

Annaka Harris

# *A Solution to the Combination Problem and the Future of Panpsychism*

**Abstract:** *This paper supports the scientific position that panpsychism is a valid category of possible resolutions to the hard problem of consciousness, and it focuses on a solution to the ‘combination problem’ in panpsychism. I argue for a new way of thinking about consciousness in which consciousness is not viewed in reference to subjects, and that the concept of a ‘subject’ is borne of the illusion of self. Therefore, we don’t face a combination problem if the notion of a subject is superfluous and consciousness itself is pervasive in the form of a field. The paper is also a more general discussion about the importance of pursuing this scientific question in the twenty-first century: is consciousness a more fundamental aspect of the universe than we have previously assumed?*

## **1. Introduction**

The great mystery of consciousness is why matter lights up with felt experience. After all, we are composed of particles indistinguishable from those swirling around in the Sun; the atoms that compose your body were once the ingredients of countless stars in our universe’s past. They travelled for billions of years to land here — in this particular configuration that is you — and are now reading these words. Imagine following the life of those atoms from their first appearance in space-time to the very moment they became arranged in such a way as to start *experiencing* something.

Correspondence:  
Email: [contact@annakaharris.com](mailto:contact@annakaharris.com)

Many assume there is probably no felt experience associated with the microscopic collection of cells that make up a human blastocyst. But over time these cells multiply and slowly become a human baby, able to detect changes in light and recognize its mother's voice, even while in the womb. And unlike a computer, which can also detect light and recognize voices, this processing is accompanied by an *experience* of light and sound. First, as far as consciousness is concerned there is nothing, and then suddenly, magically... *something*. The mystery lies in the transition. However minimal that initial something is, experience apparently ignites in the inanimate world, materializing out of the darkness.

People often use the word 'consciousness' to refer to higher-order functions, such as self-awareness and thought. But I'm addressing consciousness in the most fundamental sense — as felt experience, regardless of the content. In my 2019 book *Conscious*, I use Thomas Nagel's distinction — whether or not it's *like something* to be an organism — to get to the heart of what I'm referring to. Is it *like something* to be you in this moment? Presumably your answer is yes. Is it *like something* to be the chair you're sitting on? Your answer will (most likely) be an equally definitive no. We can all use this simple difference as a reference point for what I mean by the word 'consciousness'. Is it *like something* to be a grain of sand, a bacterium, an oak tree, an ant, a mouse, a dog? At some point along the spectrum the answer is yes, and the great mystery lies in why 'the lights turn on' for some collections of matter in the universe.

Despite the many advances in neuroscience over recent decades, the hard problem of consciousness persists, and this has caused more scientists and philosophers to wonder whether we've been thinking about the problem backward. Is it possible that consciousness, rather than arising when non-conscious matter behaves a particular way, is an intrinsic property of matter — that it was there all along? This is the type of question Philip Goff explores in depth in his recent book, *Galileo's Error* (2019). It is also the direction my own thinking has led me.

Because our only evidence of consciousness comes from self-report and can never be confirmed first-hand by an outside observer, we are forced to make one of two assumptions: (1) felt experience arises out of non-conscious processing, or (2) consciousness is a property of matter itself (or of something even more fundamental) — a category of views typically referred to as 'panpsychism'. Cases of split-brain patients, locked-in patients, and anaesthesia awareness should give us

pause when we assume that islands of consciousness cannot possibly lie beyond the awareness of the ‘reporter’ in a human brain. But we have labelled the very distinction between conscious and unconscious brain processes solely on the basis of what can be reported on and what cannot. This seems unavoidable, of course, but it’s important to remember that all our starting assumptions about consciousness are based on reportability (and to remember that they are in fact *assumptions*). I assume that the processes taking place in my liver are unconscious because ‘I’ am not conscious of them, but this doesn’t necessarily rule out the possibility that liver processing entails consciousness. The fact that a pregnant woman doesn’t experience any of the things her unborn baby is experiencing is obviously no reason to believe the baby is unconscious.

