Nicholas Humphrey
Sentience: The Invention of Consciousness
Oxford: Oxford University Press, 2022, 256 pp.
ISBN: 9780198858539

Reviewed by Alex Gomez-Marin
Instituto de Neurociencias (CSIC-UMH), Alicante, Spain
agomezmarin@gmail.com

Nicholas Humphrey is a fine thinker and a lucid writer. In his latest book, *Sentience: The Invention of Consciousness*, he expounds the latest and more complete version of his long quest to address the problem of consciousness. 'These are difficult questions. I have been asking them for fifty years', he avouches.

He kicks off with a brilliant prologue, followed by twenty-four very short chapters (some less than three pages!). Humphrey acknowledges our dual nature: 'I'm flesh and blood. I'm soul.' Right from the start he ponders how and why sensations have such an 'eerie immaterial quality'. His remarks are eventually poetic: phenomenal consciousness is 'an invention so sublime that, if it were to cease to exist, it would indeed diminish the whole of creation'. He attempts to make the mysterious, the uncanny, the ghostly, less so. He believes that 'the hard problem of consciousness becomes just an ordinary scientific problem'. It is just that, he argues, people haven't been trying to solve it appropriately.

Humphrey's autobiographical anecdotes are very relevant to grasping the plot, the argument, and their genesis. He recalls his early interests in parapsychology. As a student at Cambridge University in the 1960s, he joined 'the student branch of the Society for Psychical Research'. He went ghost-hunting and entertained claims about messages from the dead. As a psychologist, he 'wanted to understand why people are so ready to believe in things that probably aren't true'. To be an enemy of error is not quite the same as to be a lover of truth. A couple of cases were sufficient for the freshman to both causally dispel the evidence and casually dismiss any other. For Humphrey, 'the fact of the allegation was often of more interest than the fact of the facts'. He wondered 'how weird the human mind's conception of the mind is'. His stance on 'the edges' of consciousness percolated to consciousness itself throughout his career. He asks: 'Could the properties of phenomenal consciousness have arisen by some such happy accident, then to be taken up and embellished by natural

selection as something like a work of art — just because, on a surreal level, they enrich their bearers' lives?' An improvable discovery was waiting for him.

Early in the book Humphrey walks us through his remarkable liaison with Helen, a blind monkey who could see. As part of his PhD, he recorded from primate neurons to understand the role of the cerebral cortex in vision. In the lab there was a monkey called Helen with her visual cortex removed. And yet, surprisingly, Humphrey noticed, she could 'detect the spatial location of a visually salient object'. Could the monkey really see? How? And why? Together with his supervisor, Lawrence Weiskrantz, they published the landmark discovery in *Nature* in 1967. Weiskrantz found the same phenomenon in humans, where a patient who had the right visual cortex removed could not see when asked, but he could consistently guess the position, shape, and colour of objects. He called it 'blindsight'.

What is it be like to have blindsight? What did Helen lack? Did she find blindsight worse than no sight? Humphrey makes a puzzling remark: 'she didn't experience it as hers.' Helen could visually perceive without being conscious of visual sensations, an uncanny way of seeing, a kind of vision possible without any conscious sensation.

Humphrey alludes to Daniel Dennett's remarks on the 'danger of reading too much into' Helen's exceptional case. Isn't it ironic when philosophers prefer thought experiments over revealing empirical ones? Dennett, the philosopher who (believes he) explained consciousness (away), has inspired Humphrey in his philosophical quest. He indeed holds Dennett in high esteem. 'Dan is the cleverest person I know', he professes, while rejoicing in the American philosopher's 'yeses and buts' regarding his own idiosyncratic approach to the hard problem based on 'blind monkeys who could see' or 'gorillas who could read minds'.

Humphrey taught Helen to see. She was 'scheduled to be killed and her brain examined to confirm the extent of the surgery', but it was agreed Humphrey could 'go on working with her'. He confesses he 'didn't much like doing these experiments', and probably neither did she. 'It's not that I doubted their scientific value', Humphrey admits, but 'I was valuing my curiosity about how the monkey's brain works over the monkey's interest in enjoying the use of its brain'. What's the proper limit of scientific enquiry upon the creatures we submit to the violence of impediments? Humphrey seems to have written a brief consciousness love letter to Helen: 'There were many times when I tried to see things from Helen's point of view. One thing I took for

granted was that my research with her had, at some level, made her life more worth living.' He 'continually wondered what it is like to be her'.

