

Another thing in this universe that cannot be an illusion

Maaneli Derakhshani

Department of Mathematics, Rutgers University – New Brunswick, New Jersey

maanelid@yahoo.com

One of the central claims of Sam Harris's 2014 book, *Waking Up: A Guide to Spirituality Without Religion*, is that "Consciousness is the one thing in this universe that cannot be an illusion" (Harris 2014, 51-79). I can think of another: our ability to reason. I just cannot fathom how Harris, a proponent of reason, couldn't think of it! To see why we must first understand Harris's claim about consciousness.

The undoubtability of consciousness

What is "consciousness"? It's a notoriously slippery thing to define and there is no universally accepted definition among philosophers and scientists, but Harris adopts the one famously provided by philosopher Thomas Nagel, in his 1974 essay, "What Is It Like to Be a Bat?": "an organism is conscious 'if and only if there is something that it is like to *be* that organism – something that it is like *for* the organism'" (Harris 2014, 51). To flesh this out, Nagel says to imagine trading places with a bat and being left with an array of (perhaps indescribable) experiences in the form of sensations, perceptions, and feelings. That array of experiences, whatever it's like, is what consciousness is for a bat. In Harris's words, "Nagel's point is that whatever else consciousness may or may not entail in physical terms, the difference between it and unconsciousness is a matter of subjective experience. Either the lights are on, or they are not" (Harris 2014, 52).

To further illustrate the difference, Harris contrasts your experience of what you are with what our growing scientific picture of reality says that you are:

At this moment, you might be vividly aware of reading this book, but you are completely unaware of the electrochemical events occurring at each of the trillions of synapses in your brain. However much you may know about physics, chemistry, and biology, you live elsewhere. As a matter of your experience, you are not a body of atoms, molecules, and cells; you are consciousness and its ever-changing contents, passing through various stages of wakefulness and sleep, from cradle to grave. (Harris 2014, 52)

And

Consciousness—the sheer fact that this universe is illuminated by sentience – is precisely what unconsciousness is not. And I believe that no description of unconscious complexity will fully account for it. To simply assert that consciousness arose at some point in the evolution of life, and that it results from a specific arrangement of neurons firing in concert

within an individual brain, doesn't give us any inkling of *how* it could emerge from unconscious processes, even in principle. (Harris 2014, 56)

So consciousness is a subjective-experiential phenomenon, and as such it cannot be completely described in the material terms of our scientific picture of reality—as nothing but atoms, molecules, and cells in the pattern of a body and brain over time—because what we know consciousness to be from the 'inside', from *having* it, is clearly not that.

What could it mean then for consciousness to be an "illusion"? Well, we normally regard something—say an object of visual perception—as an illusion if your perception of it is somehow a misinterpretation of its actual nature; if the way it *seems* to be to you is different from the way it actually is. So, in the case of consciousness, we might say it is an illusion if your perception of it is a misinterpretation of its actual nature; if the way that consciousness *seems* to be to you is different from what it actually is. For example, maybe your consciousness really just *is* atoms, molecules, and cells in the pattern of your body and brain over time, and it only *seems* to you that there is 'something that it is like to be your brain and body'.

But hold the phone: How can consciousness "seem" to be like something that it's not, without the "seeming" being an instance of consciousness (of subjective experience) itself? An advocate for illusionism about consciousness might respond that the "seeming" is an illusion, too. But then this leads to an infinite regress—he would have to say that 'it only *seems* to seem to you that consciousness is this subjective-experiential thing that it's like to *be* your body and brain', which raises the original question again and again ad infinitum. As Harris correctly says,

To say that consciousness may only *seem* to exist, from the inside, is to admit its existence in full—for if things seem any way at all, *that* is consciousness. Even if I happen to be a brain in a vat at this moment—and all my memories are false, and all my perceptions are of a world that does not exist—the fact that I am *having an experience* is indisputable (to me, at least). This is all that is required for me (or any other sentient being) to fully establish the reality of consciousness. (Harris 2014, 53-54)

Harris then finishes this paragraph with the claim that we opened with: "Consciousness is the one thing in this universe that cannot be an illusion".

