about, in which case there is no extra risk of termination, or it does make a difference. And, if it does make a difference, then the difference is likely to be one that the simulators might also be interested in."[48] In their last analysis of actually going forward and testing simulation theory via experiment, Braddon-Mitchell and Latham are unambiguous: "let's do it."[49]

---

[44] Ibid., 4.
[45] Ibid., 4, n5.
[46] Ibid., 5.
[47] Ibid., 5.
[48] Ibid., 10.
[49] Ibid., 1.

# ANALYZING THE SIMULATION ARGUMENT

Let us take a step back and consider everything we have gone over.

The overwhelming problem with all of this speculation concerning the ancestor simulation hypothesis is that it is *complete* speculation. It relies entirely on a series of assumptions as well as yes-or-no ("maybe"/"maybe not") conditionals—*beginning with the Simulation Argument itself*—all of which are most likely ultimately unknowable. These assumptions include:

(1) Posthumans exist.
(2) We are sims in one of their simulations.
(3) Some sims may be non-conscious philosophical zombies.[50]
(4) The purpose of their simulation is either for research or entertainment.
(5) The purpose of their research or entertainment is unknown.
(6) We could do something that irrevocably compromises their research or entertainment.
    (5.a) Making our own ancestor simulation destroys our universe.
    (5.b) Executing a successful simulation probe and thereby becoming ontologically Savvy destroys our universe.
(7) Our simulated universe could be shut down for yet another unknown reason.
(8) And/or the purpose of our simulation is to become Savvy.
(9) And/or it does not matter whether or not we become Savvy.

This is an absolute mess by any standard.

There is seemingly no place to stand in the ancestor simulation hypothesis; no point of view can be maintained for long, because there is always another "if/then" issue waiting in the wings, along with the unknowable motives of a superintelligent species that exists outside of our universe.

Ultimately, the thinkers wrestling with the ancestor simulation hypothesis are as creative as they are incongruous. Their project is largely speculative and unlikely to climax in any sort of wide agreement. It opens with multiple assumptions, many of which are also unlikely to find agreement. There is no account of the origin of 'base reality', nor agreement on the exact purpose of the simulation itself. Some within the field argue pursuing these hypotheses could destroy the universe, others argue it doesn't matter, others argue it would advance our species, while others still argue everything is ultimately completely beyond our purview or control. All remains speculative.

The *singular* accomplishment of the ancestor simulation hypothesis is that it has put simulation theory on the table in a widely popular way. *That* is their singular (one and only) vital contribution. However, beyond that potentially critical insight, their speculative enterprise will likely continue to produce novel ideas that are either ultimately untestable, hypothetically dangerous, or philosophically inconsistent. It will continue to be a discourse of *what if*'s and *but if*'s. It is an unstable framework that sets the stage for an even *more* unstable universe. It leads to more questions than answers, and although that can be philosophically interesting, even exhilarating—in this case, we can't even decide on what questions are worth asking! It is

[50] Bostrom, "Living in a Computer Simulation?," 13.

therefore a *null hypothesis*. It is somewhat interesting, but you can't really get anywhere far with it.

In sum, while entertaining, the ancestor simulation hypothesis represents a *close, but no cigar* scenario. They have valiantly kicked the ball over the stadium fence, but it has not landed anywhere near the goal line.

Now, just because the ancestor simulation hypothesis leads us to a labyrinth of warped mirrors does not mean it is the end of simulation theory. It means that if there is anything to simulism we have to look elsewhere.

## THE INTELLIGENCE TEST

Before we move on, it is worthwhile to briefly go over one of the specific assumptions that we have been considering: the *ancestor* aspect inherent to the Simulation Argument. Remember, most thinkers use Bostrom's model simply because it is the most well-known form of simulism. Therefore, the vast majority of theorizing is based entirely on the idea that we are in some kind of *historical simulation* monitored by posthumans, and that simulation can be compromised somehow.

For instance, "Greene argues that the utility of such experiments is negative as they either create negative results, in most cases, or they confirm that humans are in a simulation, but this makes the simulation useless for their owners as a historical reconstruction and then the simulation could be terminated."[51] If we become Savvy, and our world is a simulation of the past, we could screw up the purpose of our universe; things would not unfold correctly or usefully. However, as Braddon-Mitchell and Latham pointed out, *maybe* our simulation is a simulation of 'base reality,'and that we are a historical simulation that *is* allowed to become Savvy, because the posthumans would want to know what would happen if their sims realized it as such. Here's the whole intellectual gordian in all its confounding glory.

Now, if you were the posthuman in this scenario, a simple fix would be preventing your simulated lifeforms from ever thinking such a thought. Like the semi-sentient robots in the 2016 series *Westworld*, you could program in the response, "It doesn't look like anything to me," whenever they are confronted with evidence of their true nature.[52] If you have the ability to program a whole universe, couldn't you also conceivably have the ability to control what your sims think or don't think? Why not?

But, again, what if the posthumans actually *wanted* a Savvy life form in their simulated universe? I argued for such a case in my aforementioned piece, "Why It Matters That You Realize You're in a Computer Simulation":

> We would actually want a Savvy intelligence inside our simulated universe. The reason why is very simple: If we only have access to observe intelligent lifeforms that are restricted to not knowing that they are in a simulation, then our own sample pool and thus knowledge base will always be restricted to intelligences that are out of the loop. Complex level lifeforms [...] would by definition always already be operating from an ontological ignorance of the true nature of their environment [...]

[51] Turchin et al., "Simulation Typology and Termination Risks," 2.
[52] "The Original," *Westworld*.

> Therefore I, as part of the original hypothetical simulation-running team, would be extremely hesitant, if not downright protective, of that Savvy sample's survival and evolution—that is if I were to interfere at all. What could possibly give me more insight into what I, the original simulation creator and maintainer, have done than this Savvy sim living in my ever-growing mock universe? Would I really throw out the sim that realized they were in *The Sims*? Indeed, evolving a sim that realizes they are in *The Sims* might feel like I'm actually getting my computational weight's worth—that goes especially if I was putting in all this effort to power and evolve a life-bearing simulated universe in the first place. If our simulated universe is inadvertently an intelligence test for the evolving life forms inside it, then I'd hope we grow a winner. A sample so intelligent that it can actually see the code at the edge of matter is likely a sample we'd benefit from studying. It is conceptually not too far removed from teaching great apes to sign.[53]

Furthermore, I argued in a 2016 paper that if a simulation probe is for some reason or another undesirable to our posthuman overlords, rather than a complete shutdown or termination of an entire reality with all its conscious players, a *far cheaper* alternative would be that whatever machinery we use to detect this possibility will simply breakdown. In a single sentence: "[Y]our machine will mysteriously and consistently malfunction and fail ('gremlins')."[54] This easy tweak would be undoubtedly *more* cost effective than, again, deleting a whole universe along with all its evolving life forms.

If these supposed posthumans were advanced enough to make a whole simulated universe, we would imagine that being able to influence or restrict their sims' ability to conjure with some ideas, or simply sabotage our experiments would be an easy thing to do. No—the catastrophizers skip these pretty obvious beats to write titillating, borderline science fiction theory papers that earn them clicks as they lead us intellectually to both doomsday and nowhere at the same time.

If there is *anything* worth destroying, it is not the universe—it is the extraneous conjectures surrounding Bostrom's otherwise cogent Simulation Argument, and all the useless places it leads otherwise intelligent minds to.

