Substrate Flexibility and the Copernican Principle of Consciousness

Jeremy Pober

Centre of Philosophy, University of Lisbon

(corresponding author: Jeremy.pober@gmail.com)

Eric Schwitzgebel

Department of Philosophy, University of California, Riverside

0. Abstract

We present a novel argument for the substrate flexibility of consciousness – that is, for the idea that conscious experiences can arise in a variety of different types of physical media, not just in biological animals as they currently exist on Earth. Some recent critiques of standard arguments for the substrate flexibility of consciousness (e.g., Cao 2022; Block 2025; Seth forthcoming) have emphasized that *humanlike* consciousness might require our specific biological substrate. However, such critiques are too narrowly focused to address the issue of consciousness in entities whose experience may be very different from ours, for example alien life forms or future AI systems designed along unfamiliar lines. Given that it's likely that functionally complex, behaviorally sophisticated entities have arisen or will arise many times in the observable universe, in diverse substrates, we argue that it would be a violation of a principle of Copernican mediocrity to hold that among these diverse entities, only we, or only we and a small proportion of others who share our substrate, are conscious.

1. Introduction

Who is conscious? We—both authors and you, the reader—are. Some non-human animals presumably are (Trestman, Birch, and Allen 1995/2026; Birch 2024). Likely, as we will argue, some extraterrestrial entities ("aliens") are. We don't think there's good evidence that current technology has produced conscious artifacts (Artificial Intelligence, or "AI"), though we don't rule out that possibility in the future (Schwitzgebel and Pober 2026; Schwitzgebel forthcoming).

Conceptually prior to "who is conscious?" is the question "what kinds of things can be conscious?" In this article, we assume physicalism (or materialism), which holds that all conscious entities are fundamentally composed of matter. We're generally inclined to think that only certain kinds of matter, organized in very specific and complex ways, can give rise to consciousness, though we won't directly argue for this claim here. What we *will* argue for is that consciousness can be realized in very different kinds of physical arrangement—different substrates. Consciousness is *substrate flexible.*

After defining substrate flexibility in Section 2, we offer the following three-premise argument in sections 3, 4, and 5, respectively:

> P1: The universe contains at least a thousand different behaviorally sophisticated species.
>
> P2: These species have substantially different substrates.
>
> P3: Some of these species with substantially different substrates are conscious.
>
> Conclusion: Therefore, consciousness is possible in substantially different substrates.

P1 is justified on cosmological grounds; P2 from astrobiological and biochemical considerations concerning the possible substrates of life, and P3 from the Copernican Principle of Consciousness (Schwitzgebel and Pober 2026). In S6 we discuss how our argument relates to the current debate on substrate flexibility, in particular the best-known argument against it, advanced by Peter Godfrey-Smith (2016, 2024), Rosa Cao (2022), and Ned Block (2025). Godfrey-Smith, Cao, and Block aim to undercut the "neural replacement" argument for substrate flexibility found in Tom Cuda (1985) and David Chalmers (1996). As we will argue, although these authors convincingly show that the precise duplication of humanlike cognitive function, and thus (presumably) humanlike consciousness, is unlikely to be possible in a substrate different from our own, their argument does not establish that *non*-humanlike consciousness is unlikely in a non-humanlike substrate. Section 7 extends our reasoning to AI cases.

2. Substrate Flexibility

Glass is a different substrate than ceramic. Since cups can be made of either glass or ceramic, cups can be realized in different substrates. Computer hard drives can be solid state, optical, or magnetic—different types of physical substrate for the long-term storage of computer information. In this section, we articulate an account of substrate flexibility, which also serves as an account of substrate dependence and substrate independence.

In our way of speaking a *target phenomenon* has, and is realized by, a physical *substrate.* Target phenomena may be properties, events, states, or entities, as may their substrates, though the categories need to correspond, i.e., properties are realized in properties, states in states, and so on (Adams 1979). We will talk of target and substrate properties as a general term except where differences between properties and states or entities matter for our argument. We also assume for ease of exposition that target and substrate properties are property kinds and not property instances, though one can sensibly use target/substrate/realizer talk for either.

A target property is *substrate flexible* to the extent it constrains its possible substrates. In theory, if it completely constrains the space of possible substrates, such that there is only one possible substrate property, it is completely substrate dependent. If it places no constraints whatsoever, so that any property may realize it, it is completely substrate independent. If the property places some constraints on its substrates, such that it lies between these two extremes, it is substrate flexible.