Humphrey's extraordinary finding marked his career and intellectual path. The existence of blindsight is a burden and a blessing for a theory of sentience since, being 'a case of pure perception in the absence of sensation', one would have to disprove the hypothesis that 'percipient but insentient animals could manage perfectly well'. In other words, if vision without conscious sensation is possible, why did evolution indulge in upgrading visual zombies? Is it more surprising that sensation is absent in blindsight or that it is present in normal sight? 'If blindsight is seeing and nothingness, normal sight is *seeing and somethingness*. And surely it's this *something* that stands in need of explanation.'

This takes us back to *sentience* (the title of the book). Humphrey draws from Thomas Reid's seminal distinction between sensation and perception. For the experimental psychologist, a 'creature is sentient if but only if it consciously experiences qualia — by virtue of which it becomes like something to be itself'. Despite being a proud physicalist, Humphrey sometimes sounds like a provisional dualist: 'let's say for now that sensations are basically mental states.' In his view, perception is to text what sensations are to illustrations. One could say that the former makes the graphic novel of life possible, while the latter make it worth reading.

Later in the book, Humphrey will offer two peculiar examples of limit cases of the dissociation between sensation and perception: proprioception (perception without sensation) and orgasm (sensation without perception). The former is much like blindsight, the latter a case of pure what-it's-like-ness. Proprioception, according to Humphrey, is 'pure perceptual knowledge', namely, the position sense 'doesn't give you the feeling of being there in any deeper sense'. In contrast, orgasm is 'about your experience of the bodily events as such' rather than about the 'objective external circumstances'.

Again, if being percipient but insentient is in principle sufficient for survival, why then did nature take the luxury to invent consciousness? What's the role of phenomenal experience? Humphrey agrees that 'nothing says that phenomenal quality is *necessary*', but he argues that consciousness gives your mind 'cognitive workspace', a 'coherent self narrative' to make sense of where you are, your history, to explain yourself to yourself and to others, and to explain other people to

yourself. Sentience 'makes all the difference in the world: the difference between being me and not being me!'.

But what's its evolutionary relevance? If it is a window on your mental life as it is played out, can something of such privacy leak publicly? Humphrey thinks so: 'I'll declare that I think there has to be. This is an article of faith. But it's supported by one very powerful argument: the evolutionary one.' It is time 'to come up with a story that will get us from insentience to full-blown sentience', he announces. Humphrey tells us his tale of the how life went 'from nothingness to somethingness', 'how our ancestors got from there to here'. Indulging in allegory, he writes: 'natural selection has created a piece of magic and planted it in the brains of billions of sentient animals like ourselves.' Evolution invented sentience, and so will Humphrey attempt here.

Facts and arguments are indeed tied together with a 'rather simple evolutionary theory'. Humphrey uses the usual bits and tape of the evolutionist, at the risk of squandering biology's greatest theory as a *post-hoc* magic wand to invent just-so stories. When it comes to consciousness, biologists' 'free miracle' is evolution (physicists have the Big Bang). I don't dispute that '[p]henomenal consciousness has evolved by natural selection'. And yet natural selection does not explain the origin of life. Can it explain, with precision, the origin of conscious minds? Analogies of a pile of sand becoming a pile out of sand do not do much explanatory work here.

Evolution 'in time' is not enough. Humphrey needs some sort of emergence 'in space'. The brain is, no doubt, the place where the word (somehow, somewhere, sometimes) becomes flesh. The core formula combines clichéd biological evolutionism with vanilla neural emergentism. In other words, evolution is the illusionist's hat and the brain the hand that pulls the rabbit of consciousness out of it. The T-shirt reads: evolution plus the brain enacts the miracle over again.

Humphrey must save his physicalist's stance with a sleight of mind. This takes place in Chapters 10 and 12, which are crucial to the argument. One must read them carefully, watching the legerdemain at slow motion. He prepares the terrain with a series of rather elliptical definitions. It requires sedulous attention: 'Whatever the brain actually does to represent sensation, the vehicle is presumably some kind of nerve cell activity. There won't be, at this stage, anything that can't be described in physical terms. It's when you form the idea of the representandum that things begin to get tricky. You describe your sensory experience to yourself in terms of phenomenal qualities… that

seem to have no counterpart in physical reality.' So, if physical processes do not have these phenomenal properties, do sensations have them?

He assumes that 'sensations, with their phenomenal properties, are feelings generated by the brain'. But he then remarks that sensory feelings 'are not constrained by the laws of physics'. He continues: 'If these properties turn out to be non-physical or even para-physical, that's just what we might expect.' What? There is more: 'sensations come to be experienced as being… made of immaterial mind stuff: in short, phenomenal.' May the card-carrying physicalist have his crypto-dualistic cake and eat it?