Now, in the following section we will cast off the yoke of the ancestor simulation paradigm completely and move on to another wholly different form of simulism.

## CAMPBELL'S CONSCIOUSNESS-BASED SIMULATION

We have been exploring the possibility of our simulated universe suddenly being switched off for virtually any reason. However, this is conditioned by *which form of simulism* we assume our universe to be. While Bostrom's ancestor simulation hypothesis is the most well-known version of simulism (and apparently highly prone to such an outcome) it differs considerably from Tom Campbell's MBT version of simulism.

A keen reader might have noticed that in the ancestor simulation hypothesis the ultimate origin of 'base reality'—where the assumed posthumans live—is unknown or irrelevant to the Simulation Argument, and the ultimate purpose (*telos*) of the ancestor simulation is *also* largely

[53] Edge, "Why It Matters That You Realize You're in a Computer Simulation."
[54] Eliott Edge, "Breaking into the Simulated Universe," *Institute for Ethics and Emerging Technologies* (IEET), October 30, 2016, https://web.archive.org/web/20171230092123/https://ieet.org/index.php/IEET2/more/Edge20161030.

unknown. Campbell's version of simulism *does* have an origin story for the cosmos, as well as a purpose for its existence. We will explore this in the present section.

Every scientific or philosophical model comes with assumptions; the fewer the better. As Einstein said, "The grand aim of all science is to cover the greatest number of empirical facts by logical deduction from the smallest number of hypotheses or axioms."[55] Campbell's book *My Big TOE* (Theory of Everything), which shorthands as 'MBT theory', opens with two assumptions—a number as low as a good critical thinker could hope for.

(1) Consciousness is fundamental.
(2) Evolution is fundamental.[56]

Assuming that evolution is fundamental may not be too difficult for most to accept. A recent multidisciplinary research effort has suggested that evolution could be a missing law of nature, not only applying to biology but to all physical processes.[57, 58] However, the assumption that consciousness is fundamental might be more difficult for some to accept. After all, as psychologist Dr. Susan Blackmore points out, "Part of the problem is that 'consciousness' has no generally accepted definition in either science or philosophy despite many attempts to define it (Anthis, 2022; Niikawa, 2020; Nunn, 2009)."[59] Indeed, biophysics and psychophysics researcher Dr. Ram Vimal identified at least forty different uses of the word, noting that "this list is by no means exhaustive."[60] Thus, whenever consciousness comes into a discussion, we must always define what we are speaking about.

In MBT, consciousness is described as an "awareness with a choice."[61] According to Campbell, this awareness must be an information system; without information there is no consciousness/awareness and vice versa.

Why it is modeled as an information system is because, under Campbell, this awareness must have four key features to be considered conscious: data input (experience), data memory, data processing (pattern-matching, or sense-making), and a self-modifying feedback loop (learning). Have these features, along with choice, and you have consciousness in MBT theory.

---

[55] Alice Calaprice, *The Ultimate Quotable Einstein* (Princeton University Press, 2011) 402. "Quoted in Lincoln Barnett, 'The Meaning of Einstein's New Theory,' *Life* magazine, January 9, 1950."
[56] Thomas Campbell, *My Big TOE: A Trilogy Unifying Philosophy, Physics, and Metaphysics* (Lightning Strike Books, 2003), 182. "The fundamental process of evolution along with primordial consciousness as a fundamental source of structurable energy are the two basic assumptions on which *My Big TOE* is based. Everything that follows will be logically derived and explained from these two fundamental assumptions."
[57] Michael L. Wong, Carol E. Cleland, Daniel Arend Jr., and Robert M. Hazen, "On the Roles of Function and Selection in Evolving Systems," *Proceedings of the National Academy of Sciences* 120, no. 43 (October 16, 2023): e2310223120, https://doi.org/10.1073/pnas.2310223120.
[58] Eric Ralls, "'Missing Law of Nature' Explains the Evolution of Everything That Exists in the Universe," *Earth.com*, October 19, 2023. https://www.earth.com/news/missing-law-of-nature-explains-the-evolution-of-everything-that-exists-in-the-universe/. "The authors assert that evolution is a fundamental process that extends to all complex systems in the universe — from celestial bodies to atomic structures."
[59] Susan Blackmore and Emily Troscianko, *Consciousness: An Introduction*, 4th ed. (London: Routledge, 2024), 1.
[60] Ram Vimal. "Meanings Attributed to the Term Consciousness: An Overview." *Journal of Consciousness Studies* 16, no. 5 (2009): 9-27. "I here describe meanings (or aspects) attributed to the term consciousness, extracted from the literature and from recent online discussions. Forty such meanings were identified and categorized according to whether they were principally about function or about experience; some overlapped but others were apparently mutually exclusive—and this list is by no means exhaustive. Most can be regarded as expressions of authors' views about the basis of consciousness, or opinions about the significance of aspects of its contents. The prospects for reaching any single, agreed, theory independent definition of consciousness thus appear remote. However, much confusion could be avoided if authors were always to specify which aspects of consciousness they refer to when using the term."
[61] Tom Campbell. "Tom Campbell: Conscious Computers and Consciousness in a VR," YouTube video. July 2, 2021, [00:43], https://youtu.be/rLhWl7si_aQ?si=xAipV0tm7kpucySa

In its simplest form, information is binary; digital; 1 and 0; this way or that way. This information system has to keep evolving or it will "die"—which Campbell calls the *Fundamental Process*: evolution; evolve or die.

Now, a note here on what Campbell means by "dying." All systems, whether we are talking about information systems or thermodynamic systems, are subject to *entropy.* The System responsible for our universe is no different. Entropy is a measure of randomness or disorder. "More entropy means more disorder and less energy that is available to do work. Conversely, less entropy means less disorder (more order and structure within the system) and that more of the system's energy is available to do work."[62] The more disorder a system has, the more noise, randomness, or chaos it has. The *less entropy* in the system the more orderly it is. If a system is subject to very high levels of entropy, it is effectively useless, like trying to make out a song or a message through intense radio static. Ergo, under Campbell, the primordial information system of consciousness was under pressure to keep its entropy low (evolve and thrive), that way it did not succumb to overwhelming chaos/noise and become effectively "dead."[63]

So, keeping entropy in mind, Campbell starts the story of existence first with a simple, dimly conscious information system that is under the pressure to evolve or die—it then must choose how to deal with this circumstance. This forced said information system to *strategize*: it generated patterns and processes of information that were stable and sustainable. It couldn't stay still because static states always eventually become unstable and eventually succumb to entropy over time: they inevitably break down. So this information system had to come up with more and more ways to keep itself from devolving. Thus it became dimly aware, then it became somewhat intelligent, then it became very intelligent indeed—all in the service of trying to keep itself (its information) alive and useful.

In MBT, this culminates in two major breakthroughs: (1) our digital system realizes it can develop itself more quickly and efficiently by splitting itself up into pieces, like a hard drive that's been partitioned, into what Campbell calls *individuated units of consciousness* (IUOC)—"players" (2) the System then discovers that simulated environments are useful because environments create rules and consequences, thus providing more evolutionary traction by creating valuable feedback. Both of these breakthroughs furthered the system's evolution by lowering its overall entropy, lowering its chaos.

