Does Your Brain Exist when Unperceived?

Alex Gomez-Marin • Instituto de Neurociencias de Alicante, Spain • agomezmarin/at/gmail.com

> Abstract • Not only does Hoffman claim that we do not see reality as it is, but that unperceived brains, trees and moons do not exist. His "interface theory of perception" is a peculiar blend of metaphorical ontology (objects are icons, space-time is a desktop) and mathematical modelling (the game-theoretical argument that fitness trumps truth). Conflating abstractions with concrete experience, evolution is used to refute everything (including evolution itself). Hoffman's sweeping iconoclasm then lands where it took off: addressing the problem of consciousness. After arguing against reality, he will tell us what it is.

Review of *The Case Against Reality: Why Evolution Hid the Truth from Our Eyes* by Donald Hoffman. W. W. Norton & Co., New York, 2019. ISBN 9780393254693 · 250 pages.

Introduction

« 1 » That our senses fool us is not breaking news (with the caveat that it is not appearances that are deceiving; what deceives is the mind). However, are we so fooled that nothing of what we see is "really real"? Donald Hoffman's answer as to how we experience reality is bold and simple: we do not. As I will discuss, he refers to a mind-independent reality. Not only does he argue that we do not see all of reality, but that we do not see reality *at all*.

« 2 » The reason, according to Hoffman, is found in evolution having hid the truth from our (and every organism's) eyes. And for a good purpose: survival – or, as he puts it, to harvest "fitness points" (29). Hoffman argues that fitness beats truth (61). Before we inspect his evolutionary argument against reality, let us concentrate on what he means by "real."

"Perception is not about truth, it is about having kids"

« 3 » Hoffman distinguishes two senses of real: "existing, and existing even when unperceived" (83). However, a clear unambiguous definition of what it means "to exist" is never provided (nor of "reality" or "truth," mentioned over 100 and 50 times respectively throughout the book). He claims that the world we see is not *objective* reality. Objective, here, means having an observer-independent existence. According to him, the moon only exists (not only in an epistemological but also in an ontological sense) when we look at it. When we look somewhere else, it ceases to be perceived, and to exist. We only need to look again for it to come back into existence. This applies to every single thing, including people, viruses and neurons. Eyes do not exist, Hoffman ensures us, when no one is looking at them.

« 4 » This is triply paradoxical. First, the organs of perception would not exist when unperceived. Second, it is not the brain that would fool us (as some neuroscientists claim), but evolution that fooled the brain. And, third, evolution itself, being the cornerstone of the argument against reality (what shall grant it its scientific status), would not be real either (as it involves matter, space, and time, all unreal, as Hoffman will try to persuade us). Is the argument designed to affect only physical entities rather than mathematical concepts?

« 5 » Berkeley's negation of material substance comes to mind: "I speak with regard to sensible things only; and of these I ask, whether by their real existence you mean a subsistence exterior to the mind and distinct from their being perceived?" (Berkeley 1713: 14). Hoffman feigns but does not quite endorse the philosopher's "subjective idealism." Instead, he calls himself a "conscious realist." For the psychologist, objective reality exists beyond doubt (spoiler alert). After his sweeping iconoclasm, he will show us what it is. Like Morpheus in the film, *The Matrix*, he offers us the red pill.

« 6 » However, will scientists and philosophers be willing to swallow it? Some readers may find Hoffman's case fascinating. After all, he is a diverting narrator. His prose skilfully combines catchy metaphors (objects are icons in the computer interface that is perception) and memorable examples (the Australian beetle that left the female for the bottle) with technical arguments (mathematically precise results that would prove his bold claims). Eager readers, no doubt, will be mind-blown. Yet who will scrutinize the logic, expose the philosophical commitments, and check the science references? Within a university professor's powerful rhetoric, science often becomes (paraphrasing Richard Feynman) the belief in the *imagination* of experts.