Substrate flexibility is a matter of degree. Being a standard-isotope hydrogen atom is minimally substrate flexible in the sense that it can be realized only in a very narrow range of physical properties. Being a cup or a hard drive is much more substrate flexible. However, they are not completely substrate independent, since a cup cannot be made of air nor a hard drive of cheese (*pace* Putnam [1975]).

Crucially, the degree to which any property is substrate flexible depends how *coarsely* or *finely* the target property is *specified*. The more finely a target property is specified, the less substrate flexible it is likely to be. The property of being a corkscrew, specified coarsely, is at least moderately substrate flexible, whereas the property of being a corkscrew made of molded plastic in such-and-such a configuration, specified more finely, is much less substrate flexible.

Properties can be specified more coarsely or finely along multiple dimensions. One such dimension pertains to the mereological parts, or mereological parts of parts (or parts of parts of parts, etc.) of a bearer of the target property, and appealing to properties of those parts; doing so is often described as giving a lower 'level' of description. Here the difference between properties and other phenomena (states, entities, and events) matters. For example, when discussing an *entity*, such as water, more finely specified entities are parts of the entity, e.g., atoms in the H_2O molecule. But when discussing the *property of being water*, properties of more finely specified parts are not parts of the property of being water (neither properties of parts nor the property of being a part are parts of properties).[1]

Nonetheless, as our examples have already shown, specifying properties more finely need not involve appealing to mereological proper parts. We can sensibly speak of the substrate of a cup as the whole lump of matter comprising it.

In our way of speaking, a target property is *realized* in a substrate property. We borrow this term from the literature on 'multiple realization', which focuses on the physical realizers (what we call substrates) of functional mental states (Putnam 1967; Fodor 1974; Kim 1992; Shapiro 2000).[2] However, there are three differences between our use of the term and its typical use in this literature. First, as we have already noted, the class of phenomena that is or can be realized is far broader than functional properties as standardly conceived. Second, while functional properties can be realized in either physical or functional properties (Shoemaker 1981; Pober, under review), substrates must be physical (including kinds of physical properties). And third, there is an issue within literature about whether properties with low substrate flexibility do or

---

[1] Discussion of mereological parts of different sizes, or after a different number of decompositions from the target entity, are often discussed in terms of different 'levels of reality' (Ioannidis et al. 2022). However, we eschew such talk here, which we take to be unnecessary for—and problematic for reasons unrelated to—the current discussion.

[2] With respect to other issues in the metaphysics of realization, we aim to be as ecumenical as possible. As discussed with the example of cups, we allow for the realization relation to hold between phenomena and substrates along "flat" and "dimensioned" lines (Gillet 2003), that is, respectively, wherein a property and its substrate property obtain of the same entity or state, or wherein the substrate property obtains of a proper part, entity, or state. We are indifferent as to whether the realization relation is identity, constitution, or some other *sui generis* relation (Morris 2018) and indeed accept that there may be multiple classes of relations which count as 'realization', e.g., physical phenomena may be identical to their substrates whereas functional states may stand in a 'role-filling' (Antony and Levine 1997) relation to their physical substrates.

do not not count as genuinely "multiply realizable" (Shapiro and Polger 2016; Piccinini 2020). Shapiro (2000), for example, has argued that although single-handled ("waiters") and double-handled ("flywheel") corkscrews count as multiple realizations of the functional kind corkscrew, otherwise identical corkscrews made of differently colored molded plastics count not as "multiple realized" but only as "variably realized". In our terminology, the property of *being a corkscrew made of molded plastic in such-and-such a configuration* is substrate flexible, albeit to a low degree.

Substrate flexibility as we have defined it is cheap, and indeed, a property's being *completely* substrate dependent is, on our conception, an extreme case. We cast our argument in terms of "substantial" substrate flexibility (P2 and P3 in S1). Life is substantially substrate flexible, in our sense, if it can arise in substrates very different from those found among terrestrial organisms. How different is "very different"? We suggest that it would be enough for the entities to have a chemical composition fundamentally different than any we have so far found on Earth—for example, being composed of different "chemical kinds" in the sense of Needham (2003, 2005). Entities are of different chemical kinds if they are composed of different inorganic molecules or different organic monomers, that is, different organic subcomponents on the scale of amino or nucleic acids (Berg et al. 2018; this allows for various polymers of the same subcomponents, such as polyesters, to count as the same chemical kind). Despite all being carbon based, ethyl alcohol, polyester, and caffeine are different types of substance, different chemical kinds (as can readily be discerned by the effects of ingestion). If we were to discover conscious entities whose basic chemical building blocks differ from ours, we could reasonably say that they differ substantially from us in substrate.