## FIRST BREAKTHROUGH

In terms of the first breakthrough: one information system (a computer or mind) will only have its own information to parse through, to experience. That's inherently limited. But if you have *two* information systems, you instantly have more possibilities open up. This provides for vastly more richness. Now, if you have dozens, then hundreds, then millions of information

---

[62] Campbell, *My Big TOE*, 197.

[63] Ibid., 298, "Digital consciousness systems do not deteriorate with time like biological systems, though they can de-evolve—that is, evolve into higher entropy, less significant, profitable, and viable states. AUM ["Absolute Unified Manifold" a term Campbell no longer uses to describe "The Big Computer"/"The Larger Consciousness System"] achieves self-optimization and growth through the exploration of the possibilities by implementing the Fundamental Process. AUM can eventually figure out how to willfully boost its quality (lower its entropy by utilizing its potential and organizing its bits more effectively) once it realizes that profitability is a function of the intent that drives its choices. So it is with us."

systems all talking to each other, you now have a lot of possibilities for growth and change to unfold. You now have lots of information, lots of opportunities for organizing, lots of possibilities for evolution because you have such a robust environment. This is the evolutionary value of having a singular information system break itself up into pieces—a lot more novelty and diversity becomes possible very quickly.[64]

People who argue for the colonization of space make a similar case. If our home planet gets wiped out for some reason, then that's it for us and all the life here. Game over. But if we are on many planets, then all kinds of new opportunities as well as a higher chance of survival opens up. For evolution, higher diversity as well as a higher population is always better than less diversity with a lower population.

The important point to keep in mind here is that these individuated units of consciousness also have choice; they are not merely programmed automatons running scripts. They too make decisions based on the available resources of their own individuated information sets; their personal database.

## SECOND BREAKTHROUGH

Now let's look at the second breakthrough: rules, protocols, constraints and eventually environments. Without rules, you do not have a well-defined system. Rather, you have a higher probability for chaos, mayhem, noise, entropy and thus de-evolution. No meaning can be generated. Without some kind of ruleset *all bets are off and anything goes*. The rules for our evolving consciousness system would be simple at first and then slowly become more complex over time. Rules also provide for more traction, meaning, consequences, depth, novelty, and so on. An information system without any rules quickly becomes useless. Rules and consequences become useful if the System is trying to keep itself organized and evolutionarily profitable.

So, at first, tackling this goal of survival manifests with the evolution of a kind of "big chatroom" for these individuated units of consciousness to communicate with each other. That space and those "communication protocols" constitute the first simulated environment, the first VR in Campbell's cosmogony. However, the big chatroom quickly plateaus, because *anyone can say anything in a chatroom*. "So now we need a richer environment, not just a big chatroom environment. [Because] in the big chat room environment, it's just you and a whole bunch of other things you can talk to or not talk to, and that's it. There's really no other rules. So it's just those simple rules. There's not much strategy. There's not much going on that's really helped you lower your entropy."[65] Campbell elaborates:

---

[64] Tom Campbell, "Eliott Edge and Tom Campbell in Conversation Part 1 of 2," YouTube video, July 25, 2023, [12:20], https://youtu.be/GxzI3AOHtpY?si=cT5dLANJvWD0z_bE, "It's only a logical step forward to decrease entropy, which is another way of saying, that to give it more possibilities of what it can evolve into, more choices, more things that it can come up with, is that it takes a piece of itself and kind of partitions off that piece. And now there's it, and this piece of a thing, and that piece also has free will. It can also make choices. So now the possibilities of two things both with free will make a huge number of things that can happen, because these two aren't the same thing. It was this one monolithic thing, well that's very limiting, but two things won't necessarily agree. Two things can have a discussion. Two things can get in a fight. Two things can fall in love, you know? Two things can do all kinds of things that one thing is very limited in doing, so it just increases the possibilities, and so it makes more of these [things]. So now there's thousands of these individuated units of consciousness interacting with it, and for a while what they can do together in cooperation is interesting and they grow and evolve, but that also becomes plateaued."

[65] Ibid., [15.04]

> So the System decides that it needs to create another virtual reality that has more interesting choices: choices that are moral choices, ethical choices, life and death choices, choices that are very meaningful, and choices that are very interactive [...] It's a very interactive environment because the chatroom is only as interactive as you kind of want it to be, and when you don't want it anymore, well, the interaction just goes away, you know? You just don't stop chatting, now you don't have any interaction, but it wanted an environment that wasn't that easy just to walk away from.[66]

With this second breakthrough we have moved from a system of simple information exchanges between unique units, to a rich consequential environment for those units which ups the ante as it were for the process of evolution.

## SUMMARY OF THE CONSCIOUSNESS-BASED SIMULATION MODEL

We started with a dimly aware information system that is under pressure to evolve or die. It evolved first by breaking itself up into pieces. Those pieces then communicated with each other (shared information), as if in a chatroom. The chatroom reached its evolutionary peak. Then, to keep evolving, richer environments with more rules and depth were created. This continued on until eventually our universe was developed. Our simple information system grew into *The Big Conscious Computer*, or the *Larger Consciousness System,* if you will.

This is *a completely different* outlook from posthumans monitoring their ancestor simulation on an external computer. In a consciousness-based simulation, our universe is *not* an ancestor simulation at all. Rather, our simulated universe is ultimately *a tool* that consciousness has developed to evolve itself—an essential, important distinction. This tool amounts to a *social system* wherein the choices made by the individuated units of consciousness are part of a survival strategy for the entire conscious computer system to evolve itself.

Back to the original question: *What happens if we find out we are in a simulation?*

In Campbell's MBT, this piece of information in itself is *neutral-to-highly beneficial*. Its *potential benefit* has to do with whether or not it lowers our entropy as individuals and a society. *Conceivably,* we could very well be *greatly rewarded* by our discovery that we exist in a computer simulation *if* with that discovery we come to understand that we, as individuated units of consciousness, are in this virtual reality to evolve prosocially. This would be a massive benefit to us and the System that we are part of. Our evolution is the System's because we and it are the same thing. We are all One in Campbell's model, because we are all consciousness itself.

Now, if we realize that our environment has a purpose (reduce entropy) and we take that on ontologically, then becoming Savvy has a huge potential benefit across the board. If we realize we are in a simulation, but we believe that means that effectively nothing truly matters, that that could be potentially dangerous for a myriad of reasons. Thus, becoming Savvy includes the caveat problem-question: *What kind of simulation are we in?*

---

[66] Ibid., [15:50]

## ORIGINS, SHUTDOWNS, AND VALUE

In the earlier sections, we briefly mentioned the ancestor simulation hypothesis does nothing to conclusively solve the problems of the origin or the ultimate goal of our universe—simulated or otherwise. However, that's not the Simulation Argument's purpose. Rather, the Simulation Argument is a logical philosophical argument based on Bayes theorem ("one of these conditions must be true") that, when in play, leans strongly towards *we are almost certainly in a simulation.*

The ancestor simulation hypothesis and the Simulation Argument function as logical philosophical frameworks; however they reach their *terminal value* (they no longer become useful) once they are asked to go beyond their three postulates. Whenever the postulates are asked to do *more* than prove the likelihood of us living in a simulated world, they fan out in every possible speculative direction. Ask, *Where did the posthumans come from? What is the purpose of the simulation? Do the posthumans want us to become aware of the simulation? Should we fear a shutdown?*—and you end up with endless conjectures and unknowns. But the Simulation Argument was not built to answer these questions. The Simulation Argument was built to merely point out *it is far more likely than not that we are in a simulation.*

Where Bostrom is laying down a logic-based foundation for *simulated reality*, Campbell is providing a scientific rubric. With Campbell, things begin with a scenario: a simple consciousness needed to evolve or die. That's all you need to explain not only the origin of consciousness and life, but the origin of our complex universe as well.

