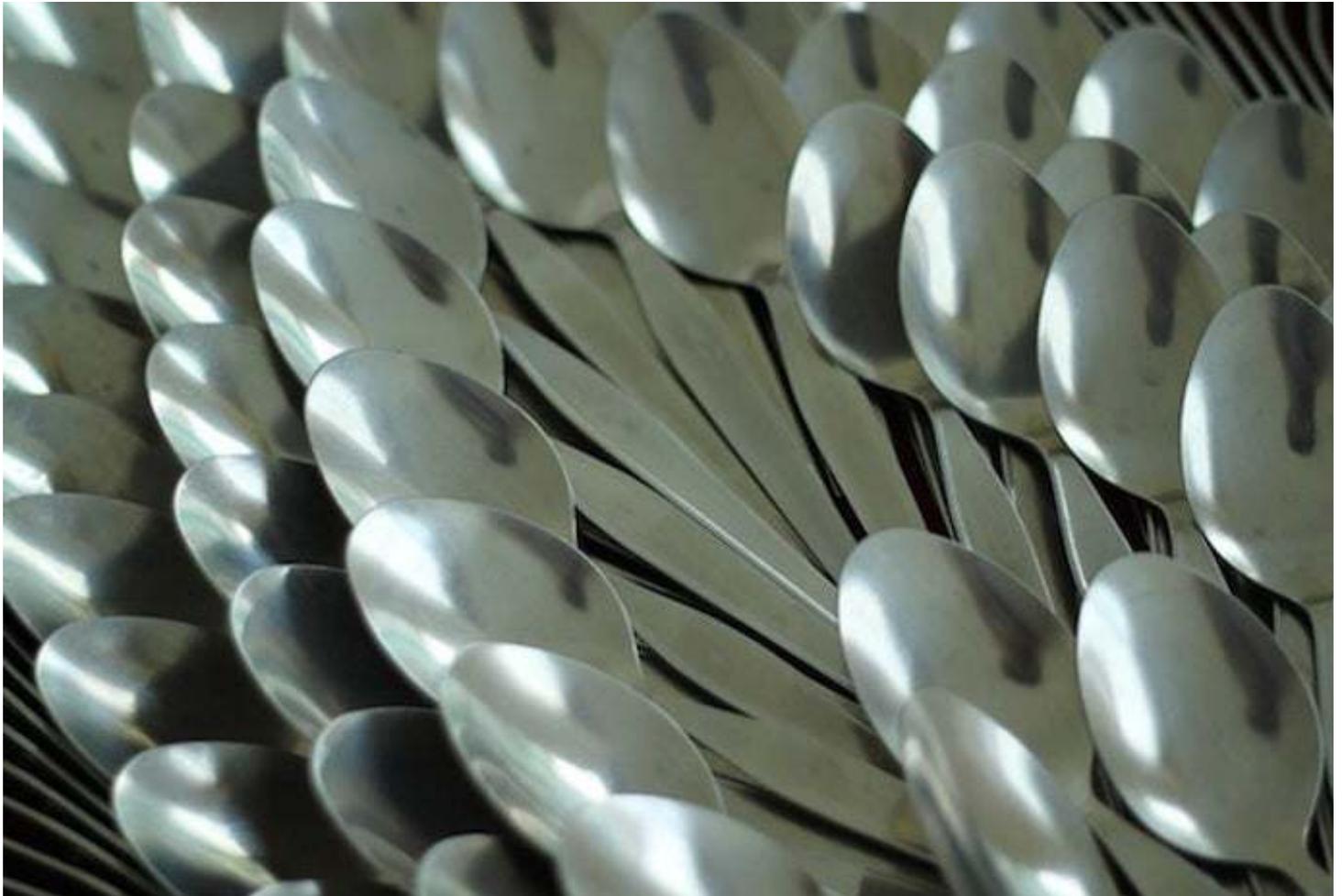


Sentient Robots, Conscious Spoons and Other Cheerful Follies

Bernardo Kastrup

How blind spots of critical thinking are distorting our collective intuitions of plausibility



Credit: [Kiattisak Promsensa Pixabay](#)

Contemporary science fiction seems obsessed with ideas such as downloading consciousness into silicon chips, sentient robots, conscious software and whatnot. Films like *Her* and *Ex_Machina* and recent episodes of series such as *Black Mirror* portray these ideas very matter-of-factly, desensitizing contemporary culture to their extraordinary implausibility.

The entertainment media takes its cue from the fact that research on artificial *intelligence*—an objectively measurable property that can unquestionably be engineered—is often conflated with artificial *consciousness*. The problem is that the presence of intelligence does not imply the presence of consciousness: whereas a computer may effectively emulate the information processing that occurs in a human brain, this does not mean that the calculations performed by the computer will be accompanied by *private inner experience*.

After all, the mere emulation of a phenomenon *isn't* the phenomenon: I can emulate the physiology of kidney function in all its excruciating molecular details in my desktop computer, but this won't make the computer urinate on my desk. Why, then, should the emulation of human information processing render a computer conscious?

When it comes to consciousness, even academics seem liable to lose touch with basic notions of plausibility. This is because, despite the prevailing assumption that consciousness is generated by arrangements of matter, we have no idea how to deduce the qualities of experience from physical parameters. There is nothing about mass, charge, spin or momentum that allows us to deduce how it feels to see red, to fall in love or to have a belly ache. Rationally, this abyssal explanatory gap should immediately lead us to question our prevailing assumptions about the nature of consciousness. Unfortunately, it has instead given license to a circus of arbitrary speculations about how to engineer, download and upload consciousness.