3. Behaviorally Sophisticated Extraterrestrial Life

Each of the next three sections defends one of our three premises. Here, we defend the first premise, that there are at least a thousand of what we call 'behaviorally sophisticated' extraterrestrial species across the span—and lifespan—of the universe. This premise follows from astrobiological estimates of 'technological civilizations' and our own account of behavioral sophistication (Schwitzgebel and Pober, 2026).

Cosmologists and astrobiologists generally estimate measurable characteristics, among which are "technological civilization" and "technosignatures" (Frank and Sullivan 2016; SETI 2021; Wright et al. 2022). The idea is that an alien civilization with technology like our own is potentially observable, for example if it broadcasts radio-frequency signals or builds observable megastructures, such as a gigantic habitat or energy collector around a star.

How common are technological civilizations? While some astrobiologists think it's possible that Earth contains the only one in the approximately one trillion galaxies that currently form the observable portion of the universe, estimates are typically several orders of magnitude higher than that (e.g., Frank and Sullivan 2016). One recent survey found median scientific estimates over one civilization per galaxy at some point in that galaxy's lifetime (Snyder-Beattie et al. 2021)—low enough to explain the "Fermi Paradox" (the question of why we haven't yet seen evidence of technological civilizations) without making technological civilizations *extremely* rare. For purposes of the present argument, the existence of at least a thousand technological civilizations scattered in time and space across the observable universe—one per billion galaxies, or 0.000000001% of the median scientific estimate—is more than sufficient; we adopt this extremely conservative estimate.

Developing a technological civilization is related to what we call behavioral sophistication. An entity is behaviorally sophisticated, in our sense, if and only if it is capable of complex goalseeking, complex communication, and complex cooperation, where those terms describe behavioral patterns without presupposing consciousness or any particular internal architecture.

Suppose an alien species constructs intricate devices that extract nutrition from its environment. These nutrition sources are stored and consumed by species members as needed. Individuals respond to environmental threats with highly specific evasive activities, often well in advance and with narrow margins for error. The aliens interact in highly complex patterns, engaging in activities that increase their chances of survival or reproduction only if several others engage in specific activities at specific times while physically remote and not in direct contact. This interaction is enabled by complex signals with flexible contents and a generative grammar. Such a species would be behaviorally sophisticated as we intend the phrase.

Plausibly, all or almost all technological civilizations would be behaviorally sophisticated, while some behaviorally sophisticated species would not be technological (for example, if humans had not advanced past first-century technology). It is thus reasonable to assume that if at least a thousand technological civilizations have existed or will exist, and if these civilizations arose independently (for example, in different galaxies), then at least a thousand behaviorally sophisticated species have evolved or will evolve.

Behavioral sophistication comes in degrees. For simplicity, we assume approximately human-level behavioral sophistication. However, as we will explain later, our argument from the Copernican Principle to substrate flexibility can be generalized for any threshold of behavioral sophistication above which almost all Earthly animals are conscious.

4. Behaviorally Sophisticated Life is Substrate Flexible

Our argument for the substrate flexibility of behaviorally sophisticated life has two parts. First, we show that life itself is substrate flexible with respect to biochemical properties and kinds. Second, we contend that the evolutionary bottlenecks between life and behaviorally sophisticated life do not add significant further restrictions on biochemical substrates. Because we invoke (bio-)chemical property kinds, our argument will suffice to show that life is substantially substrate flexible. (For ease of exposition, we drop the explicit talk of properties when discussing chemical substrates, that is, we talk of substrates *as* chemical kinds rather than as having the property of being made of such-and-such chemical composition.)

4.1. The Substrates of Life

Philosopher David Lewis (1980) imagined Martians whose cognition operates by the inflation of hydraulic sacs. Science fiction writer Greg Egan (1997) imagined a life form consisting of a single, planet-sized molecule. Scientists have hypothesized about the possibility of life based in sulfur (Keller et al. 2017) or organoborates (Petkowski, Bains, and Seager 2020; Grefensette et al. 2024), or employing alternative solvents, such as ammonia (Raulin et al. 1995) or methane (McKay 2016; see Grefensette et al. 2024 for discussion of all examples).

