

The FAQs of Consciousness

INTRODUCTORY QUESTIONS

1. What would a good definition of consciousness look like?

First off, this is not going to be a simple thing. In the 16th post in this series, I went through [a \(sorta\) brief history the definitions of consciousness](#). As I noted there, these have ranged “all the way from it being something as small as the private, ineffable, special feeling that only we rational humans have when we think about our thinking, right on down to it being a fundamental force of the universe that gives proto-feelings to an electron of what it’s like to be that electron.” The Stanford Encyclopedia of Philosophy entry on [consciousness](#) stated, “There is unlikely to be any single theoretical perspective that suffices for explaining all the features of consciousness that we wish to understand. Thus, a synthetic and pluralistic approach may provide the best road to future progress.” And as Dan Dennett noted in [one of my favourite papers](#), among many philosophers, their typical “demand for essences with sharp boundaries blinds thinkers to the prospect of gradualist theories of complex phenomena, such as life, intentions, natural selection itself, moral responsibility, and consciousness.”

So, I believe it’s clear we ought to be looking for a gradualist theory of the emergence of all the complex phenomena associated with consciousness. The famed evolutionary biologist Ernst Mayr who coined the proximate/ultimate distinction [claimed](#) that “all problems of biology, particularly those relating to emergence, are ultimately problems of hierarchical organization.” Thus, trying to reduce consciousness to a single thing is impossible, but that still leaves open the possibility of analysis. What kind of analysis? The philosopher of science Robert Cummins gave [the canonical account](#) of a *functional analysis*, which “consists in breaking down some capacity or disposition of interest into simpler dispositions or capacities, organized in a particular way.”

Therefore, we need a deep functional analysis of consciousness where elements of that emerging property are listed out for separate consideration. In this way, nuances can be captured and lassoed into an evolving understanding of all the issues.

2. What’s your definition?

That’s a tough question, but I wrote a full summary of my response in [post 23](#) of this series. In that post, I started with the background of my metaphysical hypotheses, which is just standard naturalism. Then, I laid out the theories that I like best for the two biggest mysteries for this topic — 1) the emergence of life, which I accept as happening using something like the RNA-world hypothesis, and 2) the hard problem of consciousness, which I think is most simply explained using my hypothesis of *pandynamism* (see question 4 below for details). Once chemistry makes the jump to biology, then the resulting proto-lifeforms have a defined self *and* they begin to compete for resources with other potential entrants, substitutes, or conspecifics in order to self-replicate and survive. They react to the world as if they know what they are *and* what they need. These are the building blocks for expanding the properties of subjective experience. Thus:

Consciousness, according to this evolutionary theory, is an infinitesimally growing ability to sense and respond to any or all biological forces in order to meet the needs of survival. These forces and needs can vary from the immediate present to infinite timelines and affect anything from the smallest individual to the broadest concerns (both real and imagined) for all of life.

This is intended to be a comprehensive and therefore very broad definition. Anything that is able to act to remain alive does so using aspects of consciousness. There are infinite varieties of scope and scale within this definition, so in order to map these contours I spent several posts conducting a [Tinbergen analysis](#) of the [functions](#), [mechanisms](#), [ontogeny](#), and [phylogeny](#) of consciousness. This is the standard procedure in evolutionary studies for coming to know all of the elements of any biological phenomenon, and I believe it is therefore the best method to perform a *functional analysis* as described above in question 1. The hierarchical organisation that emerged from this review is supported by logical requirements as well as empirical data from across the history of all life. That hierarchy is:

- 1) Origin of Life
- 2) Affect
- 3) Intention
- 4) Prediction
- 5) Awareness
- 6) Abstraction

In order to further elaborate this definition of consciousness, I finished my summary post by providing definitions of the following common terms in consciousness studies, which sometimes differ between technical and folk usages:

- Accessible, Attention, Bottom-up vs. Top-down, Cognition, Communication, Conscious vs. Unconscious, Emotions vs. Feelings, Evolutionary Hierarchy of Needs, Evolutionary Epistemology Mechanisms, Exteroception vs. Interoception, Intentionality vs. Intentional Stance, Involuntary vs. Voluntary, Language, Mind, Qualia vs. Something-it-is-like vs. Subjective Experience.

All of this is in keeping with the epigraph for *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness* by Peter Godfrey-Smith. That epigraph came from William James in *The Principles of Psychology* from 1890:

- “The demand for continuity has, over large tracts of science, proved itself to possess true prophetic power. We ought therefore ourselves sincerely to try every possible mode of conceiving the dawn of consciousness so that it may not appear equivalent to the irruption into the universe of a new nature, non-existent until then.”

Godfrey-Smith [has been interpreted](#) as saying that “we need a theory [of consciousness] based on continuities and comprehensible transitions; no sudden entrances or jumps.” This, of course, aligns with Darwin’s observation that nature does not jump, and I believe my theories and definitions of consciousness fit this requirement.

QUESTIONS FROM IMPARTIAL SOURCES

Now that I've laid out what a theory of consciousness should look like and what my particular theory is, let's see how that addresses the standard objections raised against other theories of consciousness. Particularly, these questions come from the Stanford Encyclopedia of Philosophy entry on [Naturalism](#) by David Papineau and the even more pointed Internet Encyclopedia of Philosophy entry on [Consciousness](#) by Rocco Gennaro.

3. Why do we think consciousness is a physical phenomenon?

According to [Papineau](#), a majority of contemporary philosophers hold that physicalism will be able to explain consciousness, although a significant minority take two other options. The first is that “conscious properties are ‘epiphenomenal’ and do not exert any influence on brain processes or subsequent behaviour.” The second route is “to embrace the ‘overdeterminationist’ view that the physical results of conscious causes are always strongly overdetermined” by both physical causes and by some other immaterial causes. Papineau declares that neither of these two positions are attractive. He says that they “posit odd causal structures,” neither of which are observed anywhere else in nature, so we're not compelled to accept them here.

In my [summary post on consciousness](#), I provide a more positive example of why the physicalist explanation for consciousness is more likely to be correct. I wrote:

- The psyche only originates and evolves along with life. This psyche expands as the living structures expand their capabilities of sensing and responding to [biological] forces. And the ‘flavour’ of experiences within this psyche are utterly dependent upon the underlying mechanisms of *which* particles of matter are being subject to *which* particular forces.
- For example, the retch of disgust from accidentally eating something harmful maps almost exactly onto the retch of moral disgust from accidentally witnessing something beyond the pale such as a mutilated dead body. These experiences come from very different sources, and they process very different bits of information, so we might expect them to feel very different, but we know from neuroscience that the brain has duct-taped the feelings of moral disgust onto the existing architecture for gustatory disgust and that is what explains the similar conscious experience. This is another striking bit of support for a materialist understanding of consciousness.

4. How could minds possibly arise from matter?

[Gennaro](#) lists this as the first standard objection to physicalist accounts of consciousness. It usually goes by one of two names. Joseph Levine (1983) coined the expression ‘the explanatory gap’ as a label for the idea that there is a key gap in our ability to explain the connection between subjective feelings (mind) and brain properties (matter). David Chalmers (1995) described something similar with the catchy phrase ‘the hard problem of consciousness’, which has come to dominate this discussion. (See an in-depth examination of this in questions 27 to 41.)

This is indeed a problem for the whole project of evolutionary explorations of consciousness. In a paper called “[The Difficulty of Fitting Consciousness in an Evolutionary Framework](#)”, the author Yoram Gutfreund noted that “the question of how the mind emerged in evolution (the mind-evolution problem) is tightly linked with the question of how the mind emerges from the brain (the mind-body problem). It seems that the evolution of consciousness cannot be resolved without first solving the ‘hard problem’. Until then, I argue that strong claims about the evolution of consciousness based on the evolution of cognition are premature and unfalsifiable.”

In my [post 19 on the functions of consciousness](#), I introduced my hypothesis for a solution to this. I wrote:

- The hard problem of consciousness is often phrased as wondering how inert matter can ever evolve into the subjective experience that we humans undoubtedly feel. I think this short-changes matter. Far from being inert, matter responds to the forces exerted on it all the time. Panpsychism says mind (*psyche*) is everywhere. But to me there can be no mind without a stable subject. In my current conception, the forces that minds feel and are shaped by are merely the chemical and physical forces that shape all matter. Until something else is found, what else could there be? So, mind is not everywhere, but forces are. The Greek for force is *dynami*, so rather than panpsychism, I would say the universe has *pandynamism*. The psyche only originates and evolves along with life.

In my [summary post on consciousness](#), I further explained this when I wrote:

- We have subjective experience. Evolutionary studies have shown us that there is an unbroken line in the history of life. But water and rocks don't appear to have anything like consciousness. So, how can inert matter ever evolve into the subjective experience that we humans undoubtedly feel? Chalmers has proposed that subjective experience may be a fundamental property of the universe, like the spin of electromagnetism. I have come to accept that as a likely hypothesis. All matter is affected by the forces of physics and chemistry. But until that matter is organised into a living subject that is capable of responding to those forces in such a way as to remain alive, it makes no sense to talk of non-living matter as ‘feeling’ or ‘experiencing’ those forces. Inert matter has no structure capable of living through subjective activities. Panpsychism claims that minds (*psyche*) are everywhere, and they don't need physics and matter to exist. But this raises innumerable difficulties, including an enormous change to one's metaphysics that supposedly cannot be detected by science. What I hypothesise instead is that the forces of physics are everywhere, and it is a fundamental property of the universe that these forces are felt subjectively when subjects emerge. Since the Greek for force is *dynami*, I would say the universe has *pandynamism* rather than panpsychism. The psyche only originates and evolves along with life.

As an example, take a very simple force. What does it take to ‘feel’ gravity? For us humans, it's registering the difference between inner ear liquids as our movements in space accelerate or decelerate. Can a rock or a photon ever experience this? No. Why not? Because there is no structure in its makeup by which it could gain such information. Panpsychism is therefore a non-starter for me, but *pandynamism* could explain how subjectivity is a fundamental feature of the universe, yet only emerges as living organisms emerge, thus bridging the explanatory gap and providing a coherent answer to the hard problem.

Is this enough to overthrow all doubts from metaphysical dualists? Not likely. But Patricia Churchland provided a wonderful quote about this in her essay “[Neurophilosophy](#)”, which was a chapter in the fantastic edited collection [How Biology Shapes Philosophy: New Foundations for Naturalism](#). She wrote:

- A methodological point may be pertinent in regard to the dualist’s argument: however large and systematic the mass of empirical evidence supporting the empirical hypothesis that consciousness is a brain function, it is always a logically consistent option to be stubborn and to insist otherwise, as do Chalmers and Nagel. Here is the way to think about this: identities—such as that temperature really is mean molecular kinetic energy, for example—are not directly observable. They are underwritten by inferences that best account for the mass of data and the appreciation that no explanatory competitor is as successful. One could, if determined, dig one’s heels in and say, “temperature is not mean molecular kinetic energy, but rather an occult phenomenon that merely runs parallel to KE.” It is a logically consistent position, even if it is not a reasonable position.

Thus, I believe conscious subjectivity appears to be another one of these identities of the universe rather than some occult phenomenon requiring an entirely new metaphysical realm.

5. Does consciousness contain non-physical information?

This is the second common objection according to [Gennaro](#) and it is usually labelled *The Knowledge Argument*. This is based on “a pair of very widely discussed, and arguably related, objections to materialism which come from the seminal writings of Thomas Nagel (1974) and Frank Jackson (1982, 1986). ... The general pattern of each argument is to assume that all the physical facts are known about some conscious mind or conscious experience. Yet, the argument goes, not all is known about the mind or experience. It is then inferred that the missing knowledge is non-physical in some sense, which is surely an anti-materialist conclusion in some sense.”

Luckily, I’ve already written about these arguments during [my series on 100 philosophy thought experiments](#). When I tackled Jackson’s thought experiment in [my physicalist response to Mary’s Knowledge Problem](#), I wrote:

- In logical form, the argument goes something like this:
 - (1) Mary has all the physical information concerning human color vision before her release.
 - (2) But there is some information about human color vision that she does not have before her release.Therefore
 - (3) Not all information is physical information.

Hogwash! The first premise is patently false because Mary does not have “all the physical information” and cannot know “all there is to know” about this subject without having experienced it first-hand. Why? Precisely because we live in a physical universe where mental imaginings are not enough to move the physical atoms that make up the nerves in

our eyes and the synapses in our brains. In philosophical terms, there is a real epistemic barrier to what we can learn no matter how much we sit in our rooms and read and think.

Later, when I tackled the thought experiment about [what it is like to be a bat](#), I wrote:

- If our epistemological stance is that knowledge can only ever come after sensory experience, then of course it would be impossible to know what it is like to be a bat because we do not share the sensory experiences of a bat. Nagel may have realised this, but he ducked the question. Buried in footnote number 8 in his [original paper](#), there is this:

“My point, however, is not that we cannot know what it is like to be a bat. I am not raising that epistemological problem. My point is rather that even to form a conception of what it is like to be a bat...one must take up the bat's point of view.”

But that's exactly the problem! The epistemological problem Nagel didn't want to raise explains the entire difficulty that his mind-body thought experiment supposedly raises. ... So, to me, the fact that we can't know what it feels like to be a bat is actually an argument that *bolsters* physicalism, rather than questions it.

Read the entire posts for those thought experiments to dig in more deeply, but [Gennaro](#) clearly agreed with me when he wrote, “Indeed, a materialist might even expect the conclusion that Nagel draws; after all, given that our brains are so different from bat brains, it almost seems natural for there to be certain aspects of bat experience that we could never fully comprehend. Only the bat actually undergoes the relevant brain processes. Similarly, Jackson’s argument doesn’t show that Mary’s color experience is distinct from her brain processes.”

