

Nicholas Humphrey

## *Easy Does It*

### *A Soft Landing for Consciousness*

**Abstract:** *Problem reports result from several misunderstandings about the nature and functions of phenomenal consciousness. I discuss some philosophical and scientific correctives that, taken together, can make the hard problem seem less hard.*

Colin McGinn provides a graphic example of someone struggling with the problem: ‘Matter is just the wrong kind of thing to give birth to consciousness... You might as well assert, without further explanation, that numbers emerge from biscuits, or ethics from rhubarb’ (McGinn, 1993, p. 160). Jerry Fodor illustrates it his way: ‘Nobody has the slightest idea what consciousness is, or what it’s for, or how it does what it’s for (to say nothing of what it’s made of)’ (Fodor, 2004, p. 31).

But consciousness<sup>1</sup> is not the only problem that makes clever people wobble at the knees. Let’s consider, since McGinn alludes to it, the problem of where *numbers* come from: natural numbers, negative numbers, irrational numbers, imaginary numbers, and so on. In the history of mathematics, every advance in number theory has given rise to what Chalmers calls ‘problem reports’. To paraphrase what he says about consciousness (Chalmers, 2018, p. 7): “There is a hard problem about some kinds of numbers”, “It is hard to see how a negative number could be physical”, “After explaining the practical uses of

Correspondence:

Nicholas Humphrey, Darwin College, Cambridge, UK.

Email: [humphrey@me.com](mailto:humphrey@me.com)

---

<sup>1</sup> I’ll follow Chalmers in not always bothering to add the term phenomenal; but it will be phenomenal consciousness that is at issue in this paper.

imaginary numbers (e.g. in applications such as electrical engineering) there remains a further question, what actually are they?”, and so on. It is easy to get ordinary people to express puzzlement about what numbers amount to, and there’s a significant body of data about the “intuitive Platonist” judgments of both children and adults.’

So with numbers, too, there may be a meta-problem: the problem of explaining why we think numbers are hard to explain. ‘Biscuits’, as McGinn might have said, ‘are just the wrong kind of thing to give birth to numbers. You might as well assert that consciousness emerges from matter.’ Exactly. Or, rather, exactly not. For McGinn, presumably, is being ironical and trading on a philosophical insider’s joke. He has in mind Gottlob Frege’s mocking remarks about how ordinary people adopt ‘the childlike ginger-biscuit viewpoint’ on the reality of numbers. Frege wrote:

To the question of the nature of cardinal numbers we are given answers such as ‘a series of kindred things’, ‘an object consisting of elements of a single kind’; in brief, a heap of ginger biscuits is a number. If a man who had never thought about the matter was woken from his sleep with the question, ‘What is a number?’, then, in his initial confusion, he would be likely to produce expressions similar to those: ‘set’, ‘heap’, ‘series of things’. (Frege, 1893/2013, p. 150)

And:

It remains only to ascribe to the flavour of the biscuit some special meaning for the concept of number. (Frege, 1884/1980, p. vi)

Frege has seen that numbers are features of the *conceptual map*, not the *material territory*. Numbers can indeed be tokens of ‘empty’ concepts — concepts that have *sense* without *reference*. And with this, much of the mystery about if and how numbers really exist evaporates. Indeed Gottfried Leibniz, two hundred years earlier, had recognized the heady possibilities: imaginary numbers, he said, are ‘an invention... a marvel of analysis, a prodigy of the world of ideas, almost an amphibian between Being and Non-Being’ (Leibniz, 1702/1989, p. 396).<sup>2</sup>

I doubt McGinn will have expected his analogy between consciousness and numbers to lead anywhere. His general line has been that a solution to the hard problem must be forever beyond our

---

<sup>2</sup> ‘Ce miracle de l’Analyse, prodige du monde des idées, objet presque amphibie entre l’Être et le Non-être.’

comprehension, out of reach. However, even if it is not what he intended, I believe we can use the analogy to advance our understanding in two ways.