We have a deeply ingrained intuition that systems that act like us are conscious and those that don’t are not. This is reasonable, because the only source of reporting or behaviour we can interpret, and therefore have access to, comes to us from organisms most similar to us. As a result, the sciences have led with assumption #1: that consciousness arises out of complex processing (namely in brains). But is this actually the most rational starting place? The idea that consciousness emerges out of non-conscious material, in fact, represents a kind of failure of the typical goal of scientific exploration: to arrive at as simple an explanation as possible. The celebrated biologist J.B.S. Haldane, for example, argued that the notion of the ‘strong emergence’ of consciousness is ‘radically opposed to the spirit of science, which has always attempted to explain the complex in terms of the simple... If the scientific point of view is correct, we shall ultimately find them [signs of consciousness in inert matter], at least in rudimentary form, all through the universe’.<sup>1</sup> The truth is that we simply don’t know if consciousness only arises only in brains (or if it is a sign of complexity at all, for that matter), and we are heavily relying on an assumption when we make that case. Each transformative shift in our understanding of the universe has delivered the ego-shattering message that we’re not special — Earth is not the centre of the universe, and life, including the human brain, is made up of the same particles as the stars. Perhaps it’s time to relinquish our last claim to specialness. If a fundamentally new aspect of matter doesn’t

---

<sup>1</sup> See Skrbina (2017).

enter the picture at some point, is it possible it was there to begin with and is simply part of the fundamental stuff that is all around us?

My own sense of the correct resolution of the hard problem of consciousness, whether or not we can ever achieve a true understanding, is split between a brain-based explanation and a panpsychist one. So although I'm not convinced that panpsychism offers the correct answer, I *am* convinced that it is a valid category of possible solutions and cannot be easily dismissed. In recent years, my primary focus has been what I see as a solution to the 'combination problem' in panpsychism. In the second and third sections of this paper, I will lay out my argument for why I think we don't necessarily face a combination problem. But first, I will discuss some of my recent thoughts about panpsychism in general.

As an umbrella term to designate any view of the universe in which consciousness is an intrinsic quality of matter, 'panpsychism' is very useful shorthand for this category of thought. But ultimately, I find it to be an obstacle to fruitful dialogue. I have, therefore, made a change in my writing recently and have chosen to omit the word going forward. I find it more effective to simply pose what is now considered by many scientists to be a legitimate scientific question: is consciousness a more fundamental aspect of the universe than the sciences have previously assumed? If we choose to follow assumption #2, where do we land?

When contemplating this category of hypotheses, it's important to first distinguish between consciousness and thought. We should be careful not to reflexively rail against the idea that rocks and spoons are conscious, which is obviously false when put in terms of a rock's being conscious *as a rock*. If consciousness is fundamental, all matter must entail consciousness by definition; but that doesn't mean it makes sense to specify such things as 'moon consciousness' and 'tree consciousness'. We would expect that the region of space-time occupied by a rock, say, entails consciousness because matter is present there. We can't imagine what that region of particles feels like (or even that it has a unified perspective at all, which seems unlikely). What we can be fairly sure of, however, is that it doesn't contain a human-like experience or even a single 'point of view'. Just as we wouldn't expect (the collection of atoms that make up) a rock to get up and walk or sing — that's not what atoms configured in such a way *do* — we wouldn't expect it to have a single, unified point of view. And we certainly wouldn't expect it to have anything like thoughts or intentions.

We should still assume that information in a complex and integrated form is required to produce experiences like ours. We shouldn't feel compelled to wonder whether there is a specific 'rock consciousness' any more than we're compelled to wonder whether there is a 'rock + the-five-blades-of-grass-the-rock-is-touching consciousness'. That description of consciousness is based on an anthropomorphic view, projecting separateness in isolated packages: me, you, rock, spoon. But perhaps felt experience *is* present — in a form we can't imagine — across the matter in any given area of space-time.