Humphrey rejects the 'neural correlates of consciousness' (NCC) programme, 'as the waters have been muddied by what has seemed to many researchers to be a good proposal'. He denounces the idea of brain processes as 'actually having' the properties of the conscious experience: 'Rather than asking how phenomenal properties are represented in the brain, they are seeking some feature of the brain that actually has these very properties. They are looking for the neural correlates of consciousness, not the neural correlates of representing consciousness.'

As Humphrey tries to clarify his position, the reader can become progressively confused. He writes: 'there could be a [hard] problem here — "an explanatory gap" between the brain and the phenomenal properties of sensations. There *could be*. But only if we cleave to the idea that sensations must somehow be *identical* to brain states.' He explicates that '[w]hen you see red, there won't be any activity of the brain that *is* phenomenally red; there will only be some activity by the brain that *creates the idea* of phenomenal redness. Hence, what we should be looking for is not the neural correlates of consciousness but the neural correlates of *representing* consciousness.' At this point, Groucho Marx comes to mind (and to brain), as Humphrey speaks about the first part of the party of the first part; why should we quarrel about a thing like this?

Whether you drop the NCC and follow Humphrey on his quest for NCRC or not, his proposal appears to be yet another epicycle refraining from giving up on the ill-posed starting point of obdurate materialist attempts, whereby what is defined as absolutely devoid of consciousness can and must produce consciousness. It seems that Humphrey is saying that sensations and ideas are not features of brains, but rather of minds, albeit formed in brains. Some tend to think of phenomenal consciousness as 'a state of mind', others wish to

conceive it as ‘a state of matter’, as ‘an intrinsic property of brain activity’. Sensations ‘have acquired phenomenal properties as a secondary overlay’, they are ‘not illusory’ but ‘veridical properties’; they are ‘ideas that represent’ how you feel about what takes place in your sense organs. For Humphrey, brain activity is the vehicle for the representation.

He is determined to loop the loop, literally and metaphorically. Let us see how. Fast-forwarding, we go from reflexes to the emergence of ‘phenomenality’ all the way to sentience, as self-entangling recurrent feedback loops give birth to Humphrey’s creation: the ‘ipsundrum’. This quasi-unpronounceable name is ‘a self-generated conundrum, echoing the real impossible triangle, the Gregundrum’. Humphrey seems to be rediscovering dynamical systems theory applied to neural processes with the hope of getting consciousness out of a middle-class mathematical abstraction, an attractor metaphorically grasped and metaphysically erected for those mathematically unversed.

Humphrey used to think that phenomenal properties were ‘pure make-believe’, until quite recently. Recall those armchair experiments where one entertains the existence of a triangle that looks like a triangle, sounds like a triangle, and moves like a triangle but that (surprise, surprise) is not a triangle. What about a strong back pain that feels like a back pain? Illusionists of consciousness inhabit a land of nonsense sprinkled with unending analytical kinks and twists. Such intellectual juggling feels like contemplating a half-pregnant woman. Luckily, Humphrey has moved on. And yet, he can’t resist the temptation to paraphrase the beautiful ‘real impossible triangle’ into his signature philosophical artwork.

Later on, Humphrey admits that ‘I’ve said little so far about what exactly might be required at the level of nerve cells to generate the attractors that *are responsible for* phenomenal consciousness. I won’t pretend I’m ready to provide a detailed anatomical and neurophysiological model’ (my italics, emphasizing the pervasive use of ‘filler verbs’). Such is the hallmark of promissory materialism: we already know ‘that’, we just need to figure out ‘how’. Humphrey’s ‘ipsundrum’ is consciousness conundrum’s murmdrum’s humdrum.

But, of course, as announced earlier in the book, ‘we had better come up with a plausible story to explain how phenomenal consciousness could possibly be generated by a physically constituted brain and body. For, without a physicalist explanation, we’ll be dogged by the philosophical nay-sayers — and there are lots of them out there — who would prefer to believe that sentience has its origin outside of

physics and didn't actually evolve as a biological phenomenon.' If you don't subscribe to his philosophical commitment, your science risks being boxed into the nonsense camp with flat-earthers and anti-vaxxers as neighbours. As we say in Spanish, '¿susto o muerte?' ('trick or treat' doesn't quite work as a translation). And yet, one must find a third way out.

Let us pause and recap the key ingredients in Humphrey's recipe for 'the invention of consciousness': a philosophical distinction (sensation versus perception), a fascinating neuropsychological finding (blind-sight), some rather vague evolutionary arguments (evolution gave us brains, and brains gave us sentience), and a mathematical object inspired from an impossible triangle (the ipsundrum).