6. And what about Hume’s missing shade of blue?

While we’re at it. There is one more famous thought experiment that may undermine physicalism and is closely related to consciousness. This is [Hume’s Missing Shade of Blue](#), which I also wrote about. This didn’t make the standard objections list, but let’s cover it quickly here before carrying on. These are the relevant snippets from my post:

- Hume argued “that all perceptions of the mind can be classed as either ‘Impressions’ or ‘Ideas’.” He further argues that: “*We shall always find that every idea which we examine is copied from a similar impression. Those who would assert, that this position is not universally true nor without exception, have only one, and at that an easy method of refuting it; by producing that idea, which, in their opinion, is not derived from this source.*”
- Just two paragraphs later though, Hume seems to provide just such a destructive idea that arises without a sense impression. He says:

“There is, however, one contradictory phenomenon, which may prove, that it is not absolutely impossible for ideas to arise, independent of their correspondent impressions. I believe it will readily be allowed that the several distinct ideas of colour, which enter by the eye...are really different from each other; though, at the same time, resembling. Now if this be true of different colours, it must be no less so of the different shades of the same colour; and each shade produces a distinct idea, independent of the rest. ... Suppose, therefore,

a person to have enjoyed his sight for thirty years, and to have become perfectly acquainted with colours of all kinds, except one particular shade of blue, for instance, which it never has been his fortune to meet with. Let all the different shades of that colour, except that single one, be placed before him, descending gradually from the deepest to the lightest; it is plain, that he will perceive a blank, where that shade is wanting, and will be sensible, that there is a greater distance in that place between the contiguous colours than in any other. Now I ask, whether it be possible for him, from his own imagination, to supply this deficiency, and raise up to himself the idea of that particular shade, though it had never been conveyed to him by his senses? I believe there are few but will be of opinion that he can: And this may serve as a proof, that the simple ideas are not always, in every instance, derived from the correspondent impressions; though this instance is so singular, that it is scarcely worth our observing, and does not merit, that for it alone we should alter our general maxim.”

- So, despite Hume's uncharacteristic dismissal of such a singular instance, “scarcely worth observing,” this stubborn little problem seems to undermine the whole underpinnings of empiricism and physicalism. And that's a really big deal!
- [After a thorough investigation of how our eyes and vision systems work, which you should read in depth if you are interested, I said that] now that we've got a consistent view of the problem across the disciplines of biology and philosophy, we understand how we can imagine particular shades of blue even if we haven't seen them yet. Physically, it's simply a matter of how excited our blue cones have been in the past. We may not be able to “know” what peaks on those cones might look like without seeing them, but we can easily imagine points in between levels of excitement we have seen. This is simply analogous to imagining what a 5 kg weight dropped on my toe would feel like once I have had a 2 kg and 10 kg weight dropped on it. We can fill in the gaps rather easily. Similarly, I might not “know” what a 200 kg weight dropped on my toe would feel like, but I could roughly extend my imagination to it once I have some experience in the matter.

So, once again, the mind is built from physical experiences and no exceptions have been found to refute that hypothesis.

7. Is consciousness so mysterious that it is beyond our ability to understand it?

This is the third standard objection noted by [Gennaro](#). In short, “[mysterians](#)” believe that the hard problem of consciousness can never be solved because of cognitive limitations we humans face. Colin McGinn is the leading proponent of this idea and has suggested we may be in the same situation with consciousness as a rat or dog is with respect to calculus. McGinn also notes that we “access consciousness through introspection or the first-person perspective, but our access to the brain is through the use of outer spatial senses (e.g., vision) or a more third-person perspective. Thus, we have no way to access both the brain and consciousness together, and therefore any explanatory link between them is forever beyond our reach.”

Gennaro notes that materialist responses to this are numerous. Rats have no concept of calculus whatsoever, so of course they cannot solve its problems. We humans, however, know a great deal about consciousness. Gennaro even quipped, “just see the references at the end of [this entry](#)!” We are clearly not in an analogous position with the ignorance of rats. And while we must acknowledge there are epistemological barriers to what any one person can know about their brains or the consciousness of others, we can “combine the two perspectives within certain experimental contexts. Both first-person and third-person scientific data about

the brain and consciousness can be acquired and used to solve the hard problem.” Scientists do this all the time.

More generally, my analysis of the evolution of consciousness places the ability for abstraction at the highest level of its hierarchy. Once the capabilities of this level are reached—and then expanded using the tools of language, writing, and symbol manipulation, which can be arranged in an infinite number of possibilities, and stored and analysed using powerful computers—it becomes very hard to see what, if anything, could limit the conceptualisations of such a consciousness. Certainty or indisputable proof for our theories of consciousness may be out of reach, but that is the case for all of our [knowledge](#). We still get by with pragmatic hypotheses that prove to be extremely robust.

To read more about how Dan Dennett finds mysterianism an embarrassment for philosophy, read [his short review of one of Colin McGinn’s books in the *Times Literary Supplement*](#).

8. What about Zombies?

We’re really grasping at straws now. The fourth and last of the standard objections listed by [Gennaro](#) is the problem of zombies. Supposedly, these are “creatures which are physically indistinguishable from us but lack consciousness entirely. ... The appeal to the possibility of zombies is often taken as both a problem for materialism and as a more positive argument for some form of dualism, such as property dualism.”

I’ve written about this argument in several places now. First, in [my response to a thought experiment about zombies](#), I focused on the poor logic in the zombie argument. The claim that zombies *may* be possible is supposed to prove that physicalism *is* false. But that’s a flawed leap. It only proves that physicalism *may* be false, and thus may also be true. I also noted there how Richard Brown’s “[zoombies](#)” (which are conceivable beings that are identical to humans in the *non-physical* realm but have *no* consciousness, therefore implying that consciousness *must* be physical) shows that zombie arguments are circular and could just as easily be constructed *against* dualism.

In this series on consciousness, I also covered a David Chalmers interview about [the Hard Problem](#) where he discusses his idea of zombies. And in my post covering all [the definitions of consciousness](#), I traced the history of the idea and some prominent responses. Finally, in my post on [the functions of consciousness](#), I focused quite a bit on Todd Moody’s “[Zombie Earth](#)” and Dan Dennett’s paper about [the unimagined preposterousness of zombies](#), which both show just how untenable the idea really is. If zombies were truly “unconscious but indistinguishable from us,” then they would display fear of upcoming public speaking events or be just as engrossed in sexual fantasies or show any number of other hallmarks of internal thought processing. They would even create and speak words in their language that describe these internal states. The fact that we think unconscious creatures couldn’t do these things blocks the intuition that zombies are a possibility that we need to concern ourselves with. So, let’s not.

Zombie proponents [Flanagan and Polger](#) thought these experiments “highlight the need to explain why consciousness evolved and what function(s) it serves. This is the hardest

problem in consciousness studies.” I agree it’s hard, but fortunately that’s what the rest of this series and these FAQs have helped to uncover.

9. How is our conscious experience bound together?

Moving on from the standard objections of philosophers, [Gennaro](#) next notes some prominent scientific holes that need to be filled. The first one listed is known as [the binding problem](#) and it relates to the unity of consciousness. In a nutshell, “How does the brain ‘bind together’ various sensory inputs to produce a unified subjective experience?”

This is a very difficult question to answer because examinations of the brain show there isn’t any one spot that could possibly act as the unifying mechanism. From an evolutionary standpoint, that’s exactly what one would expect to see in nervous systems and brains that have been built up incrementally over eons of time in lots of starts and stops down various paths of trials and errors. The philosopher Jonathan Birch even has varying degrees of unity as one of the variables in his “[Dimensions of Animal Consciousness](#).” In the summary of that fascinating paper, there is this observation:

- “For example, neuroanatomical considerations suggest that conscious experience in mammals (which have a corpus callosum) may be more highly unified than in birds (which do not) and that experience in birds may be more highly unified than in cephalopods.”

Untangling all the brain structures in the animal kingdom is taking consciousness researchers decades. And reading up on this subject quickly dives into details of brain mechanisms like v1 regions of the visual cortex, gamma-band oscillations synchronized around 40 Hz for various neuronal signals, and electromagnetic fields generated by neuronal firing. (See [here](#) and [here](#) for plenty of details like this.) It all ends up with the current state where, “There are a wide range of views on just how real this ‘unity’ is” and “the nature of, and solution to, [the binding problem] remains a matter of controversy.”

This is all okay for me and my philosophical theory of consciousness. I’m happy to wait and see how these mechanisms are mapped out. So far, the progress being made suggests that a physical solution will be found. In fact, a promising one was just discussed in April 2021 on the Brain Science Podcast with Ginger Campbell when [she interviewed Jeff Hawkins](#) about his new book [A Thousand Brains: A New Theory of Intelligence](#). The whole podcast is worth listening to, but here’s the transcript of the 4-minute clip that specifically addresses the binding problem.

- **[Ginger Campbell at 29:33]** And you automatically solve the binding problem?
- **[Jeff Hawkins]** Yes! I didn’t know if you wanted to go there or not but that’s okay. So, there’s a thing called the binding problem that’s poorly defined because people interpret it differently. You can think of it as the following. The brain has all these different sensors. Your eye, your retina, is not really one sensor; it’s thousands of sensors aligned with each other, just like your skin has thousands and thousands of sensors along your skin. Your ear has the cochlea, and it has thousands and thousands of individual sensors in there. So, you halve all this information streaming into the brain. They all have to be processed separately. All this stuff is going on, but we have this singular perception of the world. We don’t have the feeling that I’m hearing something and I’m seeing something. Your not

aware of all this complicated stuff going on in your head. You just look out into the world and say “there it is. I’m looking at something and I know what it is and what it’s supposed to feel like; I know what it’s supposed to sound like.” The question is, where does all this information get brought together in the brain? Where does it get bound together into our singular percept? If you look at the brain, you don’t see that. You don’t see everything going into one spot, which is like “that’s you.” We see connections going all over the place. There doesn’t seem to be any centralised anything. How could that be? Well, our theory, which I would be remiss in not mentioning that it is called *the thousand brains theory*, reflects the fact that you have these tens of thousands of models in your neocortex. The thousand brains theory says you have all these independent models. They’re each modelling a part of the world that they can see. And they don’t actually come together. But what they do, and we haven’t talked about this yet, is they vote. So, most of the long-range connections in the neocortex that go all over the place—from one side to the other, up and down, just all over the place, just everywhere—form connections connecting different parts of the neocortex together. We believe they’re *voting*. The different columns say things like, “I’m a touch column. I’m representing my finger’s input. I think I’m touching a coffee cup. But I’m not certain about it.” Another column in the visual column says, “well I’m looking at an edge in the scene out here and I’m trying to model it but I’m not sure if it’s a coffee cup or it could be a chair.” All these columns are not certain of what they’re looking at, but they have information, and they can vote! These long-range connections really try to reach a common consensus which is consistent with what they are all experiencing. This makes it so that all of a sudden everyone goes, “Yep! We’re all agreeing that this thing is a coffee cup, or a computer, or a bird.” So, the binding doesn’t occur in one spot. It’s essentially a voting mechanism that occurs across the brain and our perceptions are primarily of that voting. We’re not aware of all the thousands of models that are guessing what is going on in the world. But we are aware of their consensus. And the consensus says, “yes, we all agree that this is something” and I can then drill down and say, well what does that look like, what does that sound like, what does that feel like. But we all agree that it’s this bird or whatever. And so, this solves the binding problem by not binding it into one spot but by voting and reaching consensus. And so therefore we don’t have to look for a spot in the brain where everything comes together.

- [GC] This also makes sense of the fact that most of what the cortex does is not conscious.
- [JH] Yes! We’re almost totally unaware of most of what is going on in there. All the tiny inputs are constantly changing, but the consensus voting stays the same and that allows for continual experience. **[Clip ends at 33:50]**

This sounds very promising as it’s easy to see how it would be built up gradually over time, bringing more and more representational voting into the overall picture. But for now, let’s wait for the scientific method to play out before declaring any firm answers to this question.

10. What can the neural correlates of consciousness tell us?

The other major hole in our scientific understanding of consciousness that [Gennaro](#) discusses is the program to find the [neural correlates of consciousness](#) (NCCs). This project is based on the idea that consciousness originates in the brain, and “some credit for it must go to the ground-breaking 1986 book by Patricia Churchland entitled *Neurophilosophy*.” In the paper “[What is a Neural Correlate of Consciousness?](#)”, David Chalmers answers that title question thusly: “At first glance, the answer might seem to be so obvious that

the question is hardly worth asking. An NCC is just a neural state that directly correlates with a conscious state.” He goes on to elaborate, however, that, “A number of proposals have been put forward concerning the nature and location of neural correlates of consciousness. A few of these include:

- 40-hertz oscillations in the cerebral cortex (Crick and Koch 1990)
- Intralaminar nuclei in the thalamus (Bogen 1995)
- Re-entrant loops in thalamocortical systems (Edelman 1989)
- 40-hertz rhythmic activity in thalamocortical systems (Llinas et al 1994)
- Extended reticular-thalamic activation system (Newman and Baars 1993)
- Neural assemblies bound by NMDA (Flohr 1995)
- Certain neurochemical levels of activation (Hobson 1997)
- Certain neurons in inferior temporal cortex (Sheinberg and Logothetis 1997)
- Neurons in extrastriate visual cortex projecting to prefrontal areas (Crick and Koch 1995)
- Visual processing within the ventral stream (Milner and Goodale 1995)

(A longer list can be found in Chalmers 1998.)”