First, as to what consciousness is *not*. We must understand that, just as numbers are not a heap of biscuits, so consciousness is not 'a heap of nerve cells'. And theorists, however sophisticated, who assume that the quality of conscious experience simply springs into being when there's a sufficient accumulation of the right kind of brain activity have failed to grasp what kind of thing consciousness is. Tononi's integrated information theory, Penrose and Hameroff's quantum effects in microtubules, Strawson and Goff's elemental particles of consciousness are all in their way ginger-biscuit theories, in which consciousness arises from the flavour of the brain activity. We shouldn't even call them theories; they do not even have the merit of being wrong.

And then, as to what consciousness *is*. Just as numbers, in Frege's view, are objects in the conceptual landscape of arithmetic, we must recognize that conscious qualia are properties of how we *represent* a certain dimension of mental life: namely, the dimension that is often characterized as 'what it's like'. Admittedly the phrase 'what it's like' is a bit of a terminological embarrassment, which may not bear analysis. But I see no reason not to take the grammar at face value, and assume the 'it' of 'what it's like' has a perfectly good referent, which is the state of affairs — whatever this may be — that you, the subject, represent as having the conscious property in question. When, for example, you experience what it's like to look at the blue sky, you are taking something — something pertaining to the light at your eyes, presumably — to have the property of phenomenal blueness.

Still, we need to pin this down. The 'it' of what it's like is something you are representing in your mind. But what kind of thing, and where does it fit in to the rest of mental life? Chalmers is right to point out that people have introspective access to many different types of representations, but for only a subset of these can it be said it is like something to have them. Which types in particular are so privileged?

In my 1992 book *A History of the Mind* I summarized my own view on this. Following the philosopher Thomas Reid, with his account of 'the double province of the senses', I began by making a principled distinction between perception and sensation. I reviewed the evidence, physiological and psychological, that these are independent ways of interpreting what's happening at the sense organs, with different intended content and different phenomenology. While perception

provides a representation of ‘what’s happening out there’, sensation provides a representation of ‘what’s happening to me and how I feel about it’. Perception renders its object in relatively straightforward physical terms. But sensation additionally employs a palette of value-laden phenomenal properties that are not present in the content of perception or any other non-sensory object represented by the mind. So, in summary, I wrote (*ibid.*, pp. 97–8):

1. To be conscious is essentially to have *sensations*: that is, to have affect-laden mental representations of something happening here and now to me.
2. The subject of consciousness, ‘I’, is an embodied self. In the absence of bodily sensations ‘I’ would cease. *Sentio, ergo sum* — I feel therefore I am.
3. All sensations are implicitly located at the spatial boundary between me and not-me, and at the temporal boundary between past and future: that is, in the ‘present’.
4. For human beings, most sensations occur in the province of one of the five senses (sight, sound, touch, smell, taste). Hence most human states of consciousness have one or other of these qualities. There are no non-sensory, amodal conscious states.
5. Mental activities other than those involving direct sensation enter consciousness only in so far as they are accompanied by ‘reminders’ of sensation, such as happens in the case of mental imagery and dreams.
6. This is no less true of conscious thoughts, ideas, beliefs. Conscious thoughts are typically ‘heard’ as images of voices in the head — and without this sensory component they would drop away.

In the years since, I have not changed my view on this, except to add supporting arguments about the evolutionary history. I’ve argued that sensations have their origin in a primitive organism’s reflex motor responses to sensory stimulation — a form of bodily expression. I’ve called this ‘sentition’. From the beginning sentition *enacted* what the stimulation meant to the organism. This meant that, at a later stage, the organism could acquire the capacity to represent this meaning by the simple expedient of reading from an efference copy of the command signals. Eventually, in the course of evolution, the ancient responses became ‘privatized’, so that they no longer resulted in overt behaviour but became virtual responses contained within the brain. Yet, all along the subject could continue to read the virtual responses so as to get a picture of what the sensory stimulation felt like. And, even today, this is where sensations come from. Thus, when you experience what it’s like to look at the sky, you are reading *your*

*vestigial bodily response to the light at your eye, your bluing*, so as to represent how you feel about this interaction.