## 2. Subjects and Selves

We have different ways of using the word 'self'. There's the autobiographical self, which is the story of who I am: *my name is Annaka. I have two daughters. I'm a good swimmer...* If I wake up tomorrow with amnesia and don't remember my name or anything else about who I am, I've lost my sense of being the 'self' called Annaka, and I now feel like a very different person. But the sense of self I will be referring to in my discussion of the combination problem goes deeper than the autobiographical self and isn't necessarily bound up with a specific identity. A sense of self can still persist even if I lose my autobiographical self. It's the 'I' that amnesiacs refer to when they say, 'I don't know who I am! I don't remember my name or where I live!' The deeper sense of self is the experience of being a single, independently existing entity that has a precise centre or location and is doing the experiencing. And this concept of the self is an illusion.

The neuroscientist Anil Seth studies how this 'illusion of self' gets constructed. He writes, 'the predictive machinery of perception when directed at the self makes it seem as though there really is a stable essence of "me" at the centre of everything' (Seth, 2021). But he explains that this does not accurately represent the underlying reality: 'Just as experiences of redness are not indications of an externally existing "red," experiences of selfhood do not signify the existence of an "actual self" (*ibid.*). In addition to processes such as interoception and memory, which help to construct an illusion of self, the brain suffers a type of change-blindness with respect to the self, further strengthening the illusion. Seth explains, 'We are becoming different people all the time. Our perceptions of self are continually changing — you are a slightly different person now than when you started reading this chapter — but this does not mean that we perceive these changes. This subjective blindness to the changing self has

consequences. For one thing, it fosters the false intuition that the self is an immutable entity, rather than a bundle of perceptions' (*ibid.*). While it's admittedly a tough illusion to relinquish, we know that the experience of self does not offer us accurate intuitions about consciousness in general, and we know from people's experiences in meditation and under the influence of psychedelic drugs that consciousness persists when the default mode network quiets down and the illusion of self drops away.<sup>2</sup>

### 3. The Combination Problem

When considering a view of the world in which consciousness is fundamental, the question arises: if the most basic constituents of matter do indeed have some level of conscious experience, how is it that when they form a more complex system — such as a brain — those small points of consciousness combine to create a new conscious subject? For instance, if the individual atoms and cells in my brain are conscious, how do those separate spheres of consciousness merge to form the consciousness 'I' am experiencing? What's more, do all of the smaller, individual points of consciousness cease to exist after giving birth to an entirely new point of view? This is referred to as the *combination problem*, and according to the *Stanford Encyclopedia of Philosophy*, it's 'the hardest problem facing panpsychism' (Goff, 2001). The entry cites a passage from William James as the primary inspiration for the problem: 'Take a hundred of them [feelings], shuffle them and pack them as close together as you can (whatever that may mean); still each remains the same feeling it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean. There would be a hundred-and-first feeling there, if, when a group or series of such feelings were set up, a consciousness belonging to the group as such should emerge...' (*ibid.*). The combination problem has kept many scientists and philosophers who are otherwise willing to entertain the idea that consciousness is a fundamental property of matter from fully endorsing this area of study. However, it seems to me that the obstacle one faces here isn't a combination problem but the confusion of *consciousness* with the experience of *self*.

---

<sup>2</sup> For more on the illusion of self, see Chapter 5 of Harris (2019).

When discussing the combination problem, philosophers and scientists tend to speak in terms of a ‘subject’ of consciousness. My claim is that this is actually another way of pointing to the experience of self in its most basic form. In a paper on the combination problem, David Chalmers writes, ‘How could any phenomenal relation holding between distinct subjects... suffice for the constitution of a wholly new subject?’ (Chalmers, 2012). Likewise, when Greg Rosenberg discusses the related ‘boundary problem’ (also referenced by James), he writes, ‘I start with the observation that consciousness has inherent boundaries. Only some experiences are part of my consciousness; most experiences in the world are not. Arguably, these boundaries are what individuate me as an experiencing subject in the world’ (Rosenberg, 2004).