Looking at this list, you can readily understand why Gennaro said, “a detailed survey would be impossible to give here” and I would not attempt such a thing either. I’m happy to let the neuroscience play out for years to come as it maps what I think of as **the mechanisms of consciousness**, which is just one of **Tinbergen’s four questions** about any biological phenomenon. In the meantime, Chalmers’ dense list of paths for this exploration serves to highlight the two main meta-problems with this project that Gennaro notes.

First:

- “One problem with some of the above candidates is determining exactly how they are related to consciousness. For example, although a case can be made that some of them are necessary for conscious mentality, it is unclear that they are sufficient. That is, some of the above seem to occur unconsciously as well. And pinning down a narrow enough necessary condition is not as easy as it might seem.”

I think this problem of searching for a narrow condition comes from having too narrow a definition of consciousness. Researchers seem to be focused merely on conscious awareness, which comes in at level 5 in my hierarchy, and only arrived in biological life after the other levels below it were established. Such emergence never comes from a clear-cut break in evolution, so pinning down exact NCCs for that second C of “consciousness” may be a fool’s errand. As detailed above in question 1, a *functional analysis* will be required which “consists in breaking down some capacity or disposition of interest into simpler dispositions or capacities, organized in a particular way.” There just won’t be one simple answer.

Second:

- “Another general worry is with the very use of the term ‘correlate.’ ... Even if such a correlation can be established, we cannot automatically conclude that there is an identity relation. Perhaps A causes B or B causes A, and that’s why we find the correlation. Even most dualists can accept such interpretations. Maybe there is some other neural process C

which causes both A and B. ‘Correlation’ is not even the same as ‘cause,’ let alone enough to establish ‘identity.’”

This is the same problem that Patricia Churchland answered for us above in question 4. I’ll just repeat the relevant part of the quote here from her paper “[Neurophilosophy](#)”:

- Here is the way to think about this: identities—such as that temperature really is mean molecular kinetic energy, for example—are not directly observable. They are underwritten by inferences that best account for the mass of data and the appreciation that no explanatory competitor is as successful. One could, if determined, dig one’s heels in and say, “temperature is not mean molecular kinetic energy, but rather an occult phenomenon that merely runs parallel to KE.” It is a logically consistent position, even if it is not a reasonable position.

So, the results of the NCC project will have their limits, but since they are not ruling out physicalism, that hypothesis continues to hold up with all of the evidence in the universe that has ever been gathered and tested.

11. Are other animals conscious?

[Gennaro](#) starts with the obvious (to me) concession that “in the aftermath of the Darwinian revolution, it would seem that materialism is on even stronger ground provided that one accepts basic evolutionary theory and the notion that most animals are conscious.” But then he notes there is still much discussion around the question, “To what extent are animal minds different from human minds?” Well, according to my definition, *all* living beings do indeed have *some* levels of consciousness, and I can use my comprehensive hierarchy as a guide to describe how much and of what kinds. These descriptions of a being’s consciousness vary widely across all species, across individuals within a single species, and across the lifespan of individuals too.

An important outcome from this is to not think of consciousness as a single variable or an on-off switch. The philosopher Jonathan Birch has published an excellent example of this in his 2020 paper about the “[Dimensions of Animal Consciousness](#)” where he uses a [radar chart](#) (aka spider web chart) to illustrate what five dimensions might look for elephants, corvids, and cephalopods. Birch, however, recognises that this is just a starting example to get people thinking in the right way. Among the key challenges he discusses for mapping dimensions of consciousness, he says that “One is to find dimensions at the right grain of analysis. If our goal were to capture all interesting variation in conscious states, we would never have enough dimensions. We have to be pragmatic.” I agree, although I probably would have started with the 13 types of cognition listed in [Pamela Lyon’s paper](#) on the evolution of cognition (which I placed in my hierarchy when I mapped [the functions of consciousness](#)). That’s a bit more difficult to plot, though, and Birch isn’t trying to be comprehensive. I, however, do want my hierarchy to be comprehensive, so let’s see how Birch’s dimensions might be covered within my hierarchy.

1. *E-richness* (where the e stands for evaluative) is roughly equivalent to the cognition of *valance* within my level of affect, but it also looks at *motivation* according to Birch’s chart of experiments for each of his dimensions.

2. *P-richness* (where the p stands for perceptual) is equivalent to the cognition of *sense perception* that sits within my level of affect. I see p-richness and e-richness going hand in hand because one must perceive something in order to evaluate it, and living beings evaluate everything they perceive (as positive, negative, or neutral). This is why I have them on the same level in my hierarchy. Birch is right, though, that they can change in independent directions from one another.
3. *Unity* or *integration at a time* relates to the binding problem noted above in question 9. This is an interesting dimension which Birch explores with examples such as humans with split-brain syndromes, dolphins and seals sleeping with one hemisphere at a time, and the fact that birds have no structure akin to a corpus callosum. He wonders, “Could there be two subjects within one skull?” This will come up again when I discuss the nesting problem below in question 17, but for now, I see the unity dimension as a way of looking at how a few of the cognitions in my intention level actually combine together. Just how intentional can one animal (or one consciousness!) act using the *attention*, *memory*, *pattern recognition*, and *learning* that it has at its disposal. Each of sub-categories can obviously vary from one another, so I consider Birch’s *unity* as a meta-variable examining how these are combined.
4. *Temporality* or *integration across time* is another complex meta-variable to me. This one looks at just how short or long of a timespan can be considered to affect the conscious experiences and thoughts of animals. This integrates across several cognitions in my hierarchy — *sense perception* and *discrimination* in my level of affect, *memory* and *pattern recognition* in my level of intention, *anticipation* and *error detection* in my level of prediction, the *self-reference* in my level of awareness that creates the autobiographical self, and even the ability in my final level for *abstraction* to use symbols and language to help extend thinking into the distant past or future. So, I could probably make another interesting radar chart for this single variable in Birch’s dimensions.
5. *Selfhood*, according to Birch, is “the conscious awareness of oneself as distinct from the world outside.” This is equivalent to the cognitive ability for *self-reference* in my level of awareness.

So, all of Birch’s dimensions can indeed be mapped onto my hierarchy, but is there anything of mine that he’s left out? The only cognitive abilities I have listed which I don’t think he covers are the ones for *communication* and *problem solving*. These both seem to be interesting abilities that can vary widely across different individuals and different species, so perhaps they too could make for useful considerations during future analyses of animal consciousness.

12. Can machines be conscious?

The short answer is yes. According to my definition, machine consciousness is possible, although it would certainly feel different than ours. (Think how much our own consciousness changes under the influence of a few chemicals and just imagine what an entirely different substrate might cause.) In order to describe any machine’s consciousness accurately, we would need the same kind of comprehensive functional analysis as described above, which would map all of the dimensions throughout my hierarchy.

To explore this in more detail, let’s consider three more questions that [Gennaro](#) raised in his IEP article on consciousness.

1. Could an appropriately programmed machine be conscious? Yes. In a material universe without souls imbued by gods, it's hard to see why not. My theory of *pandynamism* acknowledges that all matter feels forces, but minds arise when subjects emerge. An appropriately programmed machine could conceivably recreate the conditions for a living subject, which would then feel its physical changes.

2. Could a robot really subjectively experience the smelling of a rose or the feeling of pain? Once again, yes, but only if the above conditions are satisfied. You have to have a subject before what we call subjectivity can enter into it. Cameron Harwick's long article on "[What Computer-Generated Language Tells Us About Our Own Ideological Thinking](#)" makes an important point about this. Harwick states:

- Thus, the ancient question of what separates humans from animals is the inverse of the more recent question of what separates humans from computers. With **GPT**, computers have finally worked backward (as seen in animal terms), from explicit symbol manipulation to a practically fluent generative language faculty. The result might be thought of as a human shell, missing its animal core.

This is exactly right. And that "animal core" is my hierarchy level of affect, which is what Mark Solms **calls** the hidden spring or source of consciousness. Without this innate, evolved, built-in sense of judging what is good or bad or indifferent for a self, there is no way that the sting of pain or the sweet smell of a rose can make sense. Could that be programmed into a robot or machine? Yes. But with some interesting differences worth discussing in the next question.

3. How and when does one distinguish mere 'simulation' of some mental activity from genuine 'duplication'? This question is in reference to John Searle's famous Chinese Room. I have summarised **my response to this thought experiment** by saying "emotions, definitions for good and bad, and the ability to learn to meet a hierarchy of needs are probably enough to create strong artificial intelligence. They are all we have ourselves." So, creating artificial subjects by simulating our own interactions with the world seems entirely possible, although that wouldn't *duplicate* our conscious experience of these interactions. Does that matter? Do the forces felt from the movement of my sodium-channel ions matter any more or less than the forces felt from the movement of a different set of chemicals? I'm not a bio-chauvinist so I don't see why that makes a difference morally, even if there is a difference in the raw feelings. So, duplication isn't the goal to me. Searle's Chinese Room is meant to pump the intuition that simulation of a function isn't enough to matter because "clearly" the man in the Chinese Room (or the Chinese Room as a whole system) isn't having the same experience as an individual human speaking Chinese. But that's because of all of the other activities that are also wired into our own speaking systems. If you could somehow remove all of the memories from a person, and all of the living, emotional, and other sense systems as well, but miraculously keep the auditory and speaking systems going all by themselves, would there be any "consciousness" there? Not in the way that Searle meant. Such a listening and speaking slice of a human would be just as dumb as a Chinese Room.

Here's another way to approach this issue. Let's say you programmed a computer to speak "ouch" when its vibration sensors moved too vigorously. That is *simulating* pain, but we don't think it is duplicating *our* pain. In a really sophisticated robot, would that pain matter? I don't see why not, if such a robot were programmed to be aware of its surroundings and able to

learn from them, while also striving towards open-ended goals, and simply becoming irreplaceable because of its unique prior experiences and potential for even more. And yet, the exact same physical inputs in such a computer could be easily tinkered with and re-programmed to say “yum” or “blue hippopotamus” when it was shaken, which would render its conscious simulation utterly nonsensical. There just isn’t the kind of singular match between changes in the world and felt states inside the computer that would persuade us to consider it conscious in the same way that we are. Such a computer could conceivably be constructed with a kind of consciousness that we care about, but it would be extremely fragile and fluid compared to our own. It would be subject to the whims of its programmers. Perhaps, however, we may one day learn the chemical coding that drives our own bodies to the point that we are as fluent in that as we are now in computer coding. Were such editing of our own biological codes to become so possible, our own consciousness could become just as fluid and changeable as the computer’s. Would that erase our own consciousness? I think not. It would just change what else we need to include in order to describe it.

13. So, “what is it like” to be conscious?

Question 12 was the end of the standard objections and scientific holes in the IEP entry from [Gennaro](#) about consciousness, but there were three more things he touched on in his brief introduction that I thought were worth a quick discussion. This first one is of course a reference to Thomas Nagel’s famous “what is it like” description of consciousness, which Gennaro called “perhaps the most commonly used contemporary notion of a conscious mental state.” In Gennaro’s retelling of this, “When I am in a conscious mental state, there is something it is like for me to be in that state from the subjective or first-person point of view. But how are we to understand this?”

One problem with this is that it is too narrow a search to provide much understanding. As I mentioned above in question 10 about the search for the NCCs,

- “I think this problem of searching for a narrow condition comes from having too narrow a definition of consciousness. Researchers seem to be focused merely on conscious awareness, which comes at level 5 in my hierarchy, and only arrived in biological life after the other levels below it were established.”

This “what it is like” feeling that Nagel is describing disappears whenever we are rendered unconscious, and yet much of consciousness’ processing still goes on to keep us alive (as it does when we are awake as well). In order to fully understand “what it is like”, we have to look at the long history of emergence that got to that kind of on-again / off-again state. In my post on [the evolutionary history \(aka phylogeny\) of consciousness](#), we can see the possibility that *awareness* of “what it is like” may go back a very long way. It is best tested using mirror recognition tests, which several non-human species have passed including mammals, birds, and fish, who shared a common ancestor 525 million years ago. And since cephalopods appear to have independently evolved awareness as well, it could be spread even farther and wider in the animal kingdom.

These estimates are, of course, 3rd-person conjectures using the best tools science has for studying consciousness. We cannot experience “what it is like” to be in another subject, so I

should also repeat here briefly what I mentioned above in question 5 about [my post](#) reacting to Nagel's thought experiment.

- The epistemological problem Nagel didn't want to raise explains the entire difficulty that his mind-body thought experiment supposedly raises. ... So to me, the fact that we can't know what it feels like to be a bat is actually an argument that *bolsters* physicalism, rather than questions it.

14. Do we have immortal souls?

The second extra issue raised by [Gennaro](#) with "the problem of consciousness is...related to major traditional topics in metaphysics, such as the possibility of immortality." The possibility of immaterial souls that go on forever has no evidence behind it and lots of other evidence to the contrary. Physicalists reject this idea, although I believe that ending the aging process in our human bodies in order to live [indefinitely long lives](#) is definitely an idea worth thinking and writing ([a novel](#)) about.

15. Do we have free will?

For the last of these issues raised by [Gennaro](#), there is the point that "the problem of consciousness is...related to major traditional topics in metaphysics, such as...the belief in free will." Quite luckily, while I was writing this series on consciousness, I was asked if I wanted to review Gregg Caruso and Dan Dennett's book on this subject ([Just Deserts](#)), which spurred me to dive deeply into the free will debate. After 6 posts exploring other people's positions, I wrote [a summary of my own thoughts](#). In a nutshell, I say we don't have the freest will imaginable, but we do have significant degrees of freedom, and that provides a kind of "free will worth wanting." Adding another functional analysis here using Tinbergen's four questions sheds a lot of light on the emergence and expansion of these freedoms, which are completely aligned with the emergence and expansion of consciousness. This linkage makes sense since [Dan Dennett](#) noted:

- "It is no mere coincidence that the philosophical problems of consciousness and free will are, together, the most intensely debated and (to some thinkers) ineluctably mysterious phenomena of all. As the author of five books on consciousness, two books on free will, and dozens of articles on both, I can attest to the generalization that you cannot explain consciousness without tackling free will, and vice versa."