The undoubtability of reason too

To say that consciousness is "the one thing" in this universe that cannot be an illusion is indistinguishable from saying that it is *the only thing* in this universe that cannot be an illusion. Which is to say that, except for consciousness, *everything else* in this universe could be an illusion.

Balderdash. In addition to consciousness, humans (and presumably some other animals) have a faculty of *reason*. We all have some intuitive sense of what reason is, but what is it in general terms? Harris's preferred philosopher for defining consciousness, Thomas Nagel, explains in his

1997 book, *The Last Word*: “The idea of reason...refers to nonlocal and nonrelative methods of justification—methods that distinguish universally legitimate from illegitimate inferences and that aim at reaching the truth in a nonrelative sense” (Nagel 1997, 5).

Reason cannot be an illusion, either. To see why, let’s consider a prime example of reason: rules of logic.

A basic rule of logic that we can all grasp is: “If P then Q” plus “P” implies “Q”, where “P” and “Q” are one among an infinitude of possible statements about the world. We also grasp that this rule of logic—fancily called “Modus Ponens” by philosophers—cannot fail to be valid under any circumstance. By “valid” I mean that it is impossible for the premises—“If P then Q” and “P”—to be true and the conclusion—“Q”—to nevertheless be false; it is impossible that “If P then Q” plus “P” implies “not-Q”. So we grasp that the rule of logic is valid always, everywhere, and for everyone—it is *universally* valid. At least, as long as we understand the meanings of the words “plus”, “if”, “then”, and “implies” in their normal senses.

Imagine now that some radical skeptic comes along and argues,

Maybe the rule of logic—“If P then Q” plus “P” implies “Q”—only *seems* universally valid to you, but it really isn’t. It might only seem that way to you because something you are unaware of—an evil demon scrambling your brains, or invisible aliens beaming thoughts into your head, or God divinely planting beliefs into your mind, or some genetic mutation in one of your ancestors millions of years ago—is deluding you into thinking it’s universally valid. You can’t rule out this possibility, so you can’t be sure that that rule of logic, which seems universally valid to you, really is universally valid.

This is tantamount to suggesting that the universal validity of the rule of logic might be an illusion. But this makes no sense, for a couple of reasons.

First, the argument of the radical skeptic gives us no *positive understanding* of how the rule of logic, whose invalidity we cannot imagine, might after all fail to be valid. When we look back at the rule of logic, after listening to the skeptic’s argument, we still cannot help but think that it is universally valid. This is unlike, say, the skeptical hypothesis that you are a brain-in-a-vat and an evil scientist is manipulating your brain to create false beliefs, such as believing that you have hands—here you *do* get a positive understanding of how your belief that you have hands, which is hard for you to imagine the falsity of, could be false after all.

Second, the radical skeptic’s argument actually has the form of the rule of logic that it is trying to cast doubt upon. In other words it *presupposes* the universal validity of the rule of logic. The argument asserts, “If you can’t rule out the possibility that something you are unaware of is deluding you into thinking that the rule of logic is universally valid, then you can’t be sure that the rule of logic, which seems universally valid to you, really is universally valid”. Call this “If P then Q”. It also says that “You can’t rule out the possibility that something you are unaware of is

deluding you into thinking that the rule of logic is universally valid". Call this "P". So the argument has the form, "If P then Q" plus "P" implies "Q". If the radical skeptic's argument were valid, it would apply to itself and therefore refute itself. And if the argument is not valid, then it doesn't establish the possibility that the rule of logic is an illusion.

So the universal validity of the rule of logic, "If P then Q" plus "P" implies "Q", cannot be an illusion.

Of course, reason is not restricted to a single rule of logic. Reason is a collection of rules of inference that have universal validity; or as Nagel puts it, "nonlocal and nonrelative methods of justification". Is there a more general way to see, independent of any particular rule of logic, that reason cannot be an illusion?