Handling Editor • Alexander Riegler • Vrije Universiteit Brussel, Belgium

« 7 » Versed in cognitive psychology, Hoffman uses game theory to leap into biology, neuroscience, theoretical physics, philosophy, cosmology and even spirituality. Based on numerical simulations of simple models (plus a theorem), and with virtually no empirical data, he purveys grand pronouncements about grand concepts such as truth, existence and the nature of reality.

« 8 » An augural inkling of the answer to the ultimate question of the meaning of life, the universe and everything is found (coincidentally) on page 42 of the book, when Hoffman shares with us his profound exchanges with Francis Crick. The Nobel Laureate, reacting to Hoffman's claims, replied: "Surely you believe the sun existed before there was anyone to perceive it. So why should neurons be different?" As we will see later (spoiler alert again), despite insisting that our perceptions cannot veridically describe "reality" at all, Hoffman will refuse to deny its existence. Betraying Berkeley's *esse est percipi (aut percipere)*, Hoffman thinks that a "true sun" would be a sun-in-itself (historically, the problem is rooted in Kant). A radical phenomenologist, however, would probably make the rejoinder: Yes, I do *believe* now in the sun's past existence, but for reasons that have no other basis than my present experience. Chronologically, the sun came first. Existentially, however, it is not the sun but experience that comes first. Always given in the immediacy of here-and-now, experience should have existential priority. Worries about the sun or the moon being "out there" should indeed be taken seriously, but so should a phenomenological analysis.

Despite insisting that our perceptions cannot veridically describe "reality" at all, Hoffman will refuse to deny its existence

Perception

« 9 » That tomatoes do not exist when unperceived sounds eerie, but less problematic than claiming so for brains. Hoffman's peremptory assertion is interesting insofar as it challenges the core tenets of physicalism. However, thwarting the whole of cognitive neuroscience, he tells us close to nothing about neurons. We are back to Berkeley's *Second Dialogue between Hylas and Philonous*:

> "The brain therefore you speak of, being a sensible thing, exists only in the mind. Now I would fain know whether you think it is reasonable to suppose that one idea or thing existing in the mind occasions all other ideas. And if you think so, pray how do you account for the origin of that primary idea or brain itself?" (Berkeley 1713: 52)

« 10 » As Henri Bergson (1904) remarked, to defend the view that the brain creates the world (with itself in it) leads to the self-contradictory proposition that the part is the whole or, alternatively, that the part subsists when the remainder of the whole vanishes.

« 11 » Hoffman and many others would agree that seeing is an active, constructive process (what is constructed and out of what it is constructed is often less clear). The difference is that, for Hoffman, such constructions do not exist until we construct them, and cease to exist when we cease to construct them. The Necker cube indeed makes his point (80), but how generalizable is this very example as to how we perceive every and any thing in the world?

Constructions do not exist until we construct them, and cease to exist when we cease to construct them

« 12 » As soon as one wishes to delve deeper into the actual details, caveats and subtleties of the proposal, Hoffman diverts us to analogy. This is welcome at first (as it lifts our imagination), but eventually annoying (as one needs to get back on the ground, to do the excavation work).

« 13 » The rationale goes like this: *Everything* we see is (like) an icon on a computer desktop interface. You do not frankly think that files on your desktop are genuinely white or rectangular, right? The pixels tell us *nothing* about the true nature of the file. "Its job […] is to *hide* that nature" (76). *Ergo*, the same applies for everything and anything you see.

« 14 » Apart from granting the epistemological (and also ontological) primacy of hardware over software, note the misleading traces of "smallism," or the belief that "the smaller, the realer." However, fluid dynamicists would not seek the essence of the Navier-Stokes equation in the quarks of the water molecules. The point is that quantum mechanics has virtually nothing to say about fluid dynamics. The meaning of the icons on your smartphone is not to be found in its transistors – one does not get poetry, pornography or philosophy from the Schrödinger equation.