Agreed. And tackled now.

QUESTIONS FROM OTHER NATURALISTS

In addition to the standard questions listed in the online encyclopedia articles that I cited above, I have found a few other questions worth discussing that have been raised by other naturalist philosophers. Let's go through those here.

16. Can't we just get by with a very rough definition of consciousness?

The philosopher Eric Schwitzgebel maintains an excellent blog called [The Splintered Mind](#), which often touches on topics in the field of consciousness. In 2016, Schwitzgebel published a paper called "[Phenomenal Consciousness, Defined and Defended as Innocently as I Can Manage](#)" in which he argued that the best approach for defining consciousness right now may be a "definition by example" which can work well "if one provides diverse positive and negative examples and if the target concept is natural enough that the target audience can be trusted to latch onto that concept once sufficient positive and negative examples are provided."

Let's see how this works in practice. Here are some of the positive examples Schwitzgebel lists: sensory and somatic experiences; conscious imagery; emotional experience; thinking and desiring; and dream experiences. Does that help yet? Here's a passage about the negative examples to keep at it.

- "Not everything going on inside of your body is part of your phenomenal consciousness. You do not, presumably, have phenomenally conscious experience of the growth of your fingernails, or of the absorption of lipids in your intestines, or of the release of growth hormones in your brain. Nor is everything that we normally classify as mental part of phenomenal consciousness. Before reading this sentence, you probably had no phenomenal consciousness of your disposition to answer 'twenty-four' when asked 'six times four'. ... If a visual display is presented for several milliseconds and then quickly masked, you do not have visual experience of that display (even if it later influences your behavior). ... [And] we normally think that dreamless sleep involves a complete absence of phenomenal consciousness."

Now that these have been introduced, Schwitzgebel concludes, "I suggest that there is one folk psychologically obvious concept, perhaps blurry-edged, that fits the positive and negative examples while leaving the contentious examples open and permitting wonder of the intended sort. That's the concept of phenomenal consciousness."

Is that very helpful, useful, or interesting? I don't really think so. Can we say more? Sure, but Schwitzgebel doesn't want us to go too far. He says, "At this point, it is tempting to clarify by making some epistemic or metaphysical commitments—whatever commitments seem plausible to you. You might say, 'those events with which we are most directly and infallibly acquainted' or 'the kinds of properties that can't be reduced to physical or functional role'. Please don't! Or at least, don't build these commitments into the definition. Such commitments risk introducing doubt or confusion in people who aren't sure they accept such commitments."

Okay, now we're just backing away from any of the hard work of understanding consciousness, and it's obvious that Schwitzgebel is only concerned with the very narrow conception of *conscious awareness*, which is level 5 in my hierarchy. Worse still, he's looking for the least common denominator that everyone can agree to. I'm afraid that will end up with as tiny and useless a definition as possible when more and more opinions are brought into the discussion. In fact, Schwitzgebel acknowledges, "My definition did commit me to a fairly strong claim about folk psychology: that there is a single obvious folk-psychological concept or category that matches the positive and negative examples." But this is exactly the type of

essential on/off switch that Dan Dennett warned about in his paper “[Darwin and the Overdue Demise of Essentialism](#).”

Sorry, but I don't think an overly simple and deliberately narrow definition will do. Far better to work on the comprehensive functional analysis that helps put everyone's various opinions in their own place and shows the relationships they all have to one another. Such an analysis helps us see the building blocks of consciousness and how they all emerge over evolutionary timescales. Schwitzgebel's definition does none of that work. I've set myself a lofty goal for my consciousness studies, but I do think its attainable.

Fortunately for us, Schwitzgebel's “innocent” and narrow definition doesn't actually stop him from exploring wider issues with consciousness, which I'll cover in questions 17 and 18.

17. What about the various parts of living systems? Which ones are conscious?

In a November 2020 post on his blog, Schwitzgebel laid out [the nesting problem for theories of consciousness](#). In this question I'll look at the nesting problem going down, and in the next one I'll consider it going up. First, though, what are we talking about exactly?

Schwitzgebel starts with the background that “in 2016, [Tomer Fekete, Cees Van Leeuwen, and Shimon Edelman](#) articulated a general problem for computational theories of consciousness, which they called the Boundary Problem. The problem extends to most mainstream functional or biological theories of consciousness, and I will call it the Nesting Problem.” Then, he gives this as a quick explanation:

- “Consider your favorite functional, biological, informational, or computational criterion of consciousness, criterion C. When a system has C, that system is, according to the theory, conscious. ... Unless you possess a fairly unusual and specific theory, probably the following will be true: Not only the whole animal (alternatively, the whole brain) will meet criterion C. So also will some subparts of the animal and some larger systems to which the animal belongs.”

This, then, yields some questions for Schwitzgebel:

- First: *Are* all these subsystems and groups conscious?
- Second: If we want to attribute consciousness only to the animal (alternatively, the whole brain) and not to its subsystems or to groups, on what grounds do we justify denying consciousness to subsystems or groups?
- Or maybe instead of a threshold, it's a comparative matter: Whenever systems nest, whichever has the most connectivity is the conscious system. ... Or maybe it's not really C (connectivity, in this example) alone but C plus such-and-such other features, which groups and subsystems lack. ... Or maybe groups and subsystems are also conscious — consciousness happens simultaneously at many levels of organization.

Schwitzgebel doesn't think these questions are unanswerable, just that, “this is a fundamental question about consciousness which is open to a variety of very different views, each of which brings challenges and puzzles—challenges and puzzles which philosophers and scientists of consciousness, with a few exceptions, have not yet seriously explored.”

This discussion really shows the beauty of having a comprehensive hierarchy of consciousness rather than a singular, restrictive, narrow definition. This issue started off as *the boundary problem*, but since we are dealing with a biologically emergent property, there are no clear boundaries here! It's obvious that any singular criterion C will have trouble moving up and down the story of consciousness. For me, that's not a problem.

Are all these subsystems conscious? No, your kidneys or autonomic nervous system have not reached conscious awareness in my hierarchy, but they do have the properties of consciousness that are included in my levels of affect and intention. According to Jeff Hawkins' **thousand brains theory**, they may also have local abilities for prediction too. And these subsystems contribute pieces of consciousness to other systems that may reach higher levels in my hierarchy. The point is that all of these elements can be analysed and understood for their contribution back and forth to the various levels of consciousness in living systems.

On what grounds do we justify denying consciousness to subsystems or groups? We don't deny them *all* of the levels of consciousness. We can just be clear about which ones they have and which ones they contribute to other systems that may or may not reach different levels of consciousness.

Maybe consciousness happens simultaneously at many levels of organization. That's right, as long as your definition of consciousness is as wide and flexible as mine is, yet capable of offering enough precision to describe the various varieties of consciousness that are on offer as well.

18. Is the United States conscious?

This question essentially extends the nesting problem in the upwards direction, although it is based on a paper from Schwitzgebel that is six years older than his post on the nesting problem. That paper is called, "**If Materialism Is True, the United States Is Probably Conscious.**" That title sounds ridiculous on the face of it, but let's give Schwitzgebel some benefit of the doubt and explore his claims in a bit of detail rather than just dismiss them. The explorations prove fairly illustrative for the benefit of taking an evolutionary approach here.

Schwitzgebel starts off by introducing us to two sci-fi scenarios that are meant to disabuse us of a prejudice he calls *contiguism*, which apparently stops us from believing in spatially distributed consciousnesses. The first are *Sirian Supersquids*. Here is their story:

- They can detach their limbs. To be detachable, a supersquid limb must be able to maintain homeostasis briefly on its own and suitable light-signal transceivers must appear on the surface of the limb and on the bodily surface to which the limb is normally attached. ... [Also], the limb-surface transceivers developed the ability to communicate directly among themselves without needing to pass signals through the central head. ... Despite their spatial discontinuity, they aren't mere collections. They are integrated systems that can be treated as beings of the sort that might house consciousness.

I would agree that these creatures could be spatially distributed, yet consciously integrated, but only because the information from the various parts is being integrated in one place. Any

signals not sent to the head would be analogous to the unconscious processing that goes on in our own bodies. Anyway, let's carry on. The second sci-fi creation are the *Antarean Antheads*. Here are the relevant bits of their story:

- [These are] a species of animals who look like woolly mammoths but who act much like human beings. ... Here's why I call them "antheads": Their heads and humps contain not neurons but rather ten million squirming insects, each a fraction of a millimeter across. Each insect has a complete set of minute sensory organs and a nervous system of its own, and the antheads' behavior arises from complex patterns of interaction among these individually dumb insects. ... Maybe there are little spatial gaps between the ants. Does it matter? Maybe, in the privacy of their homes, the ants sometimes disperse from the body, exiting and entering through the mouth. Does it matter? ... You might think that the individual ants would or could be individually conscious and that it's impossible for one conscious organism to be constituted by other conscious organisms. Some theoreticians of consciousness have said such things—though I've never seen a good justification of this view.

I believe the answers to these questions come from careful considerations of evolutionary biology. It's not so much that "it's impossible for one conscious organism to be constituted by other conscious organisms." That depends very much on your definition of consciousness and my answer to the previous question shows how subsystems with lower forms of consciousness integrate into higher systems that achieve higher levels of consciousness based on the extra information that is available to them. That is possible and completely consistent with the functional analysis enabled by my hierarchical theory. However, based on the evolutionary biology that has been observed in our world, it seems impossible for creatures like the *Antarean Antheads* to ever emerge.

I draw this conclusion from [*The Origins of Life*](#) by John Maynard Smith and Eors Szathmari, which covers [the major transitions in evolution](#). As I summarised in [a talk](#) I gave, "the big takeaway from this book is that each transition occurred when formerly separate and competitive biological elements figured out new ways to join up and cooperate with one another, and begin to evolve together." That sounds vaguely like what Schwitzgebel's antheads have done, but it runs afoul of this quote from p.19 of the book:

- "One feature [of major transitions] crops up repeatedly. Entities that were capable of independent replication before the transition could afterwards replicate only as part of a larger whole."

That reproductive integration is crucial! It's what actually enables natural selection to slowly work its magic on the shaping of these emergent new species. Evolution is said to require three steps: variation, selection, and retention. But there's no way for the antheads to manage this as a coherent species with such independent creatures like the ants in their heads. So, they make for a poor example whose seeming impossibility is unable to dissuade me from my so-called *contiguism*. Nevertheless, let's carry on with Schwitzgebel's paper as he combines some features from these two sci-fi creatures in order to investigate yet another one called the *Sirian Squidbit*, which brings up a few more issues.

- "The Sirian squidbits [are] a species with cognitive processing distributed among detachable limbs. ... Let me tie Sirius, Antares, and Earth a bit more tightly together. As

the squidbit continues to evolve, its central body becomes smaller and smaller—thus easier to hide—and the limbs develop more independent homeostatic and nutritional capacities, until the primary function of the central body is just reproduction of these increasingly independent limbs. Earthly entomologists come to refer to these central bodies as ‘queens’. Still later, squidbits enter into symbiotic relationship with brainless but mobile hives, and the thousand bits learn to hide within for safety. These mobile hives look something like woolly mammoths. Where is the sharp, principled line between group and individual?”

Schwitzgebel is clearly referencing the eusocial species of ants here and trying to use the fact that they are considered superorganisms to make it seem plausible that there can be something like *superconsciousness*. But once again the issue is resolved by the separability of reproductive biology from the biology of consciousness. Eusocial ants are considered superorganisms because they cannot reproduce as individuals. That is why they are only selected for at the group level. But that says nothing about the consciousness of such a group of individuals. As discussed above in question 9 about the binding problem, and in question 10 about the neural correlates of consciousness, there are several candidates for physical structures and processes that integrate elements of consciousness together. There are no features like these in ant colonies which could bind the consciousness of the individuals together even though they must reproduce as a group and are therefore selected and shaped on the basis of their collective actions. Unless consciousness is immaterial, there is no reason to believe in the consciousness of an ant colony. In fact, since there is no “spooky action at a distance” from one ant individual to another, there is no evidence for an immaterial consciousness there that is sensing and reacting to the needs of the group as a whole. Note that this is the case even though ants have been part of fiercely competitive superorganisms for millions of years! If superconsciousness were going to arise anywhere, surely it would be there. Anyway, that is how the lines between groups and individuals can be understood in ants and Sirian Squidbits.

Okay, but what about the United States? By now you must see why this is also problematic, but Schwitzgebel raises a number of other questions here (in a kind of [Gish Gallop??](#)), so let me tackle them as quickly as I can in rapid fire succession.

- You might say: The United States is not a biological organism. It doesn’t have a life cycle. It doesn’t reproduce. It’s not biologically integrated and homeostatic. Therefore, it’s just not the right *type of thing* to be conscious.

It’s not about the *type of thing*. I’m not a bio-chauvinist. It’s about the fact that the United States doesn’t have any mechanisms, phylogeny, or ontogeny—three of the four Tinbergen questions—which could contribute to any sense of a U.S. consciousness.

- Why should consciousness require being an organism in the biological sense? Properly-designed androids, brains in vats, gods—these things might not be organisms in the biological sense and yet are sometimes thought to have consciousness.

This point is fine. As I described above in question 12, consciousness may not require biology.

- Second, it’s not clear that nations aren’t biological organisms. ... other types of coordination emerge spontaneously from the bottom up, just as in ordinary animals.

If you actually look at the detailed definitions of life and organisms, it's quite clear that the United States doesn't qualify as an organism. I trust I don't need to explain this further.