Once you have all that in place, 'natural selection has a whole new design space to explore'. In his view, there are 'public consequences that can be "seen" by natural selection'. 'If being conscious of feeling this way about the situation is to show up in behaviour, it must cause changes in your mental attitudes that dispose you to act in ways you wouldn't have done otherwise.' Humphrey insists that evolution 'was destined' to give us phenomenal consciousness and the enhancement of the self. He argues that natural selection could improve survival by creating sentience, thus turning social creatures into proper individuals. Mind-reading is important indeed. But thoughts cannot ground the self, nor can perception, Humphrey believes. Sentience is the key. He transmutes the Cartesian motto into 'you feel, therefore you are'. 'What Mary didn't know' natural selection didn't know either. They both learned something when they first saw red. And yet one still wonders about the ontological status of redness? Are phenomenal properties 'really real'? Is it in the brain that the sky is blue and the rain is wet?

What about other kinds of sentience? Should we still be talking about consciousness at the scale of barracudas, bees, or bacteria? Does Humphrey's account cover them? After discussing how and why phenomenal consciousness might have evolved, Humphrey discusses where it might exist. Here we 'have to make a leap of scientific faith', he admits. According to Humphrey, there aren't 'semi-sentient creatures', 'there must be a clear threshold between insentient and sentient animals'. Humphrey doesn't believe that sentience goes all the way down either.

To draw the line, he first asks who couldn't possibly be sentient and then he postulates more stringent criteria to decide who most probably is. His exclusion and inclusion criteria are based on the right kind of

brains and lives. Once more, weaving (or perhaps waving) neuroscience and evolutionary biology, the key seems to be having neural circuits that would make phenomenal consciousness possible and a lifestyle that would make it advantageous. Humphrey suggests 'that warm-bloodedness played a double role in the evolution of sentience: on the one hand, it brought about changes in lifestyle that made sentience an essential psychological asset; on the other hand, it prepared the brain to deliver it'. Warm-bloodedness reduces the costs for the brain, allows to bring a relatively constant environment with us, insulates our bodies, and also, Humphrey suggests, our minds.

What are *the limits of sentience*, then? Who makes the cut? Humphrey wonders about how to 'possibly tell from an animal's behaviour which side of the divide it is'. One can only guess unless one has a theory. But is Humphrey's theory really a theory? I do not wish to contrast here the sound polymorphism of the word 'theory' in relativity theory versus evolutionary theory (physics and biology are different businesses), but nowadays anyone's idea seems to arrogate the moniker 'theory', despite too often being hardly more than a glorified metaphor, or a metaphysical ideology cloaked in brain scans. Humphrey doesn't seem to offer testable predictions but rather a set of post-dictions via plausible criteria to try to make sense of things.

Convinced that he's offered enough arguments, Humphrey then seeks evidence of his proposal. He seems to argue by analogy (or homology), inspecting animal behaviour based on what we humans have as traits. At the risk of being confused about what is similar and what is different (which amounts to everything), Humphrey looks for behaviours that would be sufficient proof of sentience. He hopes to read the effects of phenomenal consciousness behaviourally, pinning them down with specific tests: to look for behaviours 'that are *enabled* by having a phenomenal self and *could* not occur without it; next, behaviours that are *promoted* by having a phenomenal self and *would* not occur without it; and then, behaviours that are *required* in order to maintain the phenomenal self and so would be *irrelevant* without it'.

He proposes that 'seeking sensation for sensation's sake is something we would expect to see only in animals that are sentient. So, its presence confirms sentience (even though its absence doesn't disprove it).' This may include animals, such as birds and mammals, 'gathering experience for the fun of it, or indulging in exploratory play'. What about music? As it turns out, non-human animals are 'attracted to music or anything equivalent'. Masturbation makes the cut too, via a

poetic erotic meditation: ‘Not to put too fine a point on it, is orgasm self-made body music?’

After adducing reasons why sentience is not the exclusive province of humans, Humphrey draws a quick species argument: if any of these creatures is sentient, then most, but not all, are. Birds and mammals must be, but he is ‘a lone voice arguing that lobsters aren’t sentient’. Other kingdoms of life are not even considered by Humphrey.

Given the zeitgeist, the next obliged question is, of course, what about machines? Humphrey’s ‘warm brain theory’ cannot offer much here. He leans on the opinion that artificial devices could not be sentient. He vaguely suggests that one could create a functional duplicate, which begs the question (and, if feasible, would entail two mysteries to solve rather than one). He then sketches a flowery bet as to why humans may be motivated to build sentient robots: ‘human beings, in an act of cosmic generosity, will try to forestall the extinction of phenomenal consciousness by seeding the universe with sentient robots.’ In sum, he has little to offer here.