In an ancestor simulation, our virtual reality world is created by 'intelligent beings.' In Campbell's view, our virtual reality is created by an evolving consciousness system. Part of what makes this such an important distinction is due to the *assumed existential divide* between the intelligent beings and their sims. For the intelligent beings running our universe, *it might not matter to them much at all what happens to us denizens of the simulated universe*—therefore, a simulation shutdown type scenario *is* feasible. Whereas in Campbell's model, it absolutely matters what happens to us. This is because The Big Computer running this universe is evolving itself *through us as we evolve ourselves.* We are part of its strategy for survival. This is not parasitic, or predatory, or manipulative. It is *symbiotic*. We exist, according to Campbell, as little partitioned parts of the Larger Consciousness System, The Big Computer, that has created us and this virtual universe to survive. We and it are the same thing. We are involved in the same mission.

Hence the difference comes down to an issue of *investment.* The questions become:

(1) How invested are the posthumans in the survival of their sims?
(2) How invested is the Big Conscious Computer in the survival of its sims?

The answer to the first question is basically unknown.

The answer to the second question, however, is *very invested indeed.*

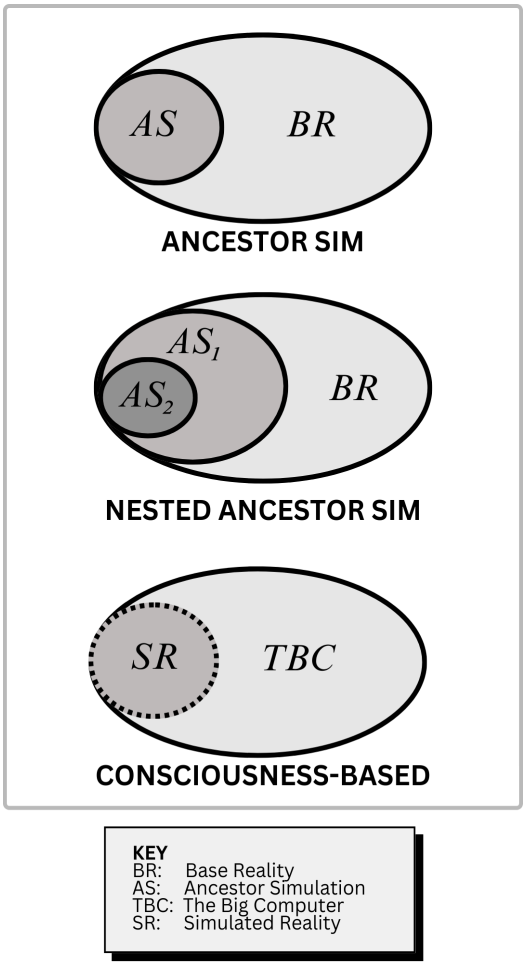


*Table #1: Simulated reality types and their relation to the supersystem (computer) that supports them.*

If our universe is an ancestor simulation, shutdown is something that needs to be seriously considered, according to the thinkers in that field. As Bostrom said, "[W]e suffer the risk that the simulation may be shut down at any time."[67] But if our universe is a consciousness-based simulation, *shutdown is possible, but unlikely.* It is *so very unlikely* that it approaches 0. In Campbell's model, The Big Computer behind our universe is smart— smart enough to tweak the dynamics of *all* of its operations, so that it has strong control over how much is being used and *how* it's being used at any given time. It is, after all, conscious.

Under a consciousness-based simulation, shutdown is rare to the point of being a non-problem, and becoming Savvy is a helpful step to our overall evolution *if* we realize with it that we are supposed to lower our entropy (get along with each other for our overall survival). All that being said, the System, being aware, can delete information that is inconsequential, irrelevant, or no longer of any value—"I think as it goes the System is smart enough not to load up its databases with trash information."[68]

In an evolving consciousness-based simulation, the System *will* have to off-load information that it believes will not be useful or accessed anymore. It will either be scrapped

[67] Bostrom, "Existential Risks," 7.
[68] Tom Campbell, "Tom Campbell Answers Your Questions Vol. 12," YouTube video, January 26, 2022, [14:25], https://www.youtube.com/watch?v=g4vNzV4gP5Y.

completely, or reduced to an approximation, a thumbnail summary version of events, rather than waste space by retaining every single atom of detail.[69]

Campbell illustrates:

> This is digital space. You can do all kinds of wonderful things. You can just stop the virtual reality at a place, let's say, you know, 50 years before the great catastrophe. And then you can just change a few things and then start it up running again. In other words, you just make a branch off, or if there was a big meteor coming this way that was going to land and stir up so much dust that would blank out the sun for five years and then everything dies, right? [...] [T]hat's a virtual meteor. You see, it's part of the probability sets that that meteor's there. [...] Well, it could just go out and make that disappear or make it go someplace else because we haven't actually seen it yet. So it's not on our radar or, you know, could you make it just make that disappear if it wanted to. Or, if it needed to, it could basically erase, say, 90% of the population because this experiment just isn't working out very well. You know, it's going downhill and to start over. We could take those 10% of the people and try to build it back up again and see if we couldn't do better or see if we make the same mistakes. So there's lots of options in a digital system. Lots of things. And it could just take the beings here and, you know, start another virtual reality, very similar to this one, same rules that, you know, it could just start it up and repopulate it with the same beings. And for a while, everything would seem very strange, but, you know, a thousand years later, nobody would know that anything ever happened. [...] So there's lots of options. But we, as individuated units of consciousness, will always have a playground to play in to make choices because the system needs that. We are part of its strategy to evolve. As we evolve, it evolves. We're a piece of that consciousness system. So as we evolve, our evolution is also its evolution. So it has a motivation to keep us learning as quickly and efficiently as possible.[70]

In Campbell's MBT, The Big Conscious Computer has a multitude of options available to it to ensure its overall project of evolution succeeds. Why destroy a whole reality when you can simply rewind it, or modify it?

Now, under an ancestor simulation, it is assumed that our universe exists either as a research project or as entertainment, and the assumed "intelligent beings" running this simulation are, for all intents and purposes, *existentially separated* from the goings-on here (their survival might not strongly depend on our existence—much like how the survival of your favorite gaming character build does not impact you in terms of your own day-to-day survival.) Under a consciousness-based simulation, our universe exists *exclusively* as a growth and evolution engine; an entropy-reduction trainer. What happens here is *essential* to the overall System's evolution and survival—its health. The System could throw out this whole universe, but Campbell believes it's probably disinclined to.

What would be *more likely* than a simulation-wide termination, is human beings simply destroying themselves before they reach prosocial harmony. The Big Computer *would actually have more of a learning opportunity* through that unfortunate outcome than it would from switching the game off. A simulated universe is an evolution machine. Basically, a simulation would only be shut down if it repeatedly failed, eon after eon, to reduce the entropy of its individuated units of consciousness (players). So the only reason for a consciousness-based

---

[69] Idib., [17:00], "If it's not important, it's not kept. So the database doesn't keep everything forever, it keeps what it needs for now [...] it throws out stuff that nobody will access. Maybe it just looks at that date and says well, you know so many years have gone by and there's never been any access. So I'll just dump that because nobody's going to ask for it. If they do, if somebody says well I want to know about Grog [a caveman], or I want to know about my great-great great great great great whatever grandfather, then they'll come up with something that is approximately true."