- Nations also reproduce—not sexually but by fission.

This badly confuses culture with biology. Nations are merely an abstract notion. They don't reproduce in any way comparable to organisms.

- According to a broad class of plausible materialist views, any system with sophisticated enough information processing and environmental responsiveness, and perhaps the right kind of historical and environmental embedding, should have conscious experience. My central claim is: The United States seems to have what it takes, if standard materialist criteria are straightforwardly applied without post-hoc noodling. It is mainly unjustified morphological prejudice that blinds us to this.

Eeks. This sounds like a blatant **category error**. Our “morphological prejudice” remains well justified, and my materialist criteria require no “post-hoc noodling” to deny the consciousness of the United States. I'll discuss the problems with linking consciousness to information processing alone in question 40 below.

- Consider, first, the sheer quantity of information transfer among members of the United States. ... Our information exchange is not in the form of a simply-structured massive internet download. The United States is a goal-directed entity, flexibly self-protecting and self-preserving. The United States responds, intelligently or semi-intelligently, to opportunities and threats. ... I am asking you to think of the United States as a planet-sized alien might, that is, to evaluate the behaviors and capacities of the United States as a concrete, spatially distributed entity with people as some or all of its parts, an entity within which individual people play roles somewhat analogous to the role that individual cells play in your body.

Yes, indeed, this is the mother of all category errors. The United States is not a “concrete, spatially distributed entity.” It's just an abstract idea. We can't draw an abstract line around every imaginable group and declare it to have its own consciousness. We don't think there is a consciousness of “left-handed NBA fans” no matter how similar that group is to a nation.

Schwitzgebel asked us to consider the consciousness of the United States, but of course the same question is often asked of other super-entities such as ecosystems or the whole earth of Gaia. Well, all of the same arguments in this question apply to those situations as well and deny any likelihood of superconsciousness there either. Let me just add this extra quote from **John Maynard Smith and Eors Szathmary** as a final piece of evidence:

- “Consider a present-day ecosystem—for example, a forest or a lake. The individual organisms of each species are replicators; each reproduces its kind. There are interactions between individuals, both within and between species, affecting their chances of survival and reproduction. There is a massive amount of information in the system, but it is information specific to individuals. There is no additional information concerned with regulating the system as a whole. It is therefore misleading to think of an ecosystem as a super-organism.”

19. How do we know we don't have "inverted qualia"?

This is a quick little issue that was mentioned in *The Guardian* in [a long review](#) of Mark Solms' recent book about consciousness. The author noted that, "the 'problem of inverted qualia' refers to the fact that the experience you call 'seeing green' could be identical to the one I call 'seeing red', and vice versa, and we'd never have any way of knowing."

Based on my response to question 13 about "what it is like" to be conscious, we physicalists admit that we can't actually know what others are experiencing. That barrier is completely consistent with physicalism, and in fact it is a consequence of the universe being confined to the physical. (If consciousness arose from immaterial mental properties, you'd think we would already have found a way to inhabit other physical bodies and therefore know what it was like in them.) However, the shared evolutionary history of all life, and the shared physical building blocks we are all made of precludes any reason to think any of us actually have inverted qualia. After all, the most common cause of [color blindness](#) is "an inherited problem in the development of one or more of the three sets of the eyes' cone cells, which sense color." Once again, changes in the subjective experience of consciousness are matched by physical changes in the body experiencing that consciousness. My reliance on evolution here leads us to the next question.

20. How do you solve the mind-evolution problem?

Based on its title, the neurobiologist Yoram Gutfreund wrote a really challenging paper for me called "[The Mind-Evolution Problem: The Difficulty of Fitting Consciousness in an Evolutionary Framework](#)." In that paper, Gutfreund described how,

- "Consciousness is one of the last biological phenomena about which we do not have a solid idea as to how and when it appeared and evolved in evolution. ... The question of how the mind emerged in evolution (the mind-evolution problem) is tightly linked with the question of how the mind emerges from the brain (the mind-body problem). It seems that the evolution of consciousness cannot be resolved without first solving the 'hard problem' (Chalmers, 1995). Until then, I argue that strong claims about the evolution of consciousness based on the evolution of cognition are premature and unfalsifiable."

I agree with Gutfreund that this mind-evolution problem is tightly linked with the mind-body problem and the hard problem. In question 4 about how minds could have emerged from matter, I explained how my theory of pandynamism fits the evidence in the world where consciousness appears to emerge and grow along with the emergence of living subjects. This answers the hard problem by naming "felt forces" as an underlying identity in the universe. These felt forces, then, grow and change in subjective consciousnesses as the subjects grow and change their structures for sensing these forces. Changes in the world that affect my five senses will change my conscious experience. If I lose a sense (e.g. if I go blind), then changes in the light around me no longer affect my consciousness.

Gutfreund had identified four possibilities for any attempts to fit consciousness into an evolutionary framework. Presumably, one of them will work for me if my theory is worth considering. Gutfreund's four possibilities are:

1. *Consciousness as a tool for behavior.* Is consciousness to an animal like wings are to a bird, i.e., a tool to enable an advantageous goal? If consciousness is a tool, what is the goal that it enables? Some answers include: to create a unified and coherent representation of all incoming information (Crick and Koch, 1998; Merker, 2005); to enable the learning of sensory and cognitive representations (Grossberg, 1999); to make complex flexible decisions (Earl, 2014); and more. ... Difficulty with this notion is that cognitive behaviors are caused by the brain's neural circuits, without the necessity to introduce conscious states to the models.
2. *Consciousness as brain identity.* One escape route around this paradox is to suggest an identity between consciousness and neuronal states (Loorits, 2014; Smart, 2017), that is, some neuronal states are conscious feelings; the two are the same, described at different levels. The biological function of the neural state then becomes the function of the feeling (Searle, 2013). A problem with such an identity approach is that evolution operates at the level of the body and not at the level of the feelings. The only things that matter from an evolutionary point of view are the animal's actions, and the neural processes that choose and elicit the actions. ... Therefore, the implication of an identity hypothesis is that consciousness becomes detached from any evolutionary theory.
3. *Consciousness as an advantageous goal.* What if consciousness is a goal in itself? In this case, neurons organized in specific ways in specific brain structures are the wings to support consciousness, and the property of being conscious improves the fitness of the animal in which it is installed, just like the properties of flying, swimming, or chewing. But, in what ways do feelings and emotions improve fitness? An antelope escaping from a lion needs to run quickly and efficiently. Why, from an evolutionary point of view, does it also need to feel the terrible feeling of fear? This is a puzzle and evolutionary theory has no answers.
4. *Consciousness as a by-product.* A different approach that bypasses the difficulties described above is to view consciousness as a byproduct of brain activity. In this case, consciousness doesn't affect behavior and has no function of its own. However, it has an adaptive value that stems from its association with a behavioral phenomenon, which in turn does have a function. ... The pitfall of such an approach is that consciousness can be removed from the model without any influence on the flow of the model.

Once again, Gutfreund appears to only be considering “consciousness” as some narrow part of conscious awareness, and this makes it quite difficult to trace the evolutionary path and usefulness of that small piece. By tracing the history of the evolution of all forms of cognition, and embracing all of those associated functions and behaviours as different aspects of consciousness, I think it becomes easier to see the slow emergence of *consciousness as an identity with living systems* (i.e. #2, but not just for brains), which impacts all behaviour and therefore *acts as a tool* (i.e. #1 but with a much broader reach of enabling and improving survival across many different routes). Consciousness is not *a goal in itself* (#3) or an *epiphenomenal by-product* (#4) since it is just an unavoidable part of life, which is unavoidably shaped by evolution and natural selection.

All of this is best traced in my post on [the functions of consciousness](#) where my hierarchy was first developed in full. Here are a couple of quick highlights:

- As soon as the origin of life takes hold in the first level of my hierarchy, the next tier of affect begins to get embedded as living entities *feel* their way through life and quickly develop associations between good and bad feelings as they relate to life and death. These are innately passed down through successful generations.

- Over time, adaptations from affective reflexes alone lead to capacities for cognition that are able to interrupt these reflexes. The capacities of attention, memory, pattern recognition, learning, and communication create a core self where organisms can be said to be acting with intention, which is the third level of my hierarchy.
- Once intentions exist, they can be taken into account. To do so is to use prediction (my fourth level) to think through what the result will be from any intentions. This requires the cognitive capacities of anticipation, problem solving, and error detection.
- As predictions and perceptions improve, organisms eventually make the connection that there is a self which has its own mind. The fifth level of awareness is achieved, along with the arrival of the cognitive capacity for self-reference. Such conscious cognition allows memories and thoughts built from the lived past and the anticipated future to create the autobiographical self. Note that this is often the level that neuroscientists concern themselves with and only a few extra abilities seem to emerge here such as “trace conditioning” and the recreation in thought of past events in order to learn from them anew in light of new information.
- Finally, in the sixth and final level of my hierarchy of consciousness, the ability of conscious and aware selves to make abstract connections gives rise to language, which immeasurably expands the scale and scope of one’s thoughts for consideration.

Note that these final two levels address what Dan Dennett calls “[the hard question of consciousness](#).” According to Dennett, “the so-called hard *problem* of consciousness is a chimera, a distraction from the hard *question* of consciousness, which is once some content reaches consciousness, ‘then what happens?’ . . . The question, more specifically, is: *Once some item or content ‘enters consciousness’, what does this cause or enable or modify?* For several reasons, researchers have typically either postponed addressing this question or failed to recognize—and assert—that their research on the ‘easy problems’ can be seen as addressing and resolving aspects of the hard question, thereby indirectly dismantling the hard problem piece by piece, without need of any revolution in science.”

Dennett is probably right that a focus on all the tools and functions of consciousness ends up dismantling the hard problem. As all of the details for this have rolled in, the only thing left for the hard problem to cover is why there is consciousness at all. Well, we can never answer all why questions. Some things appear to just be here, like gravity, or electromagnetism, or any other fundamental force in the universe. And now that we have listed out all the basic ingredients of consciousness and observed that they have been around since the very beginning of life, that makes it trivially easy for me to posit pandynamism as an underlying identity in the universe, which helps us see the bridge between the forces which affect all matter and the subjectivity those forces cause in subjects once subjects emerge. As for the question, “*what does this cause or enable or modify?*”, it clearly causes survival behaviour in ever expanding capacities towards more and more robust survival. More on that in the next question.

21. Does consciousness have a purpose?

The great evolutionary biologist [Ernst Mayr](#) is perhaps best known for helping to define the [modern synthesis](#), but as an evolutionary philosopher, I’m also very interested in the distinction he made between [proximate and ultimate causations](#). Mayr used this to show that biology just cannot be reduced to one thing; it must instead be analysed holistically.

Proximate causation “explains biological function in terms of immediate physiological or environmental factors” whereas ultimate causation “explains traits in terms of evolutionary forces acting on them.” Some examples make this clearer.

- Proximate description: “A female animal chooses to mate with a particular male during a mate choice trial. A possible proximate explanation states that one male produced a more intense signal, leading to elevated hormone levels in the female producing copulatory behaviour.”
- Ultimate description: “Female animals often display preferences among male display traits, such as song. An ultimate explanation based on sexual selection states that females who display preferences have more vigorous or more attractive male offspring.”

Note that the behaviour in these two examples is exactly the same. We just come to understand the situation better when we look at all the levels of causation. Nicholaas Tinbergen divided these two causations even further when he developed his [Four Questions](#), which I have found to be crucial for understanding the entire story of consciousness. But in a wonderful paper by the philosopher Brandon Conley about how to [disentangle and integrate Mayr and Tinbergen’s views](#), we can see how Mayr’s simpler distinctions help address a longstanding issue in the philosophy of biology. Conley writes:

- “According to Mayr, ‘The clear recognition of two types of causation in organisms has helped to solve an important problem in biology, the problem of teleology.’ A hallmark of the scientific revolution was the rejection of ancient and medieval applications of teleological reasoning to the cosmos. In slogan form, physics progressed when it came to focus on causes rather than purposes. Biology, on the other hand, and evolutionary biology in particular, appears to require reasoning about what a given trait is for, or what good it does for the organism. Biological explanation appears to be ineliminably teleological, but according to dominant conceptions of scientific reasoning, teleological reasoning is unscientific. There are three possible responses to this: (1) claim that biological explanation is not really teleological, (2) admit that biological explanation is not really scientific, or (3) claim that teleological reasoning can be scientific after all. Philosophers and scientists have tried all three, but Mayr argues that the class of processes that have been labeled as teleological are not unified and a combination of all three strategies is necessary.”

For a quick reminder of what [telos/teleology](#) is, this traces back to Aristotle and can mean *purpose, intent, end, or goal*. In particular, “Aristotle used it in a more specific and subtle sense—the *inherent* purpose of each thing, the ultimate reason for each thing being the way it is, whether created that way by human beings or nature.” As noted in the passage above, modern physics made progress when this concept was removed from the field. But it’s important to acknowledge that this was only possible because non-living matter simply *reacts* to the forces that are applied to it. Biology, on the other hand, deals with living things that can *act* too. By definition, living things act to stay alive. They have evolved an internal drive to maintain their lives. An external observer can look at these actions and say they *want* to stay alive.

Another way of describing this is by using what Dan Dennett called the *intentional stance*. In a nice [profile of Dennett in the New Yorker](#), this term was explained in the following clear and helpful way.