In *The Last Word*, Nagel provides exactly such an argument. He asks us to first consider an argument from a radical skeptic:

If my brains are being scrambled [by an evil demon], I can't rely on *any* of my thoughts, including basic logical thoughts whose invalidity is so inconceivable to me that they seem to rule out anything, including scrambled brains, which would imply their invalidity—for the reply would always be, 'Maybe that's just your scrambled brains talking.' Therefore I can't safely accord objective validity to *any* hierarchy among my thoughts. (Nagel 1997, 62)

Of this Nagel says,

But it is not possible to argue this way, because it is an instance of the sort of argument it purports to undermine. The argument proposes a possibility, purports to show that it cannot be ruled out, and draws conclusions from this. To do these things is to rely on judgments of what is and is not conceivable. There just isn't room for skepticism about basic logic, because there is no place to stand where we can formulate or think it without immediately contradicting ourselves by relying on it. (Nagel 1997, 62)

Although Nagel couches the radical skeptic's argument in terms of a specific skeptical hypothesis—an evil demon scrambling his brains—it is easy to see that Nagel's response doesn't depend on that specific skeptical hypothesis. Just replace it with some unspecified skeptical hypothesis, X (which could be invisible aliens, God, natural selection, and so on). Then the radical skeptic's argument generalizes to,

If X, I can't rely on *any* of my thoughts, including basic logical thoughts whose invalidity is so inconceivable to me that they seem to rule out anything, including X, which would imply their invalidity—for the reply would always be, 'Maybe that's just X talking.' Therefore I can't safely accord objective validity to *any* hierarchy among my thoughts.

And Nagel's rebuttal stays exactly the same.

So reason cannot be an illusion, even in principle.

What say you, Dr. Harris?

Perhaps, though, Harris views reason as *derivative* from consciousness? He never says this in *Waking Up* (nor anywhere else, as far as I'm aware). On the contrary, one passage in the book indicates that he sees reason as fundamentally distinct from consciousness, and non-mysterious in a way that consciousness is not:

We know, of course, that human *minds* are the product of human brains. There is simply no question that your ability to decode and understand this sentence depends upon neurophysiological events taking place inside your head at this moment. But most of this mental work occurs entirely in the dark, and it is a mystery why any part of the process should be attended by consciousness. Nothing about a brain, when surveyed as a physical system, suggests that it is a locus of experience. (Harris 2014, 55-56)

Reason is implied in "your ability to decode and understand this sentence", because your ability to decode and understand a sentence depends on using reason (making logical inferences). And, as Harris says, he thinks that this depends on neurophysiological events happening in your head, where "most of this mental work occurs entirely in the dark". That is, *without* the accompaniment of conscious experience.

Harris: Blind as Nagel's bat!

There is no doubt that your capacity to reason depends, at least in part, upon the neurophysiological events taking place inside your head. If taken too far, however, this can amount to claiming that reason is an illusion. That is the case if what Harris really means is that

- (a) Human minds are *entirely* the product of human brains (and other relevant physical events, such as light bouncing off objects and impinging on your retina);
- (b) Your ability to decode and understand this sentence "depends" upon neurophysiological events taking place inside your head, because your ability to decode and understand this sentence is *nothing but* the neurophysiological events taking place inside your head (along with other relevant physical events).

It's not completely clear that Harris means these things—he never says precisely what he means by "product of human brains" and "depends upon neurophysiological events". But (a) and (b) do seem to be the most straightforward interpretations of his comments. He gives no indication that

he thinks human minds are produced (even in part) by anything other than human brains, nor any indication that he thinks that your ability to decode and understand a sentence depends (in part) on anything other than neurophysiological events in your head. So let's see why (a) and (b) amount to saying that reason is an illusion.

If (a) and (b) are true then reason's universal validity *derives* from a purely physical process of biological evolution over millions of years. That's because human brains, and the neurophysiological events occurring in them, gradually evolved into their current form over millions of years, through a sequence of genetic mutations and environmental natural selection.

But this idea doesn't work. Nagel explains it best in his latest book, *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False* (Nagel 2012). He gives a general argument against the hypothesis that reason's universal validity derives from it having evolutionary-biological survival value:

in a case of reasoning, if it is basic enough, the only thing to think is that I have grasped the truth directly. I cannot pull back from a logical inference and reconfirm it with the reflection that the reliability of my logical thought processes is consistent with the hypothesis that evolution has selected them for accuracy. That would drastically weaken the logical claim. (Nagel 2012, 80)