Already in the early 20th century, Bertrand Russell observed that science says nothing about the *intrinsic nature* of the physical world, but only about its *structure* and *behavior*. A contemporary of Russell's, Sir Arthur Eddington, also observed that the only physical entity we have intrinsic access to is our own nervous system, whose nature is clearly experiential. Might this not be the case for the rest of the physical world as well? Under this "panpsychist" hypothesis, the explanatory gap disappears: consciousness isn't generated by physical arrangements but, instead, is the intrinsic nature of the physical world. The latter, in turn, is merely the extrinsic appearance of conscious inner life.

Many panpsychists, however, go a step beyond this otherwise reasonable inference: they posit that consciousness must have the same fragmented structure that matter has. This way, individual subatomic particles are posited to be conscious subjects in their own merit, in that they allegedly have private inner experiences of their own. And because the body of more complex subjects—such as you and me—is made of subatomic particles, our conscious inner life is inferred to be constituted by a combination of the conscious perspectives of countless little subatomic subjects. Some interpretations of panpsychism imply even that inanimate aggregations of matter are conscious: a [recently published essay](#) claimed, "The idea that everything from spoons to stones are conscious is gaining academic credibility."

Oh well.

The notion that inanimate objects are subject to their own experience may sound absurd; and it is. However, the reason to dismiss it is not intuition—conditioned as the latter is by unexamined cultural assumptions—but simple logic. You see, the movement from "consciousness is the intrinsic nature of the physical world" to "subatomic particles are conscious" relies on a flawed logical bridge: it attributes to *that which experiences* a structure discernible only *in the experience itself*.

Allow me to elaborate.

The concept of subatomic particles is motivated by experiments whose outcomes are accessible to us only in the form of conscious perception. Even when delicate instrumentation is used, the output of this instrumentation is only available to us as perception. Those experiments show that the images on the screen of perception can be divided up into ever-smaller elements, until we reach a limit. At this limit, we find the smallest discernible constituents of the images, which are thus akin to pixels. As such, subatomic particles are the "pixels" *of experience*, not necessarily of the *experiencer*. The latter does not follow from the former.

Therefore, that living bodies are made of subatomic particles does not necessarily say

anything about the structure of the experienter: a body is itself an image on the screen of perception and so will necessarily be “pixelated” insofar as it is perceived. Such pixelation reflects the idiosyncrasies *of the screen of perception*, not necessarily the structure of the subject itself. As an analogy, the pixelated image of a person on a television screen reflects the idiosyncrasies of the television screen; it does not mean that the person herself is made up of pixels.



Sign up for *Scientific American's* free newsletters.

I thus submit that consciousness is indeed the intrinsic nature of the physical world, but subatomic particles and other inanimate objects are *not* conscious subjects. After all, [as Freya Matthews pointed out](#), the boundaries of inanimate objects are merely nominal—where does the river stop and the ocean begin? Whereas those of conscious subjects are unambiguously determined by, for instance, the range of the subjects' internal perceptions. So inanimate objects cannot be conscious subjects.

With inanimate objects excluded, only living organisms and the inanimate universe *as a whole* can be conscious subjects (a more extensive argument for this point can be found [here](#)). This way, as a living nervous system is the extrinsic appearance of an organism's inner experiences, so the inanimate universe as a whole is the extrinsic appearance of universal inner experiences. Circumstantially, [the inanimate universe at its largest scales has indeed been found to structurally resemble a nervous system](#). Under this view, there is nothing it feels like to be a spoon or a stone, for the same reason that there is nothing it feels like to be—at least as far as you can assess through introspection—one of your neurons in and of itself. There is only something it feels like to be your nervous system *as a whole*—that is, you. Analogously, there is only something it feels like to be the inanimate universe *as a whole*.

If biology is the extrinsic appearance of conscious subjects other than the inanimate universe itself, then the quest for artificial sentient entities boils down to abiogenesis: the artificial creation of biology from inanimate matter. If this quest succeeds, the result will again be *biology*, not computer emulations thereof. The differences between flipping microelectronic switches and metabolism are hard to overemphasize, so nature gives us no reason to believe that a collection of flipping switches should be what private conscious inner life looks like from the outside; let alone stones and spoons.

Emboldened by empirically ungrounded and illogical notions being noisily contemplated in some corners of academia, the entertainment media is rendering nonsense culturally plausible. As a result, a whole generation is growing up taking folly for future likelihood.

*Note: The ideas in this essay have been elaborated in further detail in the paper “[An Ontological Solution to the Mind-Body Problem](#),” published in *Philosophies*, Vol. 2, No. 2, Article 10.*

The views expressed are those of the author(s) and are not necessarily those of Scientific

American.

Bernardo Kastrup

Most recently the author of *The Idea of the World: A Multi-disciplinary Argument for the Mental Nature of Reality*, Bernardo Kastrup has a Ph.D. in philosophy (ontology, philosophy of mind) and another in computer engineering (reconfigurable computing, artificial intelligence). He has worked as a scientist in some of the world's foremost research laboratories, including the European Organization for Nuclear Research ([CERN](#)), and authored many academic papers and books on philosophy and science. For more information, freely downloadable papers, videos, etc., please visit www.bernardokastrup.com.