There are three possible ways a biochemical kind can vary: evidence suggests that the substrates of extraterrestrial life will vary in at least two of them. The first type involves monomers made of the same set of overall elements, but put together in a different combination than the twenty amino or five nucleic acids we use. The second and third types involve distinct kinds of differences in the chemical composition of the monomers. Biochemical molecules require two types of atoms: a 'scaffolding' atom which can form long, repeating chains, and 'heteroatoms' to give different parts of the long chain molecule (different monomers of the polymer) functional diversity. Terrestrial organisms use carbon as the scaffolding atom and primarily oxygen, nitrogen, phosphorus, and sulfur as heteroatoms. The second and third types of variation involve the use of other elements as heteroatoms or as the scaffolding atom.

Restricting ourselves to elements used by human life, the space of possible amino acids is far greater than the twenty we use: using only the aforementioned heteroatoms plus carbon, there are 10^{54} possible amino acid 'alphabets' (Brown, Voracek, and Freeland 2023). Terrestrial life did not choose the twenty we wound up with by chance: there are functional properties that make this set nearly optimal (Ilardo et al. 2015). Specifically, the set covers a wide range of sizes and hydrophobic states: hydrophobia is "the driver of protein folding" and volume "determines … constraints that limit protein formation" (Brown, Voracek, and Freedland 2023, 537).[3] In other words, the set provides among its members the properties needed for diversity of polymer structure and function. Yet given the astounding number of possible amino acid alphabets, ours is not alone in being optimal with respect to certain functional properties: Brown and colleagues (ibid.) found thousands of other 'alphabets' that differ in at least one amino acid in what they consider a non-exhaustive search.

Yet life may also use other elements as heteroatoms. Terrestrial life does not use silicon in natural circumstances, yet Arnold and colleagues (Kan et al. 2016; Arnold 2018) have been able to induce *Rhodothermus marinus* bacteria to use their cytochrome c protein to synthesize silicon-carbon bonds that would be the prerequisite for using silicon as a hetereoatom. Further, other, non-terrestrial environments may be *more* amenable to the use of silicon as a heteroatom (Petkowski, Bains, and Seager 2020; Bains, Petkowski, and Seager 2024a).

To determine how effective a heteroatom an element would be in a non-terrestrial environment, scientists examine how much of its *chemical space* is *available* in that environment. An element's available chemical space is determined by two factors: the overall chemical space of that element, or the set of all possible bonds it can form across any environment, and the

---

[3] The same is likely true for the nucleic acids, which only use five combinations (four each in DNA and RNA), though we do not know of any studies which have investigated the issue.

environmental factors which limit how much of that chemical space is available. Thus, available chemical space measures the range of stable chemical bonds an atom of it can form with other elements in that environment (Lipinski and Hopkins 2004; Bains and Seager 2012). Environments consist of multiple factors such as temperature and pressure, and most chemical reactions which form bonds are only possible in a subset of temperature/pressure settings. Another aspect of the environment is the *solvent* or primary fluid in which the element in question is embedded, i.e., the primordial solvent for life on earth was water (Bains 2004; Schulze-Makuch and Irwin 2018; Bains, Petkowski, and Seager 2024b).[4]

The range of stable chemical compounds and reactions can vary considerably among environments. For example, in environments where the solvent is sulfuric acid rather than water—a situation found in our solar system in the Venusian gas clouds and likely widely throughout the universe (Ballesteros et al. 2019; Bains, Petkowski, and Seager 2024a)—the available chemical space of most elements other than silicon is vastly constricted (relative to their available chemical spaces in water), while the available chemical space of silicon is greatly expanded (Petkowski, Bains, and Seager 2020). Sulfuric acid environments might make the energy cost of bonding prohibitively high for familiar terrestrial heteroatoms oxygen, nitrogen, phosphorus, and sulfur, while facilitating the use of silicon as a heteroatom. The result would be life with a carbon scaffold, but silicon as a widely used heteroatom and sulfuric acid as the solvent. If somewhere in the vast universe there is life that employs silicon-based heteroatoms in a sulfuric acid—with all the downstream structural changes that would surely entail—life is substantially substrate flexible.

---

[4] The solvent must be a fluid: as Bains explains, "chemical life is dependent on its molecular components being suspended in a fluid of comparable density of the macromolecules" (Bains 2004, 139).