Why would it "drastically weaken" the logical claim? A rule of inference, on evolutionary theory, would (for example) have to genetically mutate into existence and then be naturally selected for because it conferred a survival advantage to an organism in its environment, not because it is *universally* valid. Natural selection only selects those mutations that confer survival advantages to organisms in their environment, nothing more. There is no reason, on evolutionary theory, why any mutation that has survival value for an organism in one environment would have survival value for that organism everywhere and at all times; nor why a mutation that has survival value for one organism should necessarily have survival value for another organism. Moreover, saying that the rule, "If P then Q" plus "P" implies "Q", is a valid form of inference *because* it confers survival advantage to an organism in its environment, is saying that the rule has local and relative validity, whereas the rule *itself* is a nonlocal and nonrelative (hence universal) statement: What it says doesn't depend on whether or not it has survival value to any organisms in any environment, nor does it depend on when it mutated into existence. So saying that a logical inference is valid *because* it confers survival advantage to an organism in its environment, is another way of saying that the logical inference is *not* universally valid; or that the universal validity that it *seems* to have, is an illusion. (The same conclusion follows if we try to regard reason as an accidental side effect of natural selection, or as a product of 'genetic drift'.) Of course, as we've seen, at least one kind of logical inference *is* universally valid—it *cannot* be any other way.

Nagel continues:

Furthermore, in the formulation of that [evolutionary] explanation...logical judgments of consistency and inconsistency have to occur without these [evolutionary] qualifications, as direct apprehensions of the truth. It is not possible to think, 'Reliance on my reason, including my reliance on *this very judgment*, is reasonable because it is consistent with its having an evolutionary explanation.' Therefore any evolutionary account of the place of reason presupposes reason's [universal] validity and cannot confirm it without circularity. (Nagel 2012, 80-81)

In other words, any attempt to justify reason's universal validity, purely in terms of some evolutionary account, runs into the same problem as attempts to doubt reason's universal validity that we saw from *The Last Word*: The attempt to doubt reason's universal validity (with a skeptical hypothesis) presupposes its independent validity, and likewise the attempt to say that reason's universal validity derives from a purely physical process of biological evolution also presupposes its independent validity.

The problem here is analogous to Harris's point that any attempt to say that consciousness only *seems* to be a matter of subjective experience, but in actuality is just atoms/molecules/cells in the shape of a brain and body over time, still *presupposes* consciousness: "To say that consciousness may only *seem* to exist, from the inside, is to admit its existence in full—for if things seem any way at all, *that* is consciousness" (Harris 2014, 53). So even though reason and consciousness are fundamentally distinct aspects of human minds, any attempt to explain what they are, in terms of some account external to them, is impossible. Any such attempt ends up presupposing them.

It follows then that the claims, "human *minds* are the product of human brains" and "your ability to decode and understand this sentence depends upon neurophysiological events taking place inside your head at this moment", are untenable if meant too strongly. And, as I have argued, the strong (hence problematic) meanings seem to be what Harris had in mind.

Waking up to Reason

The only way to make sense of Harris's claims would be to interpret "product of" and "depends upon" weakly enough to be compatible with the fact that reason cannot be an illusion, cannot be a derivative of anything else, and exists independently of any particular mind/brain/body. Then Harris is wrong to treat the capacity of human minds to reason as non-mysterious in a way that consciousness is not. When Harris (correctly) says, "Nothing about a brain, when surveyed as a physical system, suggests that it is a locus of experience" (Harris 2014, 56), he should also say, "Nothing about a brain, when surveyed as a physical system, suggests that it has a capacity to reason". Or as Nagel puts the mystery in *The Last Word*, "The problem then will be not how, if we engage in it, reason can be valid, but how, if it is universally valid, we can engage in it" (Nagel 1997, 75).

Whatever the answer to this question, it will clearly have a form that “accounts for our capacity to think these things [reason] in a way that presupposes their independent validity” (Nagel 1997, 75). Beyond that, it’s challenging to imagine what a satisfactory answer would look like. Since no description of unconscious and non-rational complexity is sufficient to account for our capacity to reason, that leaves few other options. It leaves either some theological explanation, or, as Nagel prefers (Nagel 1997, 127-143; Nagel 2012), some naturalistic explanation involving biological evolution (perhaps Darwinian or some teleological variant) plus a fundamental law of nature that whenever an organism of sufficiently high neurobiological complexity develops it ‘acquires’ a faculty of reason and consciousness.