But I think it’s wrong to talk about a ‘subject’ of consciousness. I believe that this way of framing things is entirely due to the experience of self, which we know to be an illusion. All we truly know is that content (or qualia) appears in the universe. The claim that qualia appear *to a subject* in the universe is an additional (and I think unnecessary and false) step. The most powerful sense in which there is an experience of ‘I’ or ‘me’ as a single entity to which certain experiences are presented is through the connection of experienced moments of qualia through memory. If I simply experienced green and then sharp pain and then happiness and then bright light, without memory causing these qualia to trail along, there would be no sense in which we would say this is happening to ‘me’ or to a subject at all. We would just say that qualia are appearing in the universe — like bubbles in a pot of boiling water. There is no subject that this content is appearing to. It’s just appearing. And each quale, by definition, is always limited to that specific quale.

Viewing consciousness through the illusion of self gives rise to other categories of the combination problem that don’t necessarily specify a ‘subject’ *per se*, but which I nevertheless believe still originate with the illusion of self: the *quality combination problem* and the *structure combination problem* (Chalmers, 2012). It is in this context that philosophers often refer to the ‘privacy’ of conscious experience. But again, this strikes me as a perspective from within the illusion of self. Each individual quale can be considered private, but only in the sense that it is defined by some particular qualities and not others — for changing it or combining it with another quale would, by definition, change the quale. But it is not private *for* any subject. When we look at each individual moment of experience — red, pain,

pressure, etc. — we can see that qualia arise in the universe in and of themselves (and, theoretically, there could be areas in space-time where pure consciousness exists with no qualia).

Additionally, qualia are not static, and this can be made evident empirically through introspective training such as meditation. Breaking down the flow of experiences into a stream of ‘present moments’ can help shift our sense of what we typically call ‘subjects’ or ‘selves’. In each moment, new content appears, but the content is clearly not being experienced by a subject. Some Buddhist teachings more accurately refer to the present moment as the ‘passing moment’ and, when zeroing in on these passing moments, one notices that the red of the flower (sight) and the whistle of the bird (sound) don’t arise simultaneously, nor are they solid or concrete in any real sense. Each quale is experienced sequentially and as a process, not as a static object. Then, through memory, the illusion of a full picture is given. But when one is carefully attending to each passing moment, it becomes clear that those ‘memory snapshots’ are not an accurate rendering of what the experience actually entailed.

In terms of privacy, when we realize that there is no solid centre we can label ‘you’ or ‘me’, it makes no sense to talk about where my consciousness ends and yours begins. Content is arising, and some of that content is shared across time through memory, as yet more content. Your perception of yellow isn’t ‘yours’. It’s simply an experience of yellow arising in the universe, derived from interacting forces and fields. It’s not private in the typical sense. There is no self for it to be private *for*.

#### 4. A Solution to the Combination Problem

Rather than an obstacle to theories that place consciousness in a fundamental role, the combination problem may be a reason to favour the proposition that consciousness is a fundamental feature of the universe in the form of a continuous, pervasive field, analogous to space-time. Just as space-time and gravity have an interactive relationship, consciousness might be thought of as a fundamental ‘field’ that interacts with, and is integral to, matter. We typically don’t think of space-time as bits and pieces that build on one another (it’s simply *everywhere*), and I don’t think we should be tempted to think of consciousness, if it is indeed a pervasive field, as divisible into building blocks either. Rather, it makes more sense to talk about a field that contains a range of *content* — the content depending on the other