Unfortunately for silicon enthusiasts, Petkowski and colleagues argue that "silicon-based life", i.e., life using silicon as a scaffolding atom for its polymers is "almost certainly impossible" (Ibid., 23). In short, they contend, the issue is that the conditions required for silicon to form polymer chains, cryogenic temperatures are mutually exclusive with the high-temperature conditions required for silicon to form monomers with heteroatoms (Ibid). However, synthetic protein analogues made of entirely different material, using metal oxides as the scaffolding and heteroatoms have been hypothesized (Li, Zhao, and Liu 2024).

While silicon's involvement in the biochemical basis of life has received the most attention (e.g., Bains 2004; Benner et al. 2004; Petkowski, Bains, and Seager 2020), if any chemically different systems can realistically evolve, then life is highly substrate flexible—and maybe also consciousness, if such life could develop consciousness.

.

4.2. From Life to Behaviorally Sophisticated Life

More has been said about the chemical basis of life than about the physiological basis of complex life, and the little that has been said about the latter discusses it from an evolutionary, substrate-independent perspective (e.g., Kershenbaum 2021). Yet we can ground our discussion by asking what evolutionary constraints behaviorally sophisticated life places on the space of possible substrates in addition to those posed by life itself. We suggest that there are two basic types of requirements for the evolution of behaviorally sophisticated life which might plausibly yield constraints on their substrates. We conclude that while these requirements may pose constraints on cellular substrates (i.e., substrate properties whose bearers are cells), neither offers reason to constrict the space of possible biochemical substrates.

The first constraint comes from the need for a behaviorally sophisticated organism to be large: on a scale of billions of cells (or their equivalent). Not only does the organism need to be massively multicellular, the cells need to exhibit specialization for various biological functions. This requirement appears to impose no constraints on chemical compounds involved in the cells, beyond the that the compounds have the functional flexibility necessary for life (such as carbon chains with a diversity of heteroatoms in a good solvent, as described in S4.1).

When we shift from specifying substrates in terms of molecules and monomers to specifying substrates in terms of cells, some alien life forms might be incapable of transitioning to massive multicellularity. For example, some life may use nonlipid membranes or lack membranes entirely, forming in moisture droplets or porous minerals (Grefensette et al. 2024). Such life would plausibly be unable to scale up to massive multicellularity. However, this restriction of potential substrate properties specified at a cellular grain does not imply any restriction on substrates at a biochemical grain.

A second plausible constraint is that the organism must develop a specialized type of cell for coordinating the various specialized cells throughout its body, a 'control system' (Bechtel 2012). More specifically, it must develop "some kind of central biological control system that can facilitate and manage all of the internal bodily processes … needed to keep them alive and well in their current environment, as well as successfully respond to any threats to their survival that may result from changes in their immediate external (or internal) environment" (Cranston 2015, 92). (Note that this requirement may plausibly be understood as the physiological basis for something like a mind or brain, though we do not insist on it). To achieve an organ with such a function, it needs to develop a cell type optimized for sending and receiving signals with other cells. Crucially, this signaling ability needs to be adaptable at the cellular level: it needs to be able to respond differently to the same stimulus over time in order for the organism to adapt to—and eventually learn about—its environment.

Our neurons are the cells in our body specialized for complex, coordinated, intercellular communication, and, accordingly, our brains are our control systems. Neurons have evolved to send and receive optimally in *their* environment inside (and as part of) our body: all the major adaptations in the development of our neurons, from their emergence in the first nervous systems (Munoz, Romanova, and Kohn 2021), through animal evolution (Verkhatsky et al. 2022), to the specific capacities of our cortical neurons (Galakhova et al. 2022) have involved enhancements to their signaling abilities. Yet signaling each other is hardly the purview of neurons alone. It is now widely accepted that bacteria, fungi, and plants can signal among and within each other without employing neurons (Koshland 1980; Lyon et al. 2021; Ciaunica et al. 2023), and indeed the cells neurons evolved from already had some signaling abilities (Munoz, Romanova, and Kohn 2021). In our environment, neurons have great adaptive advantages in (e.g.) speed and preservation of signal, but again we see no reason to suppose that there couldn't be fast, adaptive signaling in biochemical substrates different from our own, especially if they can propagate electrical signals. We thus see no reason to think that cognitive sophistication creates a bottleneck requiring that the biological substrate of cognitively sophisticated alien life must always be identical to our own.