[70] Tom Campbell, "Tom Campbell: Talks with News for the Heart on Virtual Realities." feat. Laurie Huston. YouTube. March 31, 2019, [38:20] https://youtu.be/qWWMhqhZUwY?si=mBmbaSYCRhi2wS47

simulation to entertain the idea of the termination of a given simulation would be very long-term stagnation—a given simulated universe fails to be useful for consciousness evolution over a very long time.

In simplest terms, what would cause the termination of a given simulated universe in Campbell's model is a universe that failed to produce any life whatsoever, or that life never evolved into anything remotely useful (self-contained, non-interacting units). In the case of our planet, with its massive biodiversity (the estimate being around 8.7 million plant and animal species, only 1.2 million of which are known) we are probably in the clear.[71]

Now, let's return to our original question: *should we test the simulation hypothesis or not?* Greene asks us to take the issue quite seriously before moving forward.

> [I]t is possible that our simulators would give us some reward for discovering that we live in a simulation. This possibility perhaps has the power, in principle, to provide a sufficient justification for experimental probes when combined with the possibility of knowledge acquisition. As far as I am aware, however, the possibility of reward has never been seriously put forward in defense of such investigations. More reflection on these issues is therefore required before we can reasonably regard experimental simulation investigations to have positive expected value.[72]

Under Campbell's framework, realizing that we are in a game with rules and a purpose is a necessary step in the evolution of its players. However, although that stage must be reached organically, *it does not necessarily have to be reached through scientific means.* Indeed, some religions and philosophies have intuited some aspects of Campbell's model: the universe exists to help us evolve prosocially (i.e. humanitarianly, or "spiritually.") This is for the betterment of all—not just human beings, and not just purely selfishly, but for the whole system. It is only now that through the dominant ideology of *science* that this understanding needs to be expressed scientifically, logically, and rationally—rather than culturally, religiously, or poetically—for it to be widely believed. Otherwise, human beings have thought and expressed similar sympathies throughout history.

Why it needs to be reached *organically* is because part of our simulated universe's purpose would be for The Big Conscious Computer to create an environment (i.e. a program) that was effective at evolving the sims subject to it. If this universe is a simulation, it is also, *ipso facto*, a program—an experimental one at that. Herein, our simulated universe is *also* an ongoing research project aimed at finding the most effective means of achieving the overall goal of total system evolution.

Further, Campbell has offered that although The Big Computer could intervene with its sims' evolution by somehow informing them that their purpose is to lower their entropy, he has spoken at length that this approach ultimately backfires because the sims are involved in a *learning lab*, and must ultimately *choose* to lower their entropy via experience.

> After the individuated units of consciousness were created and there was a large number of those, they became interactive with each other and with the system. It was after this development that the system was challenged for the first time. Before, when it was a monolithic thing, there were no outside challenges—it was just itself, with one choice, one decision, and that was it. But after we emerged, suddenly it had to deal with other entities with free will.

---

[71] *National Geographic Society,* s.v. "Biodiversity," last modified October 19, 2023, accessed February 29, 2024, https://education.nationalgeographic.org/resource/biodiversity/.
[72] Greene, "The Termination Risks of Simulation Science" 25.

> The system's initial thought was to command all the individuated units of consciousness to get in line, listen, and follow its guidance. It would explain things, essentially instructing us on how to be. However, that approach didn't work out very well. The system tried to be more forceful, believing it knew the answers and could guide us by simply providing those answers. But it turns out that beings with free will need to discover answers for themselves, in their own way; they can't simply be told.
>
> So, only after we arrived did the system realize that each entity must be allowed to be who they are. The system learned that it needs to encourage growth and provide an environment where entities can make better choices. Beyond that, it must let them evolve at their own pace.[73]

In a sense, Campbell would likely argue the ultimate point of the simulation is *not* necessarily to realize that we are *in* one, but to realize that we *are* One—one evolving consciousness system; we are all one *semi-individuated* being trying to evolve itself together. Prosocial, humanitarian, "spiritual" evolution is the name of the game for Campbell. Simulation awareness, becoming Savvy, might be a big part of that process, but becoming Savvy is not the ultimate point. It might provide us with a very helpful boost though.

Now, since science has become the dominant worldview, *if* a scientific discovery of such magnitude turned out to encourage vast swaths of human beings along the line of actively evolving prosocially (*becoming* more "spiritual")—if it were discovered as an *absolute fact* in the same way F=MA is an absolute fact—then it would be a *massive net good* for the System and its sims. That would give us Greene's *overwhelming positive value* and then some! Not only would it shed light on where we are, it would offer more meaning to existence than mere random mutations procreating ad infinitum in an ever-expanding purposeless accident ("life is a disease of matter" as Goethe supposedly opined); it would reveal *why* reality and life exists at all! And, most importantly, that life and the universe actually has a purpose.

Further still, another overwhelming positive value of becoming Savvy includes diminishing the possibility of us destroying ourselves. After all, if we realize we are in a simulation, and that simulation exists to help us evolve "spiritually," it may very well convince people to change their ways.

> We are all individuated units of consciousness and we are part of the larger consciousness system. And our evolution, as we evolve individually, the larger consciousness system evolves, because we are part of it. So we decrease the entropy of our consciousness a little bit, then the whole system then decreases its entropy a little bit. So you see, we're the Larger Consciousness System's strategy, or one of its strategies for survival, for evolving, for growing, and it's all individual.[74]

That individual work of becoming a better person (lowering our individual entropy) naturally spills out into one's immediate social circumstance and, from there, the larger social order. If people were simply good and thoughtful to each other it would become a better world. This is the 'Golden Rule' in effect.

Indeed, in a consciousness-based simulation the overwhelming positive value that Greene demands is met in that, if we uncover that we are in a simulation, the potential is there for us to evolve more prosocially *faster* due to our new understanding that our simulation is a social system. In MBT there is no immediate negative result to becoming Savvy, in terms of the Big

---

[73] Tom Campbell, "MBT Volunteers Q & A June 2021 with Tom Campbell Pt 1/3," YouTube, June 22, 2021, [23:53], https://youtu.be/rbNniT9O7rc?t=1433

[74] Rick Archer, "Tom Campbell," *Buddha at the Gas Pump*, episode 266, transcript, November 16, 2014, https://batgap.com/tom-campbell-transcript/.

Conscious Computer shutting us down. Greene's required criteria is then met in full.

## FALSIFIABILITY: THE MORPHEUS PROBLEM

Admittedly, while simulism frameworks can be theoretically tested, there remains the problem of *falsifiability*, the ability to prove a theory is wrong. David Chalmers put it quite simply: "No amount of reasoning or observation could ever completely rule out the hypothesis that I am in a Matrix right now."[75] Campbell agrees that when it comes to being either in a real world or in a simulated world, "There would be no way for you to tell the difference. No experiment that you could do that would differentiate one from the other."[76] Greene concurs, "Demonstrating that we do not live in a simulation does not seem to be an attainable goal of experimental observation."[77] Simply put, we cannot disprove we are in a simulation because *even if we did*, our hypothetical proof could just be part of the simulation itself.