« 15 » Is it not ironic that the current dominant metaphor of the brain as a computer is transfigured into the brain as a computer icon? Some argue that the odds that we may already be living in a computer simulation are tremendously high (Bostrom 2003). Attempts to explain physics in terms of computation depict a universe where apples would fall not due to gravity (or symmetry) reasons but because they are programmed to do so. In such a world there would be bits but no rabbits. For Hoffman, the cosmos is a grand simulacrum.

« 16 » To endow his metaphorical-philosophical position with scientific status, Hoffman labels it the "interface theory of perception" (ITP). Yet what, if anything, makes Hoffman's ITP a scientific theory? What empirically testable predictions has it made that are not retro-dictions or untestable predictions? When he says "ITP predicts this," or "that confirms the predictions of ITP," all he is showing is that certain (carefully picked) edgy interpretations of general relativity, evolutionary biology or cognitive science are consistent with his claims. Indeed, some strong points (but also their opposites) may be taken seriously by the respective disciplines.

« 17 » Hoffman seems as aware of the limitations of his case as he is unwilling to address them. Throughout the book he swiftly points to his own potential weak spots, only to immediately dodge them with one more captivating example or enchanting analogy. When referring to ITP, he confesses that such a "name is a bit rich for a mere metaphor, but I try in what follows to pay the promissory note" (76). Well, he does not – or, if he does, it is by borrowing from other newly issued promissory notes and charming metaphors.

Evolution

« 18 » To strengthen his psychological argument against reality, Hoffman draws upon evolution. The reader may dismiss ITP but, by summoning Charles Darwin, Hoffman hopes that scientists will take his position more seriously. Many would claim that we see (at least a part of) reality as it is. "But evolution disagrees" (20).

The current dominant metaphor of the brain as a computer is transfigured into the brain as a computer icon

« 19 » The mantra now is "Fitness-Beats-Truth" (FBT) (61). We are fooled but alive or, more precisely, we are alive precisely because we are fooled. Although we cannot apprehend anything about the transistors from the pixels on the screen, according to Hoffman, we would not need to know anything about them, anyway. The "what" of constructivist cognition is explained by the "why" of phylogenetic veridicality; "vision is no passive camera. It is an impatient hunter for fitness payoffs" (162).

« 20 » That fitness, if not veridical, at least aims at veridicality sounds reasonable. Not to Hoffman. His definition of veridicality has, nonetheless, some problems. It seems too stringent, at times a straw man, and it is also possibly tautological (accurate perceptions would, by definition, be untrue). Moreover, the conclusions he extracts from numerical simulations of very simple models are postulated to be valid for any real system (including space-time itself). Even if organisms tuned to the fitness function outcompete others, does this necessarily imply that the latter cannot survive? Some perceptions are just fit-enough. Claiming that "[p]erception is not about truth, it is about having kids" (77) oversimplifies the subtle relation between reproductive success and fitness. In a worrisome whiff of Dawkinism, Hoffman writes that "[i]t's all about struggles between genes" (25). Is the evolution of living organisms nothing more than a fitness video game?

« 21 » In a sleight of science, Hoffman surreptitiously swaps natural biological evolution for a simple model of two rabbits competing for water resources. The concrete is abstracted away, misplaced, and ultimately confounded with abstractions. Similarly, Jakob von Uexküll's (1909) concept of *Umwelt* (the meaningful world of each living organism), ends up relegated to a particular tuning of the user interface to different fitness payoffs across animals. In a word, subjectivity evaporates. Notwithstanding the great insights that mathematics can provide, when game theory replaces biology (only to declare life itself unreal and untrue), the math has become the territory.

« 22 » Are the selection pressures he talks about also unreal? What about matter and space-time, entailed by evolution itself? For Hoffman, they are not real either. Passages of the book, such as "the visual system *decodes messages* about fitness" (133, my italics), portray Hoffman's extreme computationalism, which percolates all the way down to space-time regarded as "a data-compressing and error-correcting code for fitness" (114). Risking self-refutation, Hoffman professes that evolution has made truth extinct (clearly with the exception of the truth of such a claim).