Pulling together the ideas of S4.1 and S4.2, we suggest that it's unlikely that every behaviorally sophisticated species in the universe happened to evolve the same substrate. Even if our substrate is near-optimal given familiar environmental constraints (Ilardo et al. 2015), the space of possible substrates is large, and other substrates may be favored in other environments with different constraints (Bains 2004). In this context, the constraints posed by conditions in the Venusian gas clouds are an example of such environmental constraints, albeit an extreme one: the Venusian environment is almost certainly *more* different from ours than an environment would need to be to favor life developing in other chemical kinds.

5. The Copernican Principle of Consciousness

We now turn to our third premise, that we should believe at least some of the behaviorally sophisticated extraterrestrial life is conscious. The Copernican Principle of Consciousness is a special case of the Copernican Principle in cosmology. According to the general Copernican Principle, "the Earth is not in a central, specially favored position" (Bondi 1968, p. 13) or, alternatively, "We do not occupy a privileged position in the Universe" (Barrow and Tipler 1986, p. 1; cf. Scharf 2014). We have argued elsewhere (Schwitzgebel and Pober 2026) that Copernican Principles are best understood as default principles: they are defeasibly warranted and can be defeated by further information to the contrary. If astronomical data suggest that we *are* in a specially privileged position, that would nullify any warrant for the cosmological Copernican Principle.

Copernican Principles require a reference class with respect to which they claim we are not special, and it is important to specify this reference class properly. In some respects, we *are* in a special location. The surface of the Earth is special in its ability to support life, compared to most other locations within the Solar System and a radius of at least several dozen light years (just how special depends on the conditions of exoplanets we've yet to investigate sufficiently). For this reason, recent scientific applications tend to emphasize the principle's connection to the homogeneity and isotropy of the universe at large scales (Caldwell and Stebbins 2008; Clarkson, Bassett, and Lu 2008; Camarena, Marra, Sakr, and Clarkson 2022). It is at the scale of galaxies, or large parts of galaxies, that the cosmological Copernican Principle applies because at that scale we have no data to suggest it is false.

We extend the idea of the Copernican Principle of cosmology to consciousness, as follows:

> *The Copernican Principle of Consciousness:* Among behaviorally sophisticated entities, we are not specially privileged with respect to consciousness (Schwitzgebel and Pober 2026).

Just as it's reasonable to assume, pending counterevidence, that our large-scale spatial position is not exceptional, it's reasonable to assume, pending counterevidence, that our mentality is not special relative to the mentality of other behaviorally sophisticated entities. We do not occupy the center of the trillion-galaxy-wide consciousness-is-here map. Although our mentality is special relative to rocks, rocks are a reference class to which this Copernican reasoning doesn't apply, since we already know we're more conscious than they are. As with large-scale vs. small-scale spatial position, Copernican reasoning only applies absent counterevidence.

If the Copernican Principle of Consciousness is correct, we are not specially lucky to possess consciousness-instilling Earthiform biology while other behaviorally sophisticated entities' architectures leave them entirely nonconscious. Nor do we have especially *more* or *better* consciousness. Absent some reason to think we are special, we Earthlings would then be suspiciously, inexplicably lucky.

The Copernican Principle of Consciousness thus implies that we should believe at least some of the other behaviorally sophisticated species in the universe are conscious. It does not necessarily imply that all, or even most of them, are conscious. And it does not give us an exact threshold percentage, below which consciousness would count as rare within this class, but we can make some rough stipulations: We would be strikingly lucky to be conscious if only 2% of behaviorally sophisticated species were conscious, less so if 20% were.

Suppose, then, on grounds of Copernican mediocrity that at least 20%--at least 200—of the thousand-plus behaviorally sophisticated species in the universe are conscious. It is possible that every single one of the conscious species is composed of the same sets of low-level chemical types as we are, and all of those who differ from us in substrate are nonconscious. However, for two reasons, this seems unlikely.

First, a simplicity or symmetry principle suggests that any division between conscious and nonconscious behaviorally sophisticated entities should correlate with some important functional deficit in the nonconscious ones (Schwitzgebel and Pober 2026). We see no grounds to think that such a functional deficit would inevitably be present in behaviorally sophisticated species of any substrate other than our own.

Second, in the broader spirit of the Copernican principles, we should probably also accept as a default assumption that our particular substrate is not special. Our substrate is, of course, *somewhat* special: here on Earth, it alone supports behaviorally sophisticated, conscious life. But among other substrates in the universe that can support behaviorally sophisticated life and may be favored in environments unlike ours with different resources and constraints, we have no reason to believe that our substrate is uniquely capable of supporting consciousness. Without such a positive reason, believing our substrate special in this way would again violate a plausible principle of mediocrity.