Now, in the context of the meta-problem, please note the licence this account permits us for theorizing about the exotic phenomenal features of consciousness. While there must presumably be limits to the properties we can meaningfully attribute to the physical world, there are no obvious *a priori* limits to the nature of subjective feelings. This means we have no logical reason to question, let alone object to, whatever qualities we find people attributing to the ‘it’ of what it’s like. We certainly have no warrant for saying that a particular phenomenal property is impossible, just because it’s non-physical. It may be true that properties such as ineffability, immediacy, intrinsicity, homogeneity — you name it — cannot sensibly be ascribed to the cold reality of the material world; but such properties could well be appropriate — even essential — to the ascription of meaning and value to ‘what’s happening here and now to me’.

I want to say, then, that the meta-problem of why it’s so difficult to frame a theory of *what consciousness is* is one we can put to sleep. When Fodor wrote ‘Nobody has the slightest idea what consciousness is’, he probably believed, as many do, that consciousness could not be what he knew perfectly well it was, because something with such properties would not be theoretically permissible. But he misread the situation. As properties of how you feel about what’s happening to you, phenomenal properties are just as permissible as the mind will permit.

Still, as Chalmers makes clear, meta-problems may live on at other levels. All right, so you represent how you feel about what’s happening at your sense organs by ascribing phenomenal properties. But maybe, as Fodor said, it’s still hard to explain what the representation ‘is for, or how it does what it’s for (to say nothing of what it’s made of)’. Chalmers provides as always a masterly introduction to these questions. But I think he leaves them further up in the air than, as scientists, we are obliged to.

So, first, what are subjective representations of consciousness made of? Mental representing — that’s to say, the taking of a state of affairs to have a certain meaning — will always have two separable components. There will be a representational target and a representing process, a taken and a taking, a map and a reading of the map (see my discussion in Humphrey, 2016).

Consider, for example, what happens when you look at a cartoon drawing in the *New Yorker*, and take it to be funny. There’s the

drawing that has been designed by the cartoonist to *offer up the joke*, and the process in your head that *sees the joke*. Now, in the case of consciousness, it so happens that *both* steps are in your head. You respond to sensory stimulation with an expressive response, then you read this response to discover what it's like for you to have this happening. If the response has indeed been designed by natural selection to *offer up what it's like*, what more can we say about how this works and why the results are so spectacular? What happened to sentience in the course of evolution?

In E. Nesbit's children's story, *Five Children and It* (Nesbit, 1902), the children discover a magical sand fairy which lives hidden beneath the surface of a gravel pit. The fairy, 'It', calls itself a Psammead. In writings going back to the 1970s, I've constructed a story about how sentience evolved to be a quasi-magical attractor state hidden in the brain, and suggested a plausible evolutionary trajectory by which this could have happened (Humphrey, 1992; 2006; 2011). I've called this attractor state the *ipsundrum*. What's quasi-magical about the *ipsundrum* is that when you read it to discover how you feel about the stimulation, you find your feeling has those weird and wonderful phenomenal properties.

Magic? Well, not really, of course. An illusion? No, I wouldn't say that either. It's true that the *ipsundrum*, as such, does not have phenomenal properties, and what you read into it does have them. But this doesn't mean you are *misreading* the *ipsundrum*. If I may borrow terms from semiotics (without buying into the whole works), the *ipsundrum* is a *signifier*; what it's like is the *signified*; they exist on different conceptual levels and the connection between them is *unmotivated*. The fact that what it's like for you at the level of feeling has properties not present — indeed not even conceivable — for the *ipsundrum* at the level of brain activity does not in any way imply that the feeling misrepresents what's really the case.

Consider the *New Yorker* cartoon again. The cartoon on the level of paper and ink is not funny in itself. At most we might say it's proto-funny. The funniness emerges only when the cartoon is *observed by someone with a sense of humour who gets the joke*. But the fact that the funniness depends on the observer's attitude, and is a property of the intentional object, does not of course mean it is an illusion.

Likewise the *ipsundrum* does not have phenomenal properties in its own right. The most we might say is that it's proto-phenomenal (or 'phenomenous', as I put it in an earlier paper — Humphrey, 2008). The phenomenality emerges only when this brain activity is *read by*

*an inner observer with a sense of consciousness* who gets the sensation. But again this doesn't mean the phenomenality is an illusion.