Risking self-refutation, Hoffman professes that evolution has made truth extinct

« 23 » This issue becomes more critical when we become aware that Hoffman regards his formalism as immune to his own case against reality. His FBT theorem "applies only to our perceptions of states of the world" (90), not to our perceptions of math and logic.

> "Fortunately, the FBT Theorem […] applies only to our perceptions of states of the world. Other cognitive capacities, such as our abilities with math and logic, must be studied on their own to see how they may be shaped by natural selection. It is too simplistic, and false, to argue that natural selection makes all of our cognitive faculties unreliable." (90)

Fortunately, indeed. The moon may disappear when we are not looking at it, but Hoffman's theory does not. Or, to put it comically, a man enters a bar, orders two beers, and gets number two (but no beer).

« 24 » Given the number of extraordinary claims, one cannot refrain from mentioning the relative scarcity of peer-reviewed scientific papers on the very corpus that Hoffman prolifically propounds. Except for a piece with simple model simulations, some published mathematical details, and a "theoretical review" spelling out ITP, his other papers referenced in the book are basically a rephrasing of the same mantras. In addition, his cherished mathematically precise theorem remains, to my knowledge, unpublished (except for a pdf posted on his own website; see note 18 in Chapter Four of the book). So, we got a sophisticated juggling of toy models, technological metaphors, and a bold mathematical proposition yet to undergo peer review. The burden of proof is still on the plaintiff.

The moon may disappear when we are not looking at it, but Hoffman's theory does not

Consciousness

« 25 » However, behind all negations, an affirmation lurks. We do not perceive reality as it is, and yet "[s]omething exists" (176), whether we observe it or not. These two words in the tenth chapter seem to undermine everything Hoffman has said in the previous nine. Independent of our perceptions inside *The Matrix*, there is an objective reality "out there."

« 26 » One may now be confused. Was Hoffman claiming that there is no objective reality, or that objective reality cannot be perceived? And, in any case, how can he know about it? According to Hoffman, our concepts of space, time, and causality are unfit for the job. Beneath the computer interface, therein lies the world of transistors that we cannot grasp. But perhaps Hoffman can.

« 27 » After taking science to disprove reality, Hoffman claims that science (and only science) may have access to it. At the edge of scientism, he writes:

> "I agree with Dawkins. If a system of thought, religious or otherwise, offers a claim that it wants taken seriously, then we should examine it with

our best method of inquiry – the scientific method. That's taking it seriously." (197)

Despite Hoffman's plea, it is often hard to tell when he is putting forth philosophy as science or science as philosophy.

« 28 » When referring to Plato, Hoffman proclaims his own attempt to "transform this stale philosophical chestnut into a crisp scientific claim" (54). In yet another immodest quote, we read: "Plato's famous allegory of the cave […] is a step in the right direction of ITP, but it does not go far enough" (86). Does Hoffman's allegory of the *case* claim to replace Plato's allegory of the *cave*?

« 29 » So, after his initial sweeping iconoclasm, it now turns out that (a) there is an objective reality, (b) it is graspable, by science, and science only, and (c) Hoffman is going to tell us what it is. "Can we hope to pierce that screen and see objective reality? Conscious realism says yes" (191). The thing-in-itself is not necessarily out of reach (Kant's "transcendental idealism" is forced into the lab or, more precisely, onto the whiteboard). So, what is that elusive underlying fundamental reality? It is consciousness itself.

« 30 » His take (like mine) is that consciousness is not an epiphenomenon of brain electrochemistry. Steadfast to find a solution to the "hard problem" (or should one pursue its dissolution, instead?), Hoffman dares to swim against the tide of physicalism. Matter cannot give rise to mind almost by definition (what is fundamental is fundamentally not conscious). Perhaps we can bridge the gap form the other side. However, Hoffman's is a view from nowhere, a detached objective attempt to account for the felt presence of immediate experience. Berkeley (and Goethe) would strongly disagree; the way to reality is *direct* perception, not abstract mathematics.