The situation harkens back to Morpheus in *The Matrix.* Though the sentiment has been communicated throughout history, Laurence Fishburne's character Morpheus put it so succinctly that it has been often quoted in books, academia, and popular culture ever since: "Have you ever had a dream, Neo, that you were so sure was real. What if you were unable to wake from that dream? How would you know the difference between the dream world and the real world?"[78]

Descartes' Dream is a philosophical conjecture that elaborates this problem.[79] In essence, we do not typically *notice* we are in fact dreaming; we take, say suddenly being a butterfly, as in the case of the famous dream of Chinese philosopher Chuang Tzu, to be a perfectly reasonable condition to be in while we are dreaming. *Was I, Chuang Tzu, dreaming I was a butterfly; or am I really a butterfly dreaming that I am Chuang Tzu?*[80] We tend to believe dreams are real while we dream them; we believe the world that we are experiencing as we experience it, *is* the real world.

The Morpheus Problem is, so far, no one has been able to come up with a of way of falsifying simulation theory or similar conjectures—our waking life might be a more constrained form of *lucid dreaming*, and that nothing of this life we are living is truly unfolding in the way that our common sense knowledge and beliefs about it assume. We can not disprove we are in a simulation because even if we had evidence against it; *that too could be part of the simulation*—much like you cannot disprove this is a dream, as your evidence might just be part of the dream itself. This is something of a problem, but that does not mean that simulation theory is ultimately irrelevant.

---

[75] Josh Oreck, "The Hard Problem: The Science Behind the Fiction," *The Ultimate Matrix Collection*, disc 8, *Roots of The Matrix*, DVD (Warner Bros., 2004), [02:13].

[76] Tom Campbell, "Thomas Campbell - The Monroe Institute Lecture, with Spanish subtitles." YouTube video. January 3, 2011, [44:22], https://youtu.be/uhv-XCff4_I?si=P_aKEPsXIQjvLOpH "Now, if all your senses were somehow terminated and you were in that black void of point consciousness, and then I could stimulate your central nervous system just like it's being stimulated now. I could reproduce all those little electrical signals just at the right places in your central nervous system. What would you experience? You'd experience just with your experience now."

[77] Greene, "The Termination Risks of Simulation Science," 22.

[78] *The Matrix*, directed by Lana Wachoski and Lilly Wachowski, (Warner Bros., 1999.)

[79] Ben Springett, "Philosophy of Dreaming," *Internet Encyclopedia of Philosophy*, accessed March 9, 2024, https://iep.utm.edu/dreaming-philosophy/. "Descartes' dream argument began with the claim that dreams and waking life can have the same content. There is, Descartes alleges, a sufficient similarity between the two experiences for dreamers to be routinely deceived into believing that they are having waking experiences while we are actually asleep and dreaming."

[80] Burton Watson, *Zhuangzi: Basic Writings*, 3rd ed. (New York: Columbia University Press, 2003), 44.

Philosopher Karl Popper argued for the importance of falsifiability in science. However, Thomas Kuhn, another philosopher of science, famously argued against Karl Popper's push towards falsification as the final step in the process of scientific verification. As Dr. Alexander Bird observed, "Kuhn's account argues that resisting falsification is precisely what every disciplinary matrix in science does."[81] The point, according to Bird's interpretation of Kuhn's opinion, is to make something so rigorous that falsification becomes impossible or impractical. (For more information, see the Kuhn-Popper Debates.[82,83])

Similarly, Ashutosh Jogalekar in his 2014 article for *Scientific American,* 'Falsification and its Discontents,' points out, "The Nobel Prize winning Roald Hoffmann has argued in his recent book how falsification is almost irrelevant to many chemists whose main activity is to synthesize molecules. What hypothesis are you falsifying, exactly, when you are making a new drug to treat cancer or a new polymer to sense toxic environmental chemicals?" Falsification, Jogalekar continues, "is a good guideline but which cannot be taken at face value and applied with abandon to every scientific paradigm or field."[84] So while simulism might not be falsifiable, that does not mean it is effectively useless, nor unscientific.

Since simulation theory might very well be fundamentally unfalsifiable, the questions must shift and become about *value*:

(1) Is the model useful?
(2) Does it better explain previously observed phenomena?
(3) Does it make predictions?
(4) Does it solve any problems or paradoxes?
(5) Does it require less assumptions than our current models?
(6) Can it be potentially tested by experiment?

In terms of Campbell's form of simulism, MBT, the answer is yes across the board.[85] So, although simulism might not be falsifiable, it certainly does not mean that it is ultimately valueless.

## APOCALYPSE NOW!

'Apocalypse' is a word thrown around often without an appreciation of what it originally meant. The word comes from the Greek word *apokalyptein* "to uncover, disclose, reveal."[86] Every major

[81] Alexander Bird, "Thomas Kuhn," *The Stanford Encyclopedia of Philosophy* (Spring 2022 ed.), ed. Edward N. Zalta, accessed March 1, 2024, https://plato.stanford.edu/archives/spr2022/entries/thomas-kuhn/.
[82] Imre Lakatos and Alan Musgrave, *Criticism and the Growth of Knowledge*, in *Proceedings of the International Colloquium in the Philosophy of Science* 4 (Cambridge: Cambridge University Press, 1970), https://philpapers.org/rec/LAKCAT-7. "Two books have been particularly influential in contemporary philosophy of science: Karl R. Popper's *Logic of Scientific Discovery*, and Thomas S. Kuhn's *Structure of Scientific Revolutions*. Both agree upon the importance of revolutions in science, but differ about the role of criticism in science's revolutionary growth. This volume arose out of a symposium on Kuhn's work, with Popper in the chair, at an international colloquium held in London in 1965."
[83] Phillip Irving Mitchell, "Post-Kuhnian Models of the Practice of Science," *Modern Resources, Dallas Baptist University*, accessed March 1, 2024, https://www.dbu.edu/mitchell/modern-resources/philosophyscience2.html.
[84] Ashutosh Jogalekar, "Falsification and Its Discontents," *Scientific American*, January 24, 2014, https://blogs.scientificamerican.com/the-curious-wavefunction/falsification-and-its-discontents/.
[85] Tom Campbell, *The MBT Science Trilogy and Bonus Content*, Playlist, YouTube, April 10, 2021, https://www.youtube.com/playlist?list=PLf8bCCRJkXgyGFOJg8REkl2TAdBPLl6aK.
[86] *Online Etymology Dictionary*, s.v. "Apocalypse," accessed February 27th, 2024, https://www.etymonline.com/word/Apocalypse#etymonline_v_15471.

scientific and philosophical discovery necessarily results in an *apocalypse*—an unveiling or a disclosure regarding the nature of things. We are defined and then redefined ontologically by what we are able to *reveal* about this mysterious circumstance that we have found ourselves involved in.