Note that in neither of the cases just discussed is the intentional object of the representation a material thing that the subject is wrong to believe in because no such thing actually exists in the physical world, as for example with the Impossible Triangle or Santa Claus. Rather, the intentional object is a subjective feeling that the subject is right to believe in because just such is the reality of his evaluation of events. For the subject the joke really is funny, the sensation really is phenomenally painful. True, there is nothing at the level of physical reality, within or without the brain, that satisfies the property of being phenomenally painful, or for that matter of being funny. But there *is* something at the level of psychological reality within the mind. Indeed, on this vexed question of illusionism versus realism about consciousness, I would say that on this account I lean towards realism.

Anyhow, so much for 'what consciousness is made of'. Following Fodor's agenda, we're left with the question of 'what it is for, and how it does what it's for'. The design question. Why is this so difficult? Fodor has explained: 'There are several reasons why consciousness is so baffling. For one thing, it seems to be among the chronically unemployed... What mental processes can be performed only because the mind is conscious, and what does consciousness contribute to their performance? As far as anybody knows, anything that our conscious minds can do they could do just as well if they weren't conscious' (Fodor, 2004, p. 31).

Chalmers has in the past voiced a similar belief in the functional irrelevance of consciousness, an intuition that Dennett has dubbed 'the zombic hunch' (Dennett, 2001). These days Chalmers says he has respect for evolutionary accounts, so presumably he is less convinced than he was that consciousness makes no difference at all. In fact, as he says here, he now leans towards the scientifically plausible idea that phenomenal properties prove their worth when it comes to modelling the mind from the inside. They may be helpful, possibly essential, to a workable theory of mind: '[Introspection] needs to keep track of similarities and differences in mental states, but doing so directly would be inefficient, and it does not have access to underlying physical states. So it introduces a novel representational system that encodes mental states as having special qualities' (Chalmers, 2018, p. 26).

In an essay on the 'Uses of Consciousness' thirty years ago, I made a very similar point:

The inner eye provides a picture of its information field that has been designed by natural selection to be a useful one — a user-friendly description, designed to tell the subject as much as he requires to know... We can assume that throughout a long history of evolution all sorts of different ways of describing the brain's activity have in fact been experimented with — including quite possibly a straightforward physiological description in terms of nerve cells, RNA etc. — What has happened, however, is that only those descriptions most suited to doing [introspective] psychology have been preserved. Thus the particular picture of our inner selves that human beings do in fact now have — the picture we know as 'us', and cannot imagine being of any different kind — is the description of the brain that has proved most suited to our needs as social beings. (Humphrey, 1987, p.18)

If Claude Levi-Strauss were in on this discussion, he might have quipped that phenomenal properties are *good to think with*. They provide a palette of very special properties for picturing the mental life of ourselves and, by extension, of others.

In my own more recent work I have raised a further possibility. I've argued that phenomenal properties are not only good to think *with*, they are good to think *of*. When we look at the natural history of consciousness, how it impacts people's lives, we see that people value consciousness as an extraordinary asset that enhances their own meta-physical significance. And this, I've suggested, is where the paradoxical dimensions of consciousness truly come into their own.

My suggestion is that in the course of human evolution, our ancestors who thought of their own consciousness as metaphysically remarkable — existing outside normal space and time — would have taken themselves still more seriously as Selves. The more mysterious and unworldly the qualities of consciousness, the more seriously significant the Self. And the more significant the Self, the greater the boost to human self-confidence and self-importance — and the greater the value that individuals place on their own and others' lives. In which case it is easy to see how the very qualities of consciousness that seem to render it so mysterious and magical would have been the occasion for consciousness becoming a runaway evolutionary success. In fact these qualities would soon have been designed in. (Humphrey, 2006, p. 131)

In short, human beings will have been set up by nature to relish their own inexplicability. Natural selection will have actively fostered the meta-problem.

Yet I realize, as does Chalmers, that a particular question hangs over the idea that consciousness makes mental life more vivid and distinctive at whatever level. Why should phenomenal properties be restricted to sensations? Why have the other attitudes not been invited

to the party: beliefs, desires, and so on? When it comes to modelling — and valuing — our own and others' minds, wouldn't it be extra helpful if every kind of mental state had its own phenomenal signature? As Chalmers says, 'it is not really clear why access to a [sensory] modality as opposed to an attitude should make such a striking difference' (2018, p. 24).