6. Some Possible Objections

Having now supported each of the three premises we articulated in the introduction, in this section, we respond to two possible objections about the *relevance* of our claims.

6.1. Extant Discussion of Substrate Flexibility

The recent literature on substrate flexibility (e.g., Cao 2022; Block 2025; Seth forthcoming) focuses on the plausibility of fine-grained functional equivalents to human beings existing in a different substrate. We do not. As a result, one might worry that we are not discussing the same topic. In some senses—e.g., Seth's focus on the substrate flexibility of *computation*—perhaps we are discussing a different topic. However, on what we take to be the central issue—whether a conscious being can be realized in different stuff—we are not. Rather, we simply believe we offer a better strategy for conceptualizing the substrate flexibility of consciousness.

The literatures owes its current focus to the *neural replacement argument*, from Cuda (1985) and Chalmers (1996). Cuda and Chalmers imagine swapping one neuron at a time from a biological human brain, replacing each with a functionally identical silicon chip. If each silicon chip truly is functionally identical, then the entity at the end of full neural replacement will have neural structures made entirely of silicon chips, but functionally identical to an ordinary conscious human brain. Such a hypothetical entity would inevitably *report* being conscious and never having noticed a change. After all, if each silicon chip acts exactly like one neuron, then it will signal forward to other neurons or chips in exactly the way the original neuron would have, under exactly the same conditions. All motor outputs, including speech outputs, will thereby be just the same as they would be in the unmodified human. Plausibly, if we trust human introspective reports of consciousness, we should also trust the reports of no change in experience during the swapping procedure. Therefore, the argument concludes, we should allow that an entity with a silicon-chip based neural structure could be conscious. Substrate flexibility follows. (Actually, we think it's reasonable to doubt the mid-swap introspective reports—see Udell and Schwitzgebel 2021; Schwitzgebel 2022; Block 2023—but grant introspective infallibility for the sake of argument.)

Godfrey-Smith (2016, 2024), Cao (2022), and Block (2023) argue convincingly against a crucial premise of this argument: it would *not* be possible (that is, nomically possible, or consistent with the laws of nature) to swap a silicon chip for a neuron while preserving all relevant function. The activity of neurons depends on intricate biological details. Signal speed depends on axon and dendrite lengths, and small differences of timing can have big downstream consequences. Cell membranes have tens of thousands of ion channels that are sensitive in different ways to different chemicals. Nitric oxide serves as a diffuse signal, passing freely through cell membranes to interact with intracellular structures. Blood flow matters, both in total amount and in the specific chemicals transported. Glial cells, which provide support structures, also influence neuronal behavior. Many cell changes accumulate over time without causing immediate spiking activity. And so on. The silicon chip would need to replicate not just activity at the membrane but many consequences of many changes in interior structure. To replicate all of this so precisely that the functional input-output profile matches that of a real neuron probably requires another biological neuron. Thus, a presupposition of the neural replacement argument fails: we probably cannot create silicon substitutes for biological neurons that preserve all of the functionality relevant to behavior.

If the neural replacement argument succeeded, it would establish the substrate flexibility of consciousness. A substantial chunk of the existing literature evaluates the feasibility of this argument. Despite our commitment to consciousness being substrate flexible, we think the skeptics have it right. This is not our argument.

Humanlike consciousness might require humanlike functional architecture. We don't commit to this, but we grant that it's plausible. If Cao and others are right, the property of *having humanlike consciousness* might have low substrate flexibility: It might depend on highly specific biological functions in us that can only be implemented by neurons very much like our own, down to a fine degree of detail.[5]

However, the question at hand is not whether humanlike consciousness requires a humanlike substrate. It is whether *consciousness of any sort* depends on having a specific type of substrate, or whether instead it has a moderate-to-high degree of substrate flexibility. The failure of the neural replacement argument simply does not speak to that question.