Hoffman's is a view from nowhere, a detached objective attempt to account for the felt presence of immediate experience

« 31 » Along the lines of the current renaissance of panpsychism, Hoffman claims that consciousness is fundamental. This is not a new position, but his "theory of conscious agents" (TCA) is a new stab at it. Attempting to provide the structure and dynamics of consciousness, it goes like this: "consciousness […] is properly described as a network of conscious agents" (198), whose dynamics may be "similar to the dynamics of cryptocurrencies" (195). In the same way that water is not simply H_2O, what progress would one make defining wetness in terms of wet elements? Despite their good uses, stochastic models are no substitute for Husserl's *epoché* or Heidegger's *dasein*.

« 32 » The idea that reality is made of elementary information-processing systems exchanging information is neither so radical nor novel. Hoffman's brand must distinguish itself from serious competition. He vehemently complains about others with regard to what he himself does not provide. With snide remarks about Giulio Tononi's "Integrated Information Theory" (Tononi 2004), Hoffman straw-mans it on every occasion he mentions it. Perhaps it is too much to say that consciousness is *identical* to Tononi's "Phi," but, as opposed to Tononi's, Hoffman's approach starts from *ad hoc* mathematics, rather than from experience itself. For Hoffman, perceptions are never true (but they are not illusions either: "an illusion is a perception that fails to guide adaptive behavior" 87). Thus, for him, proposing a theory of consciousness as fundamental reality comes with a price: we must relinquish concrete experience and embrace mathematical abstraction all the way down. Hoffman insists that "[a] science of consciousness no more requires divorce from living consciousness than meteorology requires naiveté about thunderstorms" (185). Yet this is the curse of any phenomenology (the science of the essence of consciousness in the first person) worth its salt.

We must relinquish concrete experience and embrace mathematical abstraction all the way down

« 33 » Periodically reminding us that science is about proposing precise theories to precisely prove them wrong, Hoffman will nevertheless choose to sin prudently. Namely, after granting that any Theory of Everything is an oxymoron – as it may explain everything except its own premises (for instance, that the whole Universe sprang into existence from nothing, at a single instant, and for no reason) –, he aspires to derive all desktop icons from the mathematics of transistors, aka, consciousness. He never tells us where mathematics comes from, though.

« 34 » Hoffman brings one more "ism" to the consciousness studies marketplace: "conscious realism" (184). Being a monism that challenges the assumptions of physicalism, he tries to distance his stance from both idealism and panpsychism. His "conscious realism," though, risks committing the mirror image of the basic mistake of physicalism. Current consciousness research seems poised at a crossroads, either we commit to the Sisyphean task of explaining consciousness in terms of everything else, or we explain everything else in terms of consciousness. Should we accept such a dilemma? It is mindless to try to eliminate mind, but is the problem solved by getting rid of matter? Mutilating dualism is not the way to eschew it.

Conclusion

« 35 » Hoffman is probably right in pointing out what is wrong (his critique of standard reductionist materialism) and wrong in pointing out what is right (his metaphorical ontology dressed as science). Ironically, his corrective to materialism falls prey to some of the same key flaws, in particular, presenting a philosophical commitment as scientific facts. Moreover, despite submitting a radical thesis, his is disappointingly conservative, drawing from popular science fiction images (computer simulations, minority reports, etc.) within the milieu of scientific "gurus." His main message that all our perceptions are, in a way, constructions is important and mostly uncontroversial. Going far, does he go far enough?