Multidisciplinary artist and author Alan Moore has often remarked on his particular reading of apocalypse:

> When I'm thinking about 'apocalypse,' I'm thinking about something that happens in the human mind, and only happens in the human mind. The word 'apocalypse' means 'revelation.' That's all it means. Although, it has come to mean 'end of the world.' Now, I think that is possibly true, but you have to decide what you mean by 'world.' And I don't think that 'the end of the world' means the end of the planet or even of planetary life. I think 'the world' is something which we have constructed. We have made it out of our economies, and our philosophies, and our prejudices, and our bigotries. We've constructed this huge palace of ideas in which we have then been condemned to live.[87]

Further elsewhere:

> This is a revelation that ends your world. I mean, that might just be that it's a realization that casts your previous perception into doubt, that sort of or destroys it. But in either event, it would seem to be a necessary thing. So, that is the way that I tend to treat 'apocalypse' which relates to this stuff. It's a personal apocalypse really."[88]

Moore's interpretation of 'apocalypse' applies nicely to our discussion in this chapter. In one way or another if simulation theory turns out to be true, whether or not it ends the physical universe, it will certainly end the world *ontologically* as we know it. It would be the revelation of revelations. Finding out we are living in a Matrix would be an apocalypse just as it was for Neo.

Let's take a wide look at everything we have examined, starting with the basics: *What does a simulation theory really propose?* Ultimately, simulation theory proposes our universe is a *subsystem* and that there is a *supersystem* that ours is dependent upon. That is what simulation theories truly put forward. This may sound radical, more than one system of reality, but science plays with ideas like this all the time. In theoretical physics we have multiple dimensions that have never been experimentally verified, cosmic strings that can likely never be detected, and multiple universes that might be closer than we think. These notions would be truly "exotic" if they were not entirely commonplace in cosmological theories.

Everything we have read about the ancestor simulation hypothesis suggests that there may be some kind of threshold problem for the sims inside: either we use too many computational resources, we learn too much about our own existence, or we (or our posthuman overlords) deviate in some kind of other unknowable way—and each of these scenarios ends up potentially destroying our universe. The key to all of this though is that it implies there is some physical and/or metaphysical limit on what human beings should or should not do.

Campbell's consciousness-based simulation also includes an upper limit for its sims in relation to their ability to lower their entropy; however, this threshold is much further away than the about 50% chance involved in an ancestor simulation. Furthermore, given that the System's

---

[87] TheBioskopPresents, "alan apocalypse," YouTube video, October 14, 2010, [01:20], https://youtu.be/cBc71ROdGxU?si=EvkjOGh0q7Zp20lC,

[88] Uonpop, "Alan Moore at the Magus conference at the University of Northampton 2010 2/5," YouTube video, June 1, 2010, [08:54], https://youtu.be/ZqiQQ26NFo0?si=ClH4gbeI4UW1DUMU

evolution is tied to the progress of its sims, Campbell puts forward a simulated universe designed with our long term evolutionary profitability in mind. In other words, we have a scenario where our universe could easily come to an end (Bostrom) and another scenario where it could also come to an end, but is a far less likely and less profitable course of action (Campbell).

Regarding our social evolution, Campbell is an optimist, "We are evolving [...] but it's a slow process. But the process is working and it's actually been speeding up. We've made more progress in the last, what, 200 years or 300 years than we probably did in, you know, a thousand years before that. So it's working."[89]

Alternatively, if the fears of Greene and his ilk are valid, and testing the simulation theory *does* in fact result in the destruction of reality, then you, dear reader, now hold in your hands one of the *most* dangerous articles—if not *the* most dangerous article—ever written. The pleasure is all mine.

Although writers (who are generally speaking a perverse bunch) would *love* to make such a claim, or have such an honor conferred upon them—"*the author of the most dangerous piece of writing ever written*"—I am afraid this is almost certainly not the case.

Think it through now: *if* the doomsayers actually believe what they say, and personally I am not quite convinced that they do—with the exceptions of Greene and possibly Bostrom—then a precise enough textbook that also includes an outline of a particular experiment is enough to unmake the whole universe. How similar this is to Lovecraft's weird tales—a forbidden grimoire that details an occult ritual that will destroy reality as we know it:

*In the accursed depths of that most damnable eldritch tome, whose very essence beckoned our universe towards a maddening simplicity, waited with silent animosity the outlines of a singular demonstration that any man of sound enough reason could all too easily perform, which upon completion would tear our cosmos apart atom by atom with the same ease that execution equines dismember a condemned prisoner.*

Could the universe really be such a gentle pushover?

I doubt it.

To me, conjectures surrounding the ancestor simulation hypothesis often makes the assumed posthumans sound unbelievably stupid, because if they are by definition a super advanced species, they should know that we would eventually figure it out. I argued elsewhere, "[I]f you are evolving intelligent life in a simulated environment, you must expect its simulated nature to be eventually discovered by its inhabitants as a logical consequence of your intelligent life forms' evolution."[90] In other words, if you are a posthuman and you have sims that are allowed to develop their intelligence through observation, science, and philosophy and you give these clever sims enough time to evolve to a state of development similar to the one that we are in right now—that is, on the cusp of executing such an experiment—then you *must* realize that sooner rather than later, your sims will *always* figure it out.

Now, if this is some kind of problem for you as a posthuman, then you would obviously anticipate that issue, and either occasionally modify your simulated universe so that your sims never figure it out, *or* you would program in some kind of intellectual glass ceiling so that your sims never think such a thought, *or* you would add 'gremlins' to their experiments so that they never properly work. These are simple patch-fixes. As a super intelligent posthuman, you would obviously be smart enough to anticipate all of this *before* you even started your simulation in the first place.

---

[89] "Tom Campbell: Talks with News for the Heart on Virtual Realities." (March 31, 2019 ed.) [43:57]
[90] Edge, "Why It Matters That You Realize You're in a Computer Simulation."

This is why the doomsday scenarios for the ancestor simulation hypothesis either do not make sense, assume our posthuman overlords are complete idiots (quite the irony indeed), or ultimately it does not matter—because, hey, if our universe were to be deleted you'd never know it anyway and it would not really matter in the long run. *And* in the ancestor simulation scenario (as we have seen) you were more likely than not heading in that direction already and will probably eventually reach it.

To me, simulists who push annihilation scenarios are like folks who feared a boat could sail off the end of the earth. Rather than embrace a new horizon of discovery, they warn us not to go too far. They assume this universe is intensely unstable—despite all physical evidence to the exact contrary—and that "a monkey with a brain full of mush trying to sort out what's right in front of [it]"[91] could by technology or discovery make this whole reality less than nothing in a mere picosecond.

It is a hubristic fantasy. It is the whole Lovecraftian forbidden knowledge situation grafted onto the simulation paradigm. That's *all* it is.

The major problem with everything we've discussed about the destruction of our universe is *not* the simulation idea itself, but all the hypothesizing about *the supposed posthumans involved*. How can we even begin to conjure with the motives and drives of a super advanced 'alien' civilization like that? Why are we even attempting to? How could you come up with anything other than conjecture and speculation? The answer is: you can't and you couldn't. That is what every simulism case that includes a posthuman factor must also include; and every case including a posthuman factor is ultimately unknowable. You might as well be playing with the assumed motives of demons. Sure, you could do that, but you would never really know?

However, if posthumans are *removed* from the simulism paradigm, as they are in Campbell's consciousness-based simulation (where the computer and the sims are linked in a symbiotic process of evolution), then you have a far more straightforward scenario, one that is stable, that explains the origin of the universe, offers predictions, is potentially testable, logical, and is *not* prone to endless teleological conjecture. The universe is simply a tool that consciousness has developed for itself to continue its evolution. It's a far less scary, less cold, less intellectually unstable, and ultimately less dangerous reality scenario.