My own preference is Nagel’s, but I cannot delve further into details here. Suffice to say these issues are extremely interesting, and surely also at the center-of-the-bull’s-eye of Harris’ interests.

How doesn’t he know all this?

You may get the impression from *Waking Up* and subsequent podcasts on the issue of consciousness (Harris 2016) that Harris simply isn’t aware of these philosophical issues arising from reason. However, in a 2015 podcast with Tim Ferris (Ferris 2015), Harris named Nagel’s *The Last Word* as one of the five books he recommends everyone should read, saying that it “champions rationality in a very compelling way” (Ferris 2015). Yet, to the best of my knowledge, he has never mentioned or analyzed the issues Nagel raises concerning reason. Maybe that’s because Harris doesn’t agree with Nagel’s take, but if so, it doesn’t come across in his unqualified recommendation of *The Last Word*.

The closest Harris has come to touching on these issues, so far as I’m aware, is in recapping his first discussion with Jordan Peterson (Harris 2017). There he says, “I have always said that the scientific worldview presupposes the validity of certain values—logical consistency (up to a point)” (Harris 2017), and subsequently denies Peterson’s claim that “all scientific truth claims can be judged on the basis of the single (Darwinian) criterion of whether the claimants survive long enough to breed” (Harris 2017). The combination of these two views implies rejection of the claim that “logical consistency (up to a point)” can be judged “on the basis of the single (Darwinian) criterion”; but Harris doesn’t say this explicitly, nor does he refer to any of Nagel’s arguments.

So it’s a puzzle why, in *Waking Up*, Harris overemphasizes the novelty of consciousness and downplays the novelty of human minds, insofar as human minds have a faculty of reason. And it remains puzzling why he has (apparently) never commented on the philosophical issues raised by reason, since the publication of *Waking Up*.

Wrapping up

Consciousness cannot be an illusion, but it is *not* “the one thing” in this universe that cannot be an illusion. Reason is another thing in this universe, distinct from consciousness, that cannot be an illusion, and this has profound implications for our understanding of how human minds emerge from the unconscious and non-rational complexity of brain function. It implies that it is impossible to explain the existence and development of reason—an aspect of human minds— purely in terms of the unconscious and non-rational complexity of brain function, and it is a mystery how we are able to engage in it at all. A mystery at least as significant as the mystery of how consciousness emerges from brain function. It is also a mystery why Harris doesn’t seem to recognize all this, given his level of familiarity with Nagel’s work. If he ever does, he will have to become more radically antireductionist about the mind-body relation than professed in *Waking Up* and elsewhere.¹

Bibliography

- [1] Ferris, Tim. 2015. “Sam Harris on Daily Routines, The Trolley Scenario, and 5 Books Everyone Should Read (#87).” July 8, 2015. *The Tim Ferris Show*. Podcast, MP3 audio, 1:00:42. <https://tim.blog/2015/07/08/sam-harris-on-daily-routines-the-trolley-scenario-and-5-bookseveryone-should-read/>
- [2] Harris, Sam. 2014. “The Mystery of Consciousness.” In *Waking Up: A Guide to Spirituality Without Religion*, 51-79. New York: Simon & Schuster.
- [3] Harris, Sam. 2016. “The Light of the Mind: A Conversation with David Chalmers.” April 16, 2016. *Making Sense*. Podcast, YouTube. 1:45:00. <https://www.samharris.org/podcasts/makingsense-episodes/the-light-of-the-mind>
- [4] Harris, Sam. 2017. “Speaking of ‘Truth’ with Jordan B. Peterson,” *Making Sense* (blog), January 23, 2017. <https://www.samharris.org/blog/speaking-of-truth-with-jordan-b-peterson>
- [5] Nagel, Thomas. 1997. *The Last Word*. New York: Oxford University Press.
- [6] Nagel, Thomas. 2012. *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False*. New York: Oxford University Press.

¹ I am grateful to Marc Andelman, Mike Laster, Dax Oliver, and Rick Sint for helpful feedback. Special thanks to Marc Andelman for excellent stylistic suggestions, most of which I have incorporated.