The underlying issue—the reason the literature has focused on specifically human consciousness—is because if the neural replacement argument succeeds, it delivers substantial substrate flexibility all in one leap; and it has seemed to many that it succeeds. Yet the resulting focus on fine-grained functional equivalents to humans necessarily limits the scope of conscious beings examined, and it invites the mistaken conclusion that consciousness requires something extremely similar to our biological configuration Indeed, Block (2025) explicitly argues that similarity to humans at the right grain of functional or biological description is the only sufficient ground we have for justifying attribution of consciousness to nonhuman entities. We disagree, and have developed our argument from the Copernican Principle of Consciousness precisely to overcome this limitation.[6]

6.2. Substantial Substrate Flexibility

---

[5] Godfrey-Smith (2016) and Seth (forthcoming) also suggest that consciousness might require, respectively, metabolism and autopoiesis. However, neither gives clear positive grounds for thinking of such biological properties as a strict requirement on consciousness, and in any case, such abstractly specified categories are plausibly highly substrate flexible (Schwitzgebel forthcoming).

[6] Figdor (MS) argues against substrate flexibility on the grounds that if mental states are substrate flexible, then it would falsely imply that phylogeny is irrelevant to mental state individuation. We don't disagree with Figdor any more that we do with Cao; we just take a broader target. The individuation of specific human or animal mental states may depend on phylogeny in the way Figdor suggests, but not consciousness or its absence per se.

We have made our case for the substrate flexibility of consciousness in terms of biochemical kinds, which are a subset of chemical kinds. One might argue that since the extant discussion is based in terms of carbon and silicon, which are elements rather than compound molecules, we are talking about substrate flexibility at the wrong level or degree of specification. We disagree in two ways. First, we disagree exegetically: in biochemistry, 'carbon-based' life just means life made of molecules using carbon as their scaffolding atom (Petkowski et al. 2020): it is not as if we are made of carbon in the way that diamonds are. The question has always been about chemical kinds.

More importantly, however, we disagree in substance (pun intended). The substrate flexibility of consciousness is clearly a matter of degree. Chimpanzee consciousness takes place in a somewhat different substrate; a chimpanzee brain works just fine for hosting consciousness. But not anything will suffice: hunks of granite are (we assume) not conscious and cannot be made so. The advantage of the phrase "substrate flexibility" over the sometimes used "substrate dependence" and "substrate independence" is that the degreed nature of the phenomenon is evident from the terminology. Although some people might be interested specifically in question of whether consciousness could ever be realized in silicon chips, this is a very specific application of a more general question. We will address this application in the next section. But proper understanding of the issue requires a broader framing.

With this broader framing in view, we think everyone should agree that consciousness is substrate flexible to some degree. The question is only how flexible and in what respects. Astrobiology gives us a scientific toehold on some of the possible variation. And by appealing to Copernican mediocrity among behaviorally sophisticated life forms, we provide grounds for endorsing substantial substrate flexibility while avoiding commitments to computational functionalism and other such hotly contested issues.

Flexibility in chemical kinds – different heteroatoms, different amino or nucleic acids – suggests flexibility at larger scales. For the same types of reasons that Cao, Block, and Godfrey-Smith emphasize, we should not expect that systems so different in their microproperties would be identical in their macroproperties and functional properties. This is speculative, of course; we're aware of no studies of how brains would be different if silicon were a primary heteroatom or if the molecule of genetic information employed different building blocks. But it's a reasonable conjecture that there would be important differences. Even here on Earth, different animal phyla have somewhat different neurochemistries and neural structures. Mollusks and vertebrates employ different (but overlapping and related) ranges of neuropeptides, and basic morphology and function can differ (Chase 2002; Nixon and Young 2003; Oliveira, Calcino, and Wanniger 2019). For example, the differentiation between axon and dendrite is often much less pronounced in mollusks, which often employ bidirectional "neurites" serving both the input (axon) and output (dendrite) roles or other specialized morphologies unfamiliar from vertebrate cases (Budelman 1995; Chase 2002; Nixon and Young 2003). Given this variation among (distantly) related animals even on Earth, it would be amazing if the control system cells (or cell-equivalents) were uniformly similar in all conscious life across the universe, in very different planetary (or non-planetary?) environments, with very different temperature ranges, solvents, pressure levels, gravitational levels, radiation types and levels, energy sources, ratios of available chemicals, and contingent chances of evolutionary history. Astrobiological evidence suggesting the possibility of different chemical kinds, combined with these more general plausibility considerations, combined with the likelihood of many independent evolutionary sites throughout spacetime, suggests that even if one accepts that all naturally evolved conscious life must be carbon based, it will be heterogenous from the chemical level on up. In any case, difference in chemical substance is sufficient for ‘substantial’ substrate flexibility, in our sense—pun (again) intended.

At this point, Copernican reasoning applies. The diversity of life forms throughout the universe will likely