- “During the course of his career, Dennett has developed a way of looking at the process by which raw matter becomes functional. Some objects are mere assemblages of atoms to us and have only a physical dimension; when we think of them, he says, we adopt a ‘physicalist stance’—the stance we inhabit when, using equations, we predict the direction of a tropical storm. When it comes to more sophisticated objects, which have purposes and functions, we typically adopt a ‘design stance’. We say that a leaf’s ‘purpose’ is to capture energy from sunlight, and that a nut and bolt are designed to fit together. Finally, there are objects that seem to have beliefs and desires, toward which we take the ‘intentional stance’. If you’re playing chess with a chess computer, you don’t scrutinize the conductive properties of its circuits or contemplate the inner workings of its operating system (the physicalist and design stances, respectively); you ask how the program is thinking, what it’s planning, what it ‘wants’ to do. These different stances capture different levels of reality, and our language reveals which one we’ve adopted.”

Getting back to the question of teleology or purpose in biology, we know that physical matter reacts to physical forces. And in my post taking us [from physics to chemistry to biology](#), I identified a set of “biological forces” that are missing from our scientific description of the world, but which clearly cause biology to react in predictable ways. Unlike with mere matter, however, living systems don’t simply react in perfectly repeatable and definitively knowable ways. Biological life learns, grows, and changes how it reacts to biological forces by using the various aspects of consciousness at its disposal to sense and respond to the environment in order to drive its behaviour toward the ultimate goal of survival. This, of course, isn’t a goal that has been designed by anyone. Nor is it even apparent to any beings in the grips of proximate goals. This is actually why Darwin faced [problems with the term natural selection](#) — it vaguely implied a selector — and so he toyed with the idea of calling the central force in evolution “natural preservation” instead. But logically, the survival goal must be the ultimate necessary outcome in a universe where things change, and nothing lives forever. Any and all proximate goals that don’t work towards this will end up going extinct.

With all of this in mind, we can now answer this question, and choose from among Mayr’s three responses. Consciousness does indeed have a purpose or telos, but it is one that emerges from selection forces rather than intentional designs. Because living beings *act* as well as *react*, it is necessary to look at *underlying causes* (biological forces) as well as *teleological purposes* (survival, ultimately) if we want to holistically understand the way that life works. In this way, teleological reasoning can be scientific after all (Mayr’s third choice), and in fact it is necessary for full scientific reckonings. (This is also why telos sits at the top of [my evolutionary hierarchy of needs](#).) Consciousness, in all its emergent and expanding properties, helps life sense and respond to the world in order to maintain its survival and make it more and more robust.

QUESTIONS FROM THOSE WHO DOUBT OR DISBELIEVE NATURALISM

For this next batch of questions, I wanted to make sure I wasn’t just preaching to the choir or responding to people who already held favourable dispositions toward the naturalist project. I

wanted to make sure I properly understood objections from the other side. To that end, I have some questions from Raymond Tallis and Philip Goff, which I'll cover in that order because it takes us through their points in increasing levels of difficulty and importance.

To start, I have three questions from Ray Tallis' recent book [*Seeing Ourselves: Reclaiming Humanity from God and Science*](#). Tallis is a retired physician and [patron of Humanists UK](#) who was once named as "[one of the top living polymaths in the world](#)." A local philosophy group really likes his work so I've had a chance to meet him in person a few times and I was once scheduled to discuss *Seeing Ourselves* with him in a public Humanist meeting, but that event fell through. After reading the first few chapters and plucking out the questions below, you may see why this cancellation was for the best.

22. Why doesn't a chair feel my bottom?

It's hard to believe this is an actual question, so let's quote Tallis directly to see what he really means by this.

- "If energy exchange entirely accounted for touch then it would be as reasonable for the chair on which I am sitting to feel my bottom as for my bottom to feel the chair: the ontological equality of myself as an object among objects does not translate into a dialogue of equal partners. That ontological equality, however, is central to materialist naturalism." ([Seeing Ourselves](#) p. 54)
- "The causal theory of perception, in which all parties are subject to the Dennettian edict of being subject to *the same physical principles, laws, and raw materials* that operate elsewhere in nature offers nothing to explain the differentiation between the perceiving subject and object of perception; between the perceiver and the perceived." ([S.O.](#) p. 54)

What an absurd caricature of the naturalist position! Tallis loves taking cheap shots at Dan Dennett like this, even though he grossly misunderstands him. (See [my review](#) of *Seeing Ourselves* for more on that.) Can naturalists explain why the blind naked mole rat doesn't see me even though I can see him? Of course we can! It is not just that "stuff feels" which explains the sense of touch. It is the *structure* of that stuff ("that ontological equality") which enables subjectivity to emerge in subjects via various [mechanisms](#).

23. How can consciousness survive sleep?

Here is yet another daft-sounding question that requires direct quotation for context.

- "One way of addressing the so-called combinatorial problem — the problem of explaining how sequins of consciousness spread through the world add up to a subject — is to deflate the subject. The subject is reduced to successive experiences, or time-slices of a flow of experience: there are no persisting subjects; each distinct experience has its own experiencer. This merely transfers the problem to that of explaining how experiences add up to a subject who has a sense of herself at a time and over time and is acknowledged to be a person by other subjects also acknowledged to be persons. It is not at all clear by what means, by who or what, the thin subjects are stitched together and how we would survive sleep or episodes of unconsciousness." ([S.O.](#) note 84 on p.393 referring to p.66)

So, Tallis is really referencing the binding problem here, which I addressed above in question 9. Admittedly, we don't have full mappings of all the neuroscience in the animal kingdom yet to give full explanations for how subjectivity is stitched together (to the extent that it is, anyway, since it **varies across the animal kingdom**). I think **Jeff Hawkins'** solution shows promise, but it's too early to say for sure.

What isn't a solution to the binding problem is "to deflate the subject [so] there are no persisting subjects." Once again, it is the *structure* of material that gives materialists their mechanisms for consciousness, and that structure clearly survives sleep and episodes of unconsciousness. (And the structures change slightly to cause those states of sleep and unconsciousness too.)

24. How could consciousness have possibly emerged from lower organisms?

There is another legitimate question, which I addressed in questions 4 and 20 above. I also described the actual evolutionary history of consciousness in much more detail in my post about **our shared history (phylogeny)**. Granted, this may be "**the hardest problem in consciousness studies**," but while Tallis grapples with it, he displays such a bewildering lack of understanding about evolution that it's no wonder he doesn't see the materialist argument. Some more direct quotes will show the paucity of Tallis' beliefs, which I'll just comment on briefly after each one so we can move on to better foes.

- "Darwinism highlights (if inadvertently) a serious objection to Darwinitis, namely, that Darwinism gives no account of the *emergence* of consciousness from the material world of which lower organisms are a part." (**S.O.** p. 66)

I have now given just such a Darwinian account, so perhaps that should be considered, but let's be clear here, no other metaphysical theories have provided an account either. And we're still learning about the universe so there's no cause to dismiss naturalism as Darwinitis just yet. Also, how dare you say that you believe in evolution but place "lower organisms" in the material world separate from humans. When did that break occur exactly? This is sheer hubris, and it's dangerous to the survival of life too.

- "There are at least two major obstacles to a materialist evolutionary account of human consciousness: the first is the question of the nature of the supposed competitive advantage conferred by being conscious; and the second is the question of how, even if consciousness *did* confer survival benefit, it could have been generated between unconscious species; that, as a result of the blood bath of natural selection, the universe could get to know itself." (**S.O.** p. 67)

That first major obstacle is legitimate and addressed in question 21 above. But that second obstacle is an embarrassment of logic by which it's hard to see how Tallis thinks *any* novel trait could emerge from evolution. Yikes.

- "Given that pre-conscious processes do so much work; there is not much useful work left for consciousness to do. To address this question properly, we need to go back to a putative moment when the first spark of consciousness was lit and ask what possible

additional advantage would an organism with (say) an array of photosensitive cells gain from being *aware* of the light it is responding to? The “obvious” benefits vanish when we acknowledge: a) That the best route to replication of the genome must be via utterly reliable mechanisms based on the (by definition) unbreakable laws of nature rather than the vagaries of (conscious) decision making; b) Evolution should favour appropriate action, but it is not evident that this should have to be mediated by true belief or indeed any belief; and c) That unconscious mechanisms have been perfectly adequate to bring about things that consciousness could not even dream of, such as the basis of the organism’s self-maintenance (including its voluntary actions), the spectacular achievement of the development of the human brain *in utero*, and the entire evolutionary process.” ([S.O.](#) p. 67)

Wow. This is just a mishmash of very muddled thinking. First, Tallis appears to consider “consciousness” to just be “conscious awareness” which we’ve established above will always get you into trouble. In my theory, “pre-conscious” processes are just lower levels of consciousness. Conscious awareness cannot arrive for a mere “array of photosensitive cells” because there is no structure there to evaluate the affect, intention, and prediction levels that are further down in my hierarchy. But they must all be there before “the first spark” of conscious awareness emerges. I could forgive such confusion about the confusing terminology used in consciousness studies, but the three-part argument in the latter half of this quote is inexcusable. a) The best route to *survival* of genomes is *not* “utterly reliable mechanisms” because that would leave no room for change and adaptation. Perfectly repeated organisms (if they ever existed) would have gone extinct at the first sign of trouble. No laws of nature stop mutations and genetic drift from happening. And conscious decision-making (to focus only on the conscious awareness that Tallis is describing) allows beings who have attained that ability to conduct mental trials and errors so their ideas can perish rather than themselves. That is hardly a vagary of living successfully. b) Useful beliefs about the world improve one’s actions. Those are favoured by evolution. c) Unconscious actions are indeed driven by lower levels of consciousness, particularly the cognitions in my levels of affect and intention. Conscious awareness emerges on top of those and enables yet further behavioural adjustments by these already very finely tuned biological machines.

- “If, say, consciousness is necessary for learning and plasticity, then we have to ask why is it not always necessary for learning and plasticity. In most cases, learning and plasticity do not require the conscious participation of the organism.” ([S.O.](#) p. 67)

As discussed above in question 20, conscious awareness seems to enable “trace conditioning” which is another form of learning that is added to all the unconscious learning that is possible using lower levels of consciousness. Not all learning and plasticity is equal.

- Natural selection can act only on what is already available. It seems inconceivable that it could *generate*, even less requisition, entirely novel properties such as consciousness. The clash between forms of organic matter over limited means to life seems hardly likely to give rise to something that goes beyond the material world, namely intentionality. ([Seeing Ourselves](#) p. 69)

This is simply more evidence of Tallis’ complete lack of understanding about evolution. This is embarrassing now, and fully illustrates why Tallis’ objections are so easily cast aside. Time to move on and see what we can learn from better foes. The next two questions come from

the philosopher Philip Goff who is the new poster boy for panpsychism. I covered his general views in [the fourth post in this series](#), but let's take a look at some of his specific objections now that my own theory has been fully elucidated.

25. Is conscious experience outside of the realm of science?

This first question comes from a short paper by Goff titled, "[Why Science Can't Explain Consciousness](#)." This is clearly related to his longer book [Galileo's Error](#), but based on this paper (and the chance I had to personally hear Goff in a small meeting in Durham) I don't think it's necessary to read that. Let's look at just a few quotes from Goff to see if you agree before I hit back with my response.

- "Here is Galileo describing his conception of matter: '...Hence I think that tastes, odours and colours, and so on are no more than mere names as far as the object in which we place them is concerned, and that they reside only in the consciousness.'"
- "In taking the qualities of consciousness not to be instantiated by material bodies, Galileo seems to be taking the qualities of consciousness to reside in an immaterial substance."
- "This rough sketch of nature was a short time later turned into a rigorous metaphysical view by Descartes. For Descartes, colours and smells and odours result from the interaction of immaterial minds with physical bodies."
- "As the result of this radical new Galilean/Cartesian metaphysics, we have, perhaps for the first time in history, a picture of the material world such that its nature can be completely captured in mathematics. Sensory qualities—the taste of the lemon, the smell of the flowers—cannot be entirely captured in mathematical language. So long as philosophers took such qualities to reside in the physical world, the scientific revolution was impossible. But once the physical world had been divested of qualitative nature, the remaining quantitative nature, concerning the way in which objects fill space, could be entirely captured in geometry. By putting sensory qualities in the conscious mind, and putting the conscious mind outside of the physical world, Galileo and Descartes provided the metaphysical underpinnings of the scientific revolution."
- "Physics, for all its virtues, gives us a radically incomplete picture of the world. It provides a description of the world that necessarily abstracts from the one aspect of concrete reality we know for certain to exist: the qualities of consciousness that are immediately and indubitably known to each of us."

There is much more in Goff's paper (and presumably in his book too), but this is enough to see that he's relying on dualist metaphysics from the 1600's that was [very poorly argued](#) at the time and has largely been discarded by modern thinkers. There is no need to think we have all placed qualitative research into some immaterial realm just because Galileo may have written about it that way. In question 7 above about whether consciousness would always remain a mystery, I wrote:

- And while we must acknowledge there are epistemological barriers to what any one person can know about their brains or the consciousness of others, we can "combine the two perspectives within certain experimental contexts. Both first-person and third-person scientific data about the brain and consciousness can be acquired and used to solve the hard problem." Scientists do this all the time.

As for “the qualities of consciousness that are immediately and indubitably known to each of us,” there are no such things and Goff’s argument evaporates once this illusion is broken. I particularly like these two quotes from Dan Dennett’s paper “[Facing Up to the Hard Question of Consciousness](#)” for dispatching this nonsense.

- “Over the past few centuries, our understanding of how vision is accomplished has grown magnificently, and one of the striking facts about what we have learned is that until scientists told us, we had no idea at all, no ‘privileged access’, to the complicated activities of the optic nerve, the occipital cortex, and even the activities of our eyeballs.”
- “The fact is, the traditional claim that our conscious minds are immediately and maybe even perfectly known to each of us is wildly false. The psychologist Karl Lashley once suggested provocatively that ‘no activity of the mind is ever conscious’, by which he meant to draw our attention to the inaccessibility of the processing that we know must go on when we think. What ‘we’ do ‘have access to’ is the contents and apparent temporal order of those contents, but how these contents, these representations of properties, objects and events, manage to represent what they do, and how they are generated when they ‘appear’ to ‘us’ is completely off-limits to introspection.”