I presume the answer must be found in evolutionary biology rather than philosophy. No doubt, in an ideal world, there could be a still better way of doing things. If we were designing robot minds to have a capacity for introspection, such that they would take themselves as seriously as human minds do, we might indeed choose to introduce phenomenal properties in areas of functioning where human beings don't have them. But natural selection, working with the human brain, can only have done its contingent best. And I suspect there was an historical practical reason why mental states other than sensations haven't been upgraded.

As described above, the representation of 'what it's like', in the case of sensations, is built on a unique neurobiological substrate: namely, the *ipsundrum*, an elaboration of an evaluative motor response to sensory stimulation. Other propositional attitudes, such as beliefs, however, do not originate in motor activity. Most likely, then, there was no available path by which selection could have enhanced them in an equivalent way. And there's something else. At an early stage of evolution the motor responses underlying sensation were still out in the open as overt expressive movements. This meant it would have been possible for another individual to *simulate* the motor pattern — and so potentially the feeling — by means of 'mirror neurons'. I've suggested (Humphrey, 2006, pp. 103–09) that this kind of mirroring may have remained a possibility even after the responses were internalized, meaning that sensations could be relatively easily channelled from mind to mind. However, this would not have been the case with other mental states. This could have provided an extra reason, then, for sensations to be selected for special treatment.

To sum up. I've discussed four features of consciousness that people find hard to explain, the four that Fodor picks out as being those nobody has the slightest idea about: 'what consciousness is, what it's made of, what it's for, and how it does what it's for.' I've tried to dispel the aura of invincibility that surrounds these questions. I've proposed candidate answers, within a materialist scientific framework, that could provide relatively easy explanations for the central phenomena, while at the same time explaining why these answers are

far from obvious. These answers may not be correct. But they provide a proof of principle that the hard problem can be solved.

Chalmers is attracted to illusionism, but confesses he still hankers after realism. I'd like him to look more closely at the solutions I've put forward. Mine may not be a realist theory in the sense he hopes for. But it's not an illusionist theory either. I'd settle for 'amphibian theory', if that would help. Still, nothing much rides on what we call it. The facts, once science discovers them, will be just what they are.

### *Acknowledgments*

Thanks to Keith Frankish and Dan Dennett for their comments, to Donald Gillies for clarification about Frege, and to François Kammerer for his detailed and insightful editorial notes.

### **References**

- Chalmers, D.J. (2018) The meta-problem of consciousness, *Journal of Consciousness Studies*, **25** (9–10), pp. 6–61.
- Dennett, D. (2001) The zombic hunch: Extinction of an intuition?, *Royal Institute of Philosophy Supplement*, **48**, pp. 27–43.
- Fodor, J. (2004) You can't argue with a novel, *London Review of Books*, 3 March, pp. 30–31.
- Frege, G. (1893/2013) *The Basic Laws of Arithmetic*, Ebert, P.A & Rossberg, M. (ed. & trans.), Oxford: Oxford University Press.
- Frege, G. (1884/1980) *Foundations of Arithmetic*, Austin, J.L. (trans.), New York: Harper.
- Humphrey, N. (1987) *The Uses of Consciousness*, New York: American Museum of Natural History.
- Humphrey, N. (1992) *A History of the Mind*, London: Chatto & Windus.
- Humphrey, N. (2006) *Seeing Red: A Study in Consciousness*, Cambridge, MA: Harvard University Press.
- Humphrey, N. (2008) Getting the measure of consciousness, *Progress of Theoretical Physics Supplement*, **173**, pp. 264–269.
- Humphrey, N. (2011) *Soul Dust: The Magic of Consciousness*, Princeton, NJ: Princeton University Press.
- Humphrey, N. (2016) A riddle written on the brain, *Journal of Consciousness Studies*, **23** (7–8), pp. 278–287.
- Leibniz, G.W. (1702/1989) *La naissance du calcul différentiel*, Parmentier, M. (ed. & trans.), Paris: Vrin.
- McGinn, C. (1993) Cosmology and consciousness: Hyperdualism ventilated, in Davies, M. & Humphreys, G.W. (eds.) *Consciousness*, pp. 155–177, Oxford: Blackwell.
- Nesbit, E. (1902) *Five Children and It*, London: Unwin.