« 36 » Entertainingly devious, Hoffman's statements are mostly conjectural. As a reader, I felt enchanted, at first, subsequently confused, and ultimately frustrated. And yet, agreement is overrated. What Hoffman is trying to do is both noble and

necessary. The lure of his claims may attract a great deal of followers, mostly outside academia. However, given his level of ambition, one must be unforgiving. Claiming that science now shows that reality does not exist is astonishing, even seriously risible. It is one thing to blow someone else's mind, another to become a professional mindblower.

« 37 » Technological metaphors are making not only life and mind but also reality obsolete. The history of science would be hard to conceive without analogies. In physics, we say that rocks fall obeying laws as if they were responsible citizens. In neuroscience, many are convinced that the brain is a computer (although, initially, computers were like brains). In Hoffman's case, it is a figure-of-speech overdose: *icons, desktops, headsets, videogames, social networks, the twitter-verse, the kid in a candy store, the red pill, fitness points, cryptocurrencies,* etc. Hoffman exhorts us to take his metaphors seriously, rather than literally. May he use them concretely? Imagination is indeed important for science, until it starts turning science into fiction.

« 38 » Let us recap. Hoffman's double gambit consists of hanging two sets of arguments together: ITP and FBT telling us what reality is not, and TCA telling us what it is. The blue pill brings us back to ever-deceiving perceptual experiences; the red pill leads to abstract mathematic enlightenment. However, for all this to count as new scientific theories, one should forbear to disseminate folk metaphysics, and point to predictions that are genuinely novel and empirically testable. Serious work remains to be done.

« 39 » Hoffman's metaphorical tour de force literally captures the zeitgeist, within science and without. Perception is a walled window (with a plasma screen hanging on top). Reality is illusory, simulated or hallucinated. Nature is not a garden of mysteries anymore (nor a book of secrets to be forcefully extracted by Promethean scientists), but a bag of tricks discovered by illusionists and employed by mavericks. The world is a grand cheat, the cosmos a deep fake. Truth has gone extinct. Hungry for "likes," compelling is selected over correct. The virtual is the new real.

Nature is a bag of tricks discovered by illusionists and employed by mavericks

« 40 » Let us recall the words of the Spoon Boy in *The Matrix*: "Do not try to bend the spoon. That's impossible. Instead only try to realize the truth. There is no spoon. Then you'll see that it is not the spoon that bends, it is only yourself." Hoffman has bent the spoon (and broken it). There is a certain dogmatism in claiming, from one's own mind, that what is real is outside the mind. When presented with the forced choice between unpleasant truth and blissful ignorance, we can still pick the green pill.

References

Bergson H. (1904) Le paralogisme psychophysiologique. Revue de Métaphysique et de Morale 12(6): 895–908. English translation: Bergson H. (1920) Brain and thought: A philosophical illusion. In: Mind-energy: Lectures and essays. Translated by H. Wildon Carr. Henry Holt, New York: 231–255.

Berkeley G. (1713) Three dialogues between Hylas and Philonous. http://www.gutenberg.org/ebooks/4724

Bostrom N. (2003) Are you living in a computer simulation? Philosophical Quarterly 53(211): 243–255. ► **https://cepa.info/4957**

Tononi G. (2004) An information integration theory of consciousness. BMC Neuroscience 5(1): 42. https://bmcneurosci.biomedcentral.com/articles/10.1186/1471-2202-5-42

Uexküll J. von (1909) Umwelt und Innenwelt der Tiere [Umwelt and inner world of animals]. Julius Springer, Berlin.

Alex Gomez-Marin is a theoretical physicist turned behavioral neurobiologist. He is currently a *Ramón y Cajal Fellow* and the Principal Investigator of the Behavior of Organisms Laboratory at the Instituto de Neurociencias de Alicante, Spain. Combining computational ethology, real-world cognitive neuroscience, and process bio-philosophy, his research aims to establish shared principles of animal behavior across species. https://behavior-of-organisms.org Twitter: @behaviOrganisms.

Received: 9 September 2020
Accepted: 16 November 2020