26. Are minds everywhere? What about panpsychism?

Here’s one more quick question, based on Goff’s notorious essay “[Panpsychism is crazy, but it’s also most probably true](#).” The argument is very basic, so let me give it in a few quotes.

- “According to panpsychism, the smallest bits of matter—things such as electrons and quarks—have very basic kinds of experience; an electron has an inner life. The main objection made to panpsychism is that it is ‘crazy’ and ‘just obviously wrong’. It is thought to be highly counterintuitive to suppose that an electron has some kind of inner life, no matter how basic, and this is taken to be a very strong reason to doubt the truth of panpsychism.”
- “Scientific support for a theory comes not merely from the fact that it explains the evidence, but from the fact that it is the *best* explanation of the evidence, where a theory is ‘better’ to the extent that it is more simple, elegant, and parsimonious than its rivals.”
- “In fact, the only thing we know about the intrinsic nature of matter is that some of it—the stuff in brains—involves experience. We now face a theoretical choice. We either suppose that the intrinsic nature of fundamental particles involves experience, or we suppose that they have some entirely unknown intrinsic nature.”

That’s it?! Well, there’s actually a third choice that is just as simple, elegant, and parsimonious, which doesn’t result in the just obviously wrong and counterintuitive notion that an electron has some kind of inner life. That’s my theory of pandynamism, which I explained above in question 4 about how we might be able to get minds from matter. It’s not just that matter feels subjectivity. It’s that you need the right *structure* for that subjectivity to emerge in actual subjects. I’ll repeat my comparison of these two theories here:

- As an example, take the simplest force. What does it take to ‘feel’ gravity? For us humans, it’s registering the difference between inner ear liquids as our movements in space accelerate or decelerate. Can a rock or a photon ever experience this? No. Why not?

Because there is no structure in its makeup by which it could gain such information. Panpsychism is therefore a non-starter for me, but *pandynamism* could explain how subjectivity is a fundamental feature of the universe, yet only emerges as living organisms emerge, thus bridging the explanatory gap and providing a coherent answer to the hard problem.

QUESTIONS FROM DAVID CHALMERS

Okay, that's enough from those two foes of naturalism. Now for the full arguments of the man best known for throwing up stumbling blocks for consciousness studies, David Chalmers. He coined the phrase *the hard problem*, but what's so hard about it and what else does he object to and worry about? To find out, I've carefully gone through his 30-page paper "The Problem of Consciousness" ([TPoC](#), hereafter). According to the abstract, "this paper is an edited transcription of a talk at the 1997 Montreal symposium on 'Consciousness at the Frontiers of Neuroscience.'" I found it to be an incredibly useful paper and would like to finish up this long list of FAQs (and my entire consciousness series!) by going through it in detail.

27. What are the easy problems of consciousness?

According to [TPoC](#),

- The easy problems of consciousness include those of explaining the following phenomena:
 - the ability to discriminate, categorize, and react to environmental stimuli;
 - the integration of information by a cognitive system;
 - the reportability of mental states;
 - the ability of a system to access its own internal states;
 - the focus of attention;
 - the deliberate control of behavior;
 - the difference between wakefulness and sleep.
- All of these phenomena are associated with the notion of consciousness. For example, one sometimes says that a mental state is conscious when it is verbally reportable, or when it is internally accessible. Sometimes a system is said to be conscious of some information when it has the ability to react on the basis of that information, or, more strongly, when it attends to that information, or when it can integrate that information and exploit it in the sophisticated control of behavior. We sometimes say that an action is conscious precisely when it is deliberate. Often, we say that an organism is conscious as another way of saying that it is awake.
- In each case, an appropriate cognitive or neurophysiological model can clearly do the explanatory work. If these phenomena were all there was to consciousness, then consciousness would not be much of a problem. Although we do not yet have anything close to a complete explanation of these phenomena, we have a clear idea of how we might go about explaining them. This is why I call these problems the easy problems.

That second bullet point illustrates the very wide variance in the usage of the term *consciousness*, which is another reason why I've done my best to rope them all into a comprehensive hierarchy. Researching the Tinbergen history of all of these easy problems is

precisely what enabled me to set out the hierarchy as I have, while still recognising there are plenty of details to be filled in yet.

28. What is the hard problem of consciousness?

According to [TPoC](#),

- The really hard problem of consciousness is the problem of experience. When we think and perceive, there is a whirl of information-processing, but there is also a subjective aspect. As Nagel (1974) has put it, there is something it is like to be a conscious organism. This subjective aspect is experience.
- Why is it that when our cognitive systems engage in visual and auditory information-processing, we have visual or auditory experience: the quality of deep blue, the sensation of middle C? How can we explain why there is something it is like to entertain a mental image, or to experience an emotion? It is widely agreed that experience arises from a physical basis, but we have no good explanation of why and how it so arises. Why should physical processing give rise to a rich inner life at all? It seems objectively unreasonable that it should, and yet it does.
- If any problem qualifies *as* the problem of consciousness, it is this one. In this central sense of ‘consciousness’, an organism is conscious if there is something it is like to be that organism, and a mental state is conscious if there is something it is like to be in that state. Sometimes terms such as ‘phenomenal consciousness’ and ‘qualia’ are also used here, but I find it more natural to speak of ‘conscious experience’ or simply ‘experience’. Another useful way to avoid confusion (used by e.g., Newell 1990 Chalmers 1996) is to reserve the term ‘consciousness’ for the phenomena of experience, using the less loaded term ‘awareness’ for the more straightforward phenomena described earlier. If such a convention were widely adopted, communication would be much easier; as things stand, those who talk about ‘consciousness’ are frequently talking past each other. ([Chalmers](#))

Agreed! As we’ve seen throughout this series, researchers and philosophers frequently are talking about ‘awareness’ while others have something else in mind for ‘consciousness’ so they are indeed talking past one another or not getting to the root of the problem. The entire “phenomena of experience” is what I’m after with my comprehensive hierarchy of consciousness. And when we see how those phenomena exists across the entire spectrum of life, and over life’s entire evolutionary history, but it does not seem to extend into any non-living organic systems, then it makes sense to posit *pandynamism* (see question 4 for details) as the theory for why subjectivity is a fundamental identity of the universe but it only arises in subjects.

29. What does it take to solve the easy problems of consciousness?

According to [TPoC](#),

- The easy problems are easy precisely because they concern the explanation of cognitive *abilities* and *functions*. To explain a cognitive function, we need only specify a mechanism that can perform the function. The methods of cognitive science are well-suited for this sort of explanation, and so are well-suited to the easy problems of consciousness. By

contrast, the hard problem is hard precisely because it is not a problem about the performance of functions. The problem persists even when the performance of all the relevant functions is explained.

- Once we have specified the neural or computational mechanism that performs the function of verbal report, for example, the bulk of our work in explaining reportability is over. ... All it could *possibly* take to explain reportability is an explanation of how the relevant function is performed; the same goes for the other phenomena in question.

Not quite! **Functions** and **mechanisms** are only half of **Tinbergen's four questions**. We gain a lot of insight from looking through the **ontogeny** and **phylogeny** of these phenomena too. Seeing these evolutionary histories is precisely how we see the logic and empirical data for putting everything into the ordered hierarchy as I have done.

30. Is the hard problem really different than the easy ones?

According to **TPoC**,

- When it comes to conscious experience, this sort of [easy] explanation fails. What makes the hard problem hard and almost unique is that it goes beyond problems about the performance of functions. ... even when we have explained the performance of all the cognitive and behavioral functions in the vicinity of experience—perceptual discrimination, categorization, internal access, verbal report—there may still remain a further unanswered question: *Why is the performance of these functions accompanied by experience?*
- If someone says, “I can see that you have explained how DNA stores and transmits hereditary information from one generation to the next, but you have not explained how it is a gene”, then they are making a conceptual mistake. All it means to *be* a gene is to be an entity that performs the relevant storage and transmission function. But if someone says, “I can see that you have explained how information is discriminated, integrated, and reported, but you have not explained how it is *experienced*”, they are not making a conceptual mistake.
- We know that conscious experience does arise when these functions are performed, but the very fact that it arises is the central mystery. There is an *explanatory gap* (a term due to Levine 1983) between the functions and experience, and we need an explanatory bridge to cross it.

When I first discussed the hard problem in **post 3** of this series, I noted that “I'd like to make a distinction for Chalmers' hard problem between the how and the why. *How* do physical processes lead to subjective experience? *Why* do physical processes lead to subjective experience? The ultimate why is ultimately an impossible problem.” Chalmers' hard problem is clearly a why problem, and perhaps an impossible why.

In the opening of **a recent Brain Science podcast**, the neuroscientist Anil Seth said much the same and pushed back on Chalmers by saying,

- “It's not essential for a branch of science to explain why the phenomenon is there in the first place. Physics...doesn't tell us why there is a universe in the first place to explain. We often set a higher bar for consciousness than we do for other things. Partly because we are conscious. We want that intuitive *a ha* that makes sense. There's absolutely no reason why

a scientific account of consciousness should be intuitively satisfying. It would be nice if it were, but that's not strictly necessary.”

In some respects, Chalmers is playing a game of eternal regression here by just continuing to ask *why* for consciousness. But by doing so, he ends up driving home the point that perhaps the experience of subjectivity is just fundamental to the universe. More on this later.

31. Can we see an example? Is the binding problem hard or easy?

According to [TPoC](#),

- Binding is the process whereby separately represented pieces of information about a single entity are brought together to be used by later processing, as when information about the color and shape of a perceived object is integrated from separate visual pathways.
- Crick and Koch hypothesize that binding may be achieved by the synchronized oscillations of neuronal groups representing the relevant contents. When two pieces of information are to be bound together, the relevant neural groups will oscillate with the same frequency and phase.
- Such a theory would be valuable, but it would tell us nothing about *why* the relevant contents are experienced. ... Even if it is accepted, the explanatory question remains: *Why* do the oscillations give rise to experience?

So, the binding problem is unsolved for now, but it is still easy. This passage perfectly illustrates how Chalmers uses the question of *why* to keep the hard problem out of reach.

32. How have people tried to answer the hard problem?

According to [TPoC](#),

- In placing this sort of work with respect to the problem of experience, a number of different strategies are available. It would be useful if these strategic choices were more often made explicit.
- The first strategy is simply to *explain something else*. Some researchers are explicit that the problem of experience is too difficult for now, and perhaps even outside the domain of science altogether.
- The second choice is to take a harder line and deny the phenomenon. According to this line, once we have explained the functions such as accessibility, reportability, and the like, there is no further phenomenon called “experience” to explain.
- In a third option, some researchers *claim to be explaining* experience in the full sense. These researchers (unlike those above) wish to take experience very seriously; they lay out their functional model or theory and claim that it explains the full subjective quality of experience (e.g., Flohr 1992 Humphrey 1992). The relevant step in the explanation is usually passed over quickly, however, and usually ends up looking something like magic.
- A fourth, more promising approach appeals to these methods to *explain the structure of experience*. ... At best, it takes the existence of experience for granted and accounts for some facts about its structure, providing a sort of nonreductive explanation of the

structural aspects of experience (I will say more on this later). This is useful for many purposes, but it tells us nothing about why there should be experience in the first place.

- A fifth and reasonable strategy is to *isolate the substrate of experience*. ... the strategy is clearly incomplete. For a satisfactory theory, we need to know more than *which* processes give rise to experience; we need an account of why and how.

I agree with Chalmers that these strategies do not answer his hard problem. But that is, of course, because he has probably placed it out of reach with his infinite regression of why questions. Still, it is interesting to see the various strategies that have been employed so far by people who don't seem to fully grasp what Chalmers is getting at. I don't believe the theory I have developed in this series misunderstands Chalmers' hard problem, however, nor does it 1) explain something else, 2) deny the phenomenon, 3) pass over it like magic, 4) take it for granted, or 5) assume it is isolated to one substrate.

33. So, what else is needed and why do physical accounts fail?

According to [TPoC](#),

- To account for conscious experience, we need an *extra ingredient* in the explanation. This makes for a challenge to those who are serious about the hard problem of consciousness: What is your extra ingredient, and why should *that* account for conscious experience?
- At the end of the day, the same criticism applies to *any* purely physical account of consciousness. For any physical process we specify there will be an unanswered question: Why should this process give rise to experience?
- A physical account can *entail* the facts about structures and functions: once the internal details of the physical account are given, the structural and functional properties fall out as an automatic consequence. But the structure and dynamics of physical processes yield only more structure and dynamics, so structures and functions are all we can expect these processes to explain.

This is the same issue for all fundamental properties of the universe. We don't know *why* matter, gravity, or electromagnetism exist and behave the way that they do. One cannot get outside of all frames of reference to understand what is going on inside them. To paraphrase the eco-philosopher Arne Næss, one cannot blow a balloon up from the inside. This appears to be the same issue for explaining the subjective phenomena of consciousness. It just seems to happen in all living things, and my theory of *pandynamism* explains why this might be so for us, but not be so for non-living things.

34. Is this the same problem we faced with vitalism?

According to [TPoC](#),

- This might seem reminiscent of the vitalist claim that no physical account could explain life, but the cases are disanalogous. ... Once it turned out that physical processes could perform the relevant functions, vitalist doubts melted away. ... With experience, on the other hand, physical explanation of the functions is not in question. The key is instead the

conceptual point that the explanation of functions does not suffice for the explanation of experience.