Towards the end, when it comes to the ethical corollaries of his proposal, Humphrey refrains from sermonizing. And yet he does not shy from claiming that ‘dogs matter to themselves in a way that octopus do not… We should care *for* them even if their not being sentient means there is less reason to care *about* them’. He gently declines the idea that one can derive an ought from an is: ‘human ethics come from somewhere else.’

Let me add some final reflections.

Humphrey’s book is the latest in a prolific series of attempts to solve the hard problem. His work spans decades and disciplines. It is impressive. His intimate style and fine pen bail the time spent with such cleverly fatigued ideas. His pedigree is impressive as well. He comes from an illustrious family of intellectuals, including Maynard Keynes and A.V. Hill. Furthermore, he has been riding on Dennett’s tails, curling new winds. Altogether, remarkable.

And yet, given that Humphrey’s intellectual adventure has been going on for half a century, I cannot but wonder whether he has ever seriously entertained that the problem may not be consciousness but materialism. Why not allow oneself to look for the missing key elsewhere, rather than under the mainstream streetlight? Some people think that you ought to be a materialist because you are smart (and the converse). Indeed, one has the feeling that Humphrey thinks that he’s smarter than previous and contemporary physicalists. He wants to solve the hard problem while others can’t. But, like a dove propelled

towards the sun, one has the impression his cultivated and multi-faceted thought cannot achieve the escape velocity required to leave behind the gravitational pull of his metaphysical commitments. As a result, his accounts are colourful, but trapped in the same walls of the old citadel of promissory materialism, just painted in a different yellow.

Stuck with the physicalist credo, the orthodoxy must constantly re-invent scenarios to save not only the appearances but their worldview, which in turn has become a sign of identity. In other words, materialism is less a habit of life than of academic posture. Humphrey is a precious exemplar of an open-minded scholar unable to take alternatives to materialism seriously.

Materialism is the tail that still wags the science of consciousness dog. But things are changing. You can see it in all those hyperbolic adjectives currently used to 'make matter great again'. Perhaps, if card-carrying materialists wondered better about what matter is (and is not), they would stop pretending that it can give rise to mind and consciousness and could start entertaining the reality of soul and spirit. Perhaps, leaving aside overused evolutionist and emergentist views for a moment, they could acknowledge that physicists themselves do not know what matter is and that, after diving into its supposed brick-like constituents, they discovered (more than a century ago!) that it is more mind-like than they could have imagined. Perhaps, if they considered alternative starting points that question the very question they seek to answer (how minds bootstrap themselves from mindlessness), consciousness studies would look less like rationality on steroids making a fool of itself.

Paraphrasing the singular words of Nietzsche (on the relationship between art and nature), Humphrey articulates a prodigious koan: 'Sensations have evolved to be a metaphysical supplement to the reality of our embodiment… we have a phenomenal self *in order not to die of materialism*' (my italics). I refrain from unpacking this lustrous phrase here, but it suggests a final intriguing consideration.

At the end of the day, Humphrey's proposal enacts a kind of 'terminal lucidity' of materialism, namely, the last surge of mental clarity of a journey that has served science well for about four centuries but whose end is near. The hard problem of consciousness, if I may pamper the allegory further, is materialism's admission to the palliative care unit of logic. And with it, a panoply of sad but predictable symptoms unfolds, including the grumpy dismissal of panpsychism ('even a teacup has a smidgeon of conscious feeling'),

the crypto-dualistic pirouettes of card-carrying physicalists (I quote a few above), or the snub to one of the very few mathematical theories of consciousness out there, namely, integrated information theory, a scientific theory that every angry materialist worth their ideology must scorn (some even call it a 'pseudoscience'). On that end, Humphrey's remarks say more about himself than about the theory ('I'm not sure I fully understand the maths') and also about his failure, or unwillingness, to grasp what IIT is about ('Why should we engage with a theory of subjective phenomenal experience that unashamedly leaves out the subject and leaves out the experience?'). Science should not be a monoculture of ideas.

Borrowing his paraphrasing of the American poet Mary Oliver, let me add that 'it would be terrible to be wrong about Humphrey's proposal. Terrible to be wrong indeed, but irresponsible not to be right.' Humphrey's '*invention of consciousness*' tells us less about consciousness itself than about what conscious beings, particularly faithful materialists, can (and must) invent about it. It isn't just that the emperor has no clothes. When it comes to mainstream consciousness studies, the clothes have no emperor.