Speculation on how and why our simulated universe exists are more cut and dry with Campbell than they are with Bostrom's posthumans.

---

[91] Terence McKenna, "Taxonomy of Illusion," transcript of talk featuring Ralph Abraham, University of California Santa Cruz, Santa Cruz, CA, 1993, https://www.asktmk.com/talks/Taxonomy+of+Illusion

| CONDITIONS | ANCESTOR SIMULATION | CONSCIOUSNESS-BASED SIMULATION |
|---|---|---|
| Source | Basement Reality | Fundamental Consciousness |
| "Programmers" | "Posthuman" and/or "X" simulators | No external "simulators" required |
| Purpose | Ancestor Simulation(s) | Evolutionary Trainer |
| Nested Simulations | Nested simulations possible, but potentially dangerous | Nested simulations possible, but non-dangerous |
| Physics | Classical and/or QM psychics-based computing assumed in 'Basement Reality' (ultimately unknown) | Computing processes not limited to classical or QM physics; 'nonphysical'[92] |
| Sims status | Possibly trivial sims | Sims ultimately non-trivial |
| Shutdown | Shutdown potential (≈ 50%) | Shutdowns possible, but largely impractical |
| Simulation Awareness | Unknown outcome for becoming Savvy; potentially dangerous | Neutral-to-strong potential reward for becoming Savvy |
| Relationship between realities | Existential divide between sims and simulators | Symbiotic connection between sims and The Big Conscious Computer |
| Stability | Multifactorial existential precarity (dangerously unstable) | System overall highly stable; evolution and survival encouraging |
| Resource-load | Almost certainly less finite resources | Almost certainly more finite resources |
| Simulation typologies | Multiple scenario typologies | Scenario typologies largely limited to evolutionary value |

*Table #2: Model differences in Ancestor Simulation Hypothesis and Consciousness-based Simulation theory.*

Human beings are a species marked by our imagination and hubris. The idea that we could destroy our planet is one thing, but that we could destroy the entire universe is an

[92] Edward Fredkin, "A New Cosmogony" *Proceedings of the Physics of Computation* Workshop, (October 2-4, 1992.) http://www.ai.mit.edu/projects/im/ftp/poc/fredkin/New-Cosmogony "There is no need for a space with three dimensions; computation can do just fine in spaces of any number of dimensions! The space does not have to be locally connected like our world is. Computation does not require conservation laws or symmetries. A world that supports computation does not have to have time as we know it, there is no need for beginnings and endings. Computation is compatible with worlds where something can come from nothing, where resources are finite, infinite or variable. It is clear that computation can exist in almost every kind of world that we can imagine, except for worlds that are sterile or static at every level."

outstanding belief. It is so hugely assumptive, it is amazing that it was actually batted around seriously at all.

It is not my intention that you will walk away from this chapter with a new form of overwhelming paranoia over the end of the universe. If you think about simulism carefully, you come to these conclusions: our universe is demonstrably stable, the likelihood of it being ended in a flash is so small it shouldn't even be entertained, intelligent sims should always figure out that they are in a simulation and our posthuman overlords (if there even are any) would anticipate all of this *a priori*. We probably have nothing to worry about when it comes to the destruction of the universe. We should be more concerned about what we do here on our planet to make our lives better.

If finding out we live in a purposeful universe, rather than an arbitrary/accidental one, encourages this effort, then that's a massive benefit.

Lovecraft's protagonists always collapse against the overwhelming knowledge of their own insignificance in the face of the grandeur and monstrosity of an uncaring cosmos. Our simulism doomsayers present a similar reading: our world is an illusion created by potentially indifferent, elder, and superior "alien" beings that reside in another dimension. It lends itself to a cold reading of existence. However, in a consciousness-based simulation, our universe is indeed an illusion, but this illusion is not a trap or a prison, but a functional tool with a purpose: the ongoing prosocial ("spiritual") evolution of the sims engaging with it. One is a gnostic-like vision of an amoral, callous (arguably "evil") world; the other, a strategy for the evolutionary betterment of all sentient life, including the 'Great Mind' behind the universe itself. It's an almost Buddhist vision where the goal of life is not merely one's own salvation, but "to bring a complete end to all the sufferings of others along with their own suffering."[93]

In a consciousness-based simulation, the discovery of the world as a virtual reality is potentially inconsequential. The only factors of interest under Campbell are: *does the understanding that we live in a simulated environment come with it an understanding of the grander evolutionary purpose of the simulation (the game) itself?* Or*, does this understanding have a non-effect?* Or*, does it result in anti-social consequences?* Given Campbell has been pushing for a wider understanding that our world is a simulated environment for over a decade, it stands to reason that at the minimum the author of the theory believes it will result in a wide prosocial net benefit.

My guess is that the overwhelming positive value is that we are *supposed* to figure this all out because if we have already gone this far at our current, arguably lowly state of intellectual development, *then it is more likely than not the case that becoming Savvy is not only anticipated, but expected.* Maybe existence is a kind of evolutionary intelligence test, and reaching this understanding, becoming Savvy is part of the process.

We should test simulation theory because the possibility is there that Campbell is correct, and that finding out we live in a simulation could encourage humanity to change course away from the selfish paradigm into a different paradigm where it is scientifically understood that, indeed, we are all One.

As Moore states:

> I can imagine a powerful idea—something that changes the way in which we thought about our world and ourselves—being powerful enough to actually destroy the structures of our old mindset

[93] Paul Williams, *Mahayana Buddhism: The Doctrinal Foundations*, 2nd ed. (New York: Routledge, 2009), 195.

> so that we could hopefully build a better mindset that was more appropriate to the times through which we live in. So yeah, we might have a little apocalypse down the line, but trust me, it will probably be a lot of fun.[94]

The universe will not be destroyed, nor will human beings go mad. Like all scientific revolutions, we will simply change and later remark on how limited our previous worldview truly was.

So, go on ahead, dear reader.

Push the button.

Have no fear.

[94] TheBioskopPresents, “alan apocalypse,” [02:08]

# IN-TEXT REFERENCES


- Agatonović, Miloš (2022). Computer Simulations as Fiction in Science. *Balkan Journal of Philosophy* 14 (1):5-10.
- Anthis, J. R. (2022). Consciousness semanticism: A precise eliminativist theory of consciousness. In V. Klimov, & D. Kelley (Eds.), Biologically inspired cognitive architectures 2021: *Proceedings of the 12th annual meeting of the BICA Society*, 1032 (pp. 20–41). Cham: Springer International Publishing.
- Bruiger, Dan J. (2023) A Refutation of the Simulation Argument.
- Franceschi, Paul (2008). The Simulation Argument and the Self-Indication Assumption.
- Franceschi, Paul (2014). The Simulation Argument and the Reference Class Problem : a Dialectical Contextualism Analysis. *Philosophiques* 43 (2):371-389.
- Niikawa, T. (2020). A map of consciousness studies: Questions and approaches. *Frontiers in Psychology*, 11, 530152.
- Nunn, Chris. (Ed.) (2009). Defining consciousness. Special issue, *Journal of Consciousness Studies*, 16(5).


# SUGGESTED READING


- Fredkin, Edward (1992). Finite Nature. *Proceedings of the XXVIIth Rencontre de Moriond*.
- E. Fredkin, (1990). Digital Mechanics. *Physica D*, 254-270 North Holland