forces or fields it's interacting with. In the same way that gravity is a two-way street — matter warps space-time and the shape of space-time determines how matter moves — a consciousness field would imbue matter with another property, giving rise to the range of content experienced. In this view, content is perhaps divisible, but consciousness isn't. Therefore, consciousness is also not interacting *with itself*, as it would be in the act of 'combining'. Considering consciousness to be fundamental allows for matter to have a specific internal character everywhere, in all its various forms. It doesn't necessitate a reductionist explanation in which all qualia are made up of little 'quale building blocks'. It also allows for content to overlap in space, similar to the way sound waves at different frequencies coexist in the same space by essentially 'passing through' one another. The science journalist George Musser uses an analogy to sound to help explain the stratification of nature by scale: 'Sounds of long and short wavelengths are oblivious to each other; if you sound a deep bass note and a high treble pitch simultaneously, each ripples through the room as though it were the only sound in the world... These waves overlap in the three dimensions of space through which they propagate, yet they're independent of each other, as if they were located in different places' (Musser, 2015).

If consciousness is fundamental, then the questions that prompt the combination problem are potentially the same as all the other questions we might ask about space-time in which we *don't* anticipate this problem. All matter would entail consciousness, and complex systems, such as human brains, would give rise to certain types of content in those locations in space-time. Even if each individual atom has its own experience, *consciousness itself* is not necessarily isolated. The matter might be isolated, and therefore the *content* associated with the consciousness at that location would be isolated. But consciousness itself would not be said to be isolated. Again, we can think of consciousness as analogous to space-time: how it's affected by matter depends on the matter in question (its mass, in the case of space-time). Similarly, a consciousness field might be 'shaped' by matter in terms of experiential quality or content. This line of thinking yields interesting questions: how does the content that appears in an area of consciousness depend on the configuration of matter present in that location in space-time? Are there sometimes countless experiences of overlapping or merging content in the same location?

In a related conversation I had with the neuroscientist Christof Koch, we discussed what might result from a hypothetical experiment

in which two brains were connected as successfully as the two hemispheres of a single brain are connected. Since various experiments with split-brain patients — people whose right and left hemispheres have been surgically disconnected — have shown that the contents of consciousness can be separated, would two brains wired together produce a new, integrated mind? If Christof and I had our brains wired together, for instance, would it create a new Christof-Annaka consciousness — a new single point of view? Would a new mind be produced, with access to all the content that had previously been experienced separately by our brains — all our thoughts, memories, fears, abilities, etc. — constituting a new ‘person’?

Even if the answer is yes, I don’t think we encounter a combination problem in this thought experiment. We run into problems only if we see my and Christof’s conscious experiences as ‘selves’ or ‘subjects’ — permanent structures of consciousness with fixed boundaries. In the instance of connecting two brains, we might simply have an example of consciousness changing its content or character, in the same way the content of your consciousness changes predictably along with the properties of the physical world. For example, when your eyes are open and then you close them, the trees and sky are available to your field of view and then they’re not (or you could say that an *experience of trees and sky appears, followed by an experience of darkness*). When you dream, your brain states produce an experience of environments quite different from your actual surroundings, and perhaps even the feeling that you are a different person altogether. During both my pregnancies, I found myself experiencing drastic variations in the contents of my consciousness — sensations in my uterus I had never before known were on the menu of experience, an obsession with tomatoes and tomato sauces in every form, feelings of panic and other more amorphous emotions, physical pain, insomnia. I didn’t feel like ‘myself’. And I expect I wouldn’t feel like myself during a mind meld with a 62-year-old male neuroscientist either. But it doesn’t necessarily point to a *combination problem* for consciousness. Even in our daily lives, content comes and goes, and consciousness itself can seem to flicker in and out.

In the case of connecting two brains, the part of one brain that gives rise to the memory of mailing a letter yesterday, for example, could now give rise to that memory alongside all the qualia (including memories) that are possible in other areas of this new, merged brain. But in connecting two brains, no selves disappear, nor does a new ‘self’ appear. When I don’t remember something I’ve done in the past,

I don't tend to think of the 'past me' as someone else. Similarly, am I a new self when I learn something new? The past me had no idea what it was like to ski; the present me does. We're not at all tempted to say that the old me has ceased to exist, giving birth to a new mind — an updated version of me that can ski: 'The old me died and gave birth to a new me who skis!' This is clearly not the correct way to think about it. As long as we're speaking in terms of selves, we are essentially new selves in each new moment. I am a different self now from the self I was 20 minutes ago, with new content appearing because my brain is in a different state than it was in 20 minutes ago (less so than if I've ingested LSD during that time, rather than having just written a few paragraphs, but new just the same). So, too, for merged brains — they would share content and, therefore, some content would appear together (or in new forms) that wasn't possible before. It may seem legitimate to ask, 'Where did those two original subjects go?' But they didn't go anywhere. They weren't there to begin with!