Chalmers is right that the phenomenon of life (the fact that living beings act as living beings) is an objective observation that can be explained away once the mechanisms of life are understood. The internal subjective feeling of consciousness is not like this. There is an abundance of evidence for subjectivity in living organisms, as explained in question 8 above about zombies, but it is not an obvious phenomenon from the outside and we certainly cannot crawl into another's physical embodiment to truly know "what it feels like" to be them. Still, understanding the physical processes of life melted away any thoughts of extra non-physical ingredients for life. And similarly, understanding the physical processes for all aspects of consciousness in my comprehensive hierarchy is melting away any thoughts for any extra non-physical ingredients for consciousness. What is left? Just the simplest observations that subjectivity *does* occur in living things, and it does *not* appear to occur in non-living things.

35. So, is consciousness just fundamental?

According to [TPoC](#),

- Although a remarkable number of phenomena have turned out to be explicable wholly in terms of entities simpler than themselves, this is not universal. In physics, it occasionally happens that an entity has to be taken as *fundamental*. Fundamental entities are not explained in terms of anything simpler. Instead, one takes them as basic, and gives a theory of how they relate to everything else in the world. For example, in the nineteenth century it turned out that electromagnetic processes could not be explained in terms of the wholly mechanical processes that previous physical theories appealed to, so Maxwell and others introduced electromagnetic charge and electromagnetic forces as new fundamental components of a physical theory. To explain electromagnetism, the ontology of physics had to be expanded. New basic properties and basic laws were needed to give a satisfactory account of the phenomena.
- Other features that physical theory takes as fundamental include mass and space-time. No attempt is made to explain these features in terms of anything simpler. But this does not rule out the possibility of a theory of mass or of space-time. There is an intricate theory of how these features interrelate, and of the basic laws they enter into. These basic principles are used to explain many familiar phenomena concerning mass, space, and time at a higher level.
- I suggest that a theory of consciousness should take experience as fundamental. We know that a theory of consciousness requires the addition of *something* fundamental to our ontology, as everything in physical theory is compatible with the absence of consciousness.

I agree with Chalmers that the subjective feeling of consciousness is fundamental in this way. But, as explained above, it does not arise in non-living matter because there is no structure there that constitutes a *subject*, which could then experience *subjectivity*. Our physical theories are compatible with the *reaction* of all physical matter to physical forces. But Chalmers is wrong about our biological observations. Those require something else to explain the *actions* that living organisms take. (See question 8 above for a discussion of the preposterousness of non-conscious zombies.) Defining consciousness as I have ("an infinitesimally growing ability to sense and respond to any or all biological forces in order to meet the needs of survival"),

and then explaining what these biological forces are, and how *pandynamism* gave rise to feeling them, gives us a coherent physical explanation for all of our observations—both the objective ones and subjective ones, in physics, chemistry, and biology.

36. If we accept consciousness is fundamental, then what?

According to [TPoC](#),

- We might add some entirely new nonphysical feature, from which experience can be derived, but it is hard to see what such a feature would be like. More likely, we will take experience itself as a fundamental feature of the world, alongside mass, charge, and space-time. If we take experience as fundamental, then we can go about the business of constructing a theory of experience.
- Where there is a fundamental property, there are fundamental laws. A nonreductive theory of experience will add new principles to the furniture of the basic laws of nature. These basic principles will ultimately carry the explanatory burden in a theory of consciousness.
- Just as we explain familiar high-level phenomena involving mass in terms of more basic principles involving mass and other entities, we might explain familiar phenomena involving experience in terms of more basic principles involving experience and other entities.
- Of course, by taking experience as fundamental, there is a sense in which this approach does not tell us why there is experience in the first place. But this is the same for any fundamental theory. Nothing in physics tells us why there is matter in the first place, but we do not count this against theories of matter. Certain features of the world need to be taken as fundamental by any scientific theory. A theory of matter can still explain all sorts of facts about matter, by showing how they are consequences of the basic laws. The same goes for a theory of experience.
- Nothing in this approach contradicts anything in physical theory; we simply need to add further *bridging* principles to explain how experience arises from physical processes. There is nothing particularly spiritual or mystical about this theory—its overall shape is like that of a physical theory, with a few fundamental entities connected by fundamental laws. It expands the ontology slightly, to be sure, but Maxwell did the same thing.

Yes! This is the route I have taken, and I have started to sketch these new principles and fundamental laws of pandynamism and biological forces.

37. Is this fundamental view a sort of dualism?

According to [TPoC](#),

- In particular, a nonreductive theory of experience will specify basic principles telling us how experience depends on physical features of the world. These *psychophysical* principles will not interfere with physical laws, as it seems that physical laws already form a closed system. Rather, they will be a supplement to a physical theory. A physical theory gives a theory of physical processes, and a psychophysical theory tells us how those processes give rise to experience.

- This position qualifies as a variety of dualism, as it postulates basic properties over and above the properties invoked by physics. ... If the position is to have a name, a good choice might be *naturalistic dualism*.

Except that it's not dualism! There isn't a dualism of matter + space-time + electromagnetism + any other fundamentals of physics. It's all just the list of properties in a monist physical universe. Adding subjectivity as a fundamental feeling that emerges in physical material once that material attains the form of self-sustaining life does not change this monistic view.

Furthermore, Chalmers is right that “a *physical theory* gives a theory of *physical processes*” but he is wrong about what a *psychophysical theory* then gives us. To extend the comparison logically, a *psychophysical theory* gives us ...wait for it... a theory of *psychophysical processes*! That is exactly what my theory of biological forces helps us to understand—the *psychophysical processes* going on in living organisms, which drives their actions over and above the simple reactions of the physical and chemical laws of nature. If subjective consciousness is truly taken as fundamental, there is no need to “tell us how processes give rise to experience.” That's fundamental!

38. If consciousness is fundamental, shouldn't it be simple to describe?

According to [TPoC](#),

- If this view is right, then in some ways a theory of consciousness will have more in common with a theory in physics than a theory in biology. Biological theories involve no principles that are fundamental in this way, so biological theory has a certain complexity and messiness to it; but theories in physics, insofar as they deal with fundamental principles, aspire to simplicity and elegance. The fundamental laws of nature are part of the basic furniture of the world, and physical theories are telling us that this basic furniture is remarkably simple. If a theory of consciousness also involves fundamental principles, then we should expect the same. The principles of simplicity, elegance, and even beauty that drive physicists' search for a fundamental theory will also apply to a theory of consciousness.
- Finally, the fact that we are searching for a fundamental theory means that we can appeal to such nonempirical constraints as simplicity, homogeneity, and the like in developing a theory. We must seek to systematize the information we have, to extend it as far as possible by careful analysis, and then make the inference to the simplest possible theory that explains the data while remaining a plausible candidate to be part of the fundamental furniture of the world.

Yes! I think my theory is pretty simple. I'm glad that is a feature and not a bug.

39. What about Chalmers' own theories?

According to [TPoC](#),

- In what follows, I present my own candidates for the psychophysical principles that might go into a theory of consciousness. The first two of these are *nonbasic principles* — systematic

connections between processing and experience at a relatively high level. These principles can play a significant role in developing and constraining a theory of consciousness, but they are not cast at a sufficiently fundamental level to qualify as truly basic laws. The final principle is my candidate for a *basic principle* that might form the cornerstone of a fundamental theory of consciousness.

- The principle of structural coherence: this is a principle of coherence between the *structure of consciousness* and the *structure of awareness*. ... If we accept the principle of coherence, the most *direct* physical correlate of consciousness is awareness: the process whereby information is made directly available for global control. ... This principle reflects the central fact that even though cognitive processes do not conceptually entail facts about conscious experience, consciousness and cognition do not float free of one another but cohere in an intimate way.
- The principle of organizational invariance: this principle states that any two systems with the same fine-grained *functional organization* will have qualitatively identical experiences. If the causal patterns of neural organization were duplicated in silicon, for example, with a silicon chip for every neuron and the same patterns of interaction, then the same experiences would arise.
- The double-aspect theory of information: I understand information in more or less the sense of Shannon (1948). Where there is information, there are *information states* embedded in an information space. An *information space* has a basic structure of *difference* relations between its elements, characterizing the ways in which different elements in a space are similar or different, possibly in complex ways. ... To borrow a phrase from Bateson (1972), physical information is a *difference that makes a difference*. The double-aspect principle stems from the observation that there is a direct isomorphism between certain physically embodied information spaces and certain *phenomenal* (or experiential) information spaces.

Regarding Chalmers' first principle of structural coherence, this is so typical of a philosopher to focus on such a high level of conscious awareness rather than starting at the bottom of consciousness. Chalmers says "the most *direct* physical correlate of consciousness is awareness: the process whereby information is made directly available for global control" but by tracing the evolutionary history of consciousness, we see that this comes far after all the cognitive abilities in my hierarchies of affect, intention, and prediction. In fact, I would go so far as to say that awareness *can only* arise after these other abilities are present. I agree that "consciousness and cognition do not float free of one another," so the general principle of structural coherence is fine, but you have to do a Tinbergen analysis to see *all* of the cognitions that are built into consciousness. And this affects Chalmers' other theories.

The second principle of organizational invariance is very hard to accept given the impact that tiny bits of chemical drugs can have on our conscious experience. The matter seems to matter! Perhaps the carbon lifeforms that have slowly, slowly arisen over the billions of years of Earth's evolutionary history have found their way here precisely because their structure yields experiences that drive towards survival and away from extinction. Maybe a silicon-based replica would love the feeling of electricity coursing through its body too much and would quickly zap itself into oblivion like a moth to a flame. We certainly don't know that, but it seems just as possible as Chalmers' speculation. And given the fact that no other substrates for life *have* arisen here, it seems more likely that functional organization is not enough for "qualitatively identical experiences."

Finally, I see Chalmer's basic principle about information as a simple truism. Yes, there is "a direct isomorphism between certain physically embodied information spaces and certain *phenomenal* (or experiential) information spaces." But this is exactly because the universe is physical. Any changes in experience *are* associated with physical changes. And both of these can be expressed as information. But information can be abstracted from everything! There isn't anything that follows from this about information itself. More on this in the next question.

40. Is consciousness all about information processing?

According to [TPoC](#),

- This [basic principle] leads to a natural hypothesis: that information (or at least some information) has two basic aspects, a physical aspect and a phenomenal aspect. This has the status of a basic principle that might underlie and explain the emergence of experience from the physical. Experience arises by virtue of its status as one aspect of information, when the other aspect is found embodied in physical processing.
- If the principle of organizational invariance is to hold, then we need to find some fundamental *organizational* property for experience to be linked to, and information is an organizational property *par excellence*.
- Wheeler (1990) has suggested that information is fundamental to the physics of the universe. According to this "it from bit" doctrine, the laws of physics can be cast in terms of information, postulating different states that give rise to different effects without actually saying what those states *are*. It is only their position in an information space that counts. If so, then information is a natural candidate to also play a role in a fundamental theory of consciousness. We are led to a conception of the world on which information is truly fundamental, and on which it has two basic aspects, corresponding to the physical and the phenomenal features of the world.

To me, this has it exactly backwards. The bit comes from the it! Information is just abstraction from the physical world. And I found abstraction to be the final level of consciousness that emerges in my hierarchy. The ability to have abstract thoughts (i.e., the ability to represent the physical with language and symbols) is what gives a consciousness the freedom to think infinitely far and wide, formulate imagined hypotheses about the world, and communicate our thoughts about all of these thoughts. Only physicists, mathematicians, and philosophers who spend their whole lives in this abstract world could actually think that this is the primary cause of reality. We must not follow them down this hole.

In [a long Psychology Today post](#) about the spirituality of [Integrated Information Theory](#), we see the trouble this leads people into. The author there said,

- "Let's follow the logic of this idea and see how it holds up. We know that certain brain states feel like something. Brain states are just information states. Therefore, information feels like something. Sounds pretty solid.

No! Brain states are not *just* information states. They are *specific* information processors that are processing *specific* information. Information and information states are everywhere because everything in reality (and in imagination!) can be abstracted. One could similarly

claim that boulders are *just* information states, therefore information feels like *nothing*. The bad logic is the same! So, contrary to what Chalmers states, information is not an *organizational property par excellence*. Information is that map in a joke which says “**Scale: 1 mile = 1 mile.**”

Chalmers says, “Experience arises by virtue of its status as one aspect of information, when the other aspect is found embodied in physical processing.” I find it much easier to say that experience arises from physical processing, but this subjectivity only emerges and grows along with the emergence of subjects as the requisite structures form that can capture and register these experiences.

41. So, can we make progress and answer the hard problem of consciousness?

According to **TPoC**,

- Most existing theories of consciousness deny the phenomenon, explain something else, or elevate the problem to an eternal mystery. I hope to have shown that it is possible to make progress on the problem even while taking it seriously. To make further progress, we will need further investigation, more refined theories, and more careful analysis. The hard problem is a hard problem, but there is no reason to believe that it will remain permanently unsolved.

Yes, but then no. We can indeed make progress on all the easy problems of consciousness, and that appears to melt away the hard problem in the same way that a hard problem of electromagnetism melted away. Given that all the building blocks of consciousness are widespread across the entirety of life throughout its evolutionary history, it is likely that the experience of subjectivity is a fundamental property of our universe. But any further questions about *why* this universe, or all universes, are like that appear to be permanently unsolvable. One cannot always get outside of one’s frame of reference in order to understand everything within that frame. **Godel’s incompleteness theorems** and **Tarski’s undefinability theorem** are good examples of this principle. But hopefully, once answers like the ones I have proposed to these frequently asked questions about consciousness have been developed, debated, and widely accepted, then the hard problem of consciousness will no longer be considered any more of a mystery than gravity.