The experience of a human brain is largely one of confusing consciousness with the experience of self, and we run into a combination problem only when we drag the concept of a 'self' or a 'subject' into the equation. The solution to the combination problem is that no 'combining' at all is going on with respect to *consciousness itself*. Consciousness could persist as is, while the character and content changes depending on the arrangement of the specific matter in question. In my analogy to the pot of boiling water, the bubbles are the content, and the water is consciousness. Maybe content is sometimes shared across large, intricately connected regions and sometimes confined to very small ones, perhaps even overlapping. If two human brains were connected to each other, both people might feel as if the content of their consciousness had simply expanded, with each person feeling a continuous transformation from the content of one person's consciousness to the whole of the two, until the connection was more or less complete. It's only when you insert the concepts of 'him', 'her', 'you', and 'me' as discrete entities that the expanding or merging of content becomes a combination problem.

Here I'm reminded of the classic device of characters switching places in a story or a film. But when we look closely at what this actually entails, it becomes obvious that there's no 'self' to transport from one person to another. Being someone else would be no different from what it's already like to be that person. It seems paradoxical, but we end up simply stating the obvious: 'That's what it's like to be over there as that configuration of atoms, and this is what it's like to be

over here as this configuration of atoms.’ It’s analogous to saying: ‘The configuration of atoms that compose a leaf result in all its expected leaf properties, but if you take all of those atoms and reassemble them into a liquid, they will take on the expected properties. That’s what molecules *do* in that configuration, and this is what they *do* in this configuration. Likewise, that’s what molecules *feel like* in that configuration, and this is what they *feel like* in this configuration.’

Before we can move forward with a theory in which consciousness plays a fundamental role in the universe, I think it’s imperative that we weed out and untangle the illusion of self from our description of consciousness. If consciousness itself doesn’t combine, then we no longer face a combination problem. The experience of consciousness need not be continuous or maintained as an individual self or subject. Nor is it necessarily extinguished when the smaller constituents of matter combine to make more complex systems, like brains. The human sense of being a self, along with an experience of continuity over time through memory (in which trails of previous qualia remain in circulation), may in fact be a very rare form of content. Is it possible that alongside the conscious experience of ‘me’ is a much dimmer experience of each individual neuron, or of different collections of neurons and cells in my body and beyond, like an orchestra of sound waves? Perhaps the universe is literally teeming with consciousness — with content flickering in and out, connecting through memory, separating, overlapping, flowing, in ways we can’t quite imagine — ruled by physical laws we don’t yet understand.

## References

- Chalmers, D.J. (2012) The combination problem for panpsychism, [Online], <http://consc.net/papers/combination.pdf>.
- Goff, P. (2001) Panpsychism, in Zalta, E.N. (ed.) *Stanford Encyclopedia of Philosophy*, [Online], <https://plato.stanford.edu/entries/panpsychism/>.
- Goff, P. (2019) *Galileo’s Error*, London: Rider.
- Harris, A. (2019) *Conscious*, New York: Harper.
- Musser, G. (2015) *Spooky Action at a Distance*, New York: Scientific American/FSG.
- Rosenberg, G. (2004) *A Place for Consciousness*, Cambridge: Cambridge University Press.
- Seth, A. (2021) *Being You*, New York: Dutton.
- Skrbina, D. (2017) *Panpsychism in the West*, Cambridge, MA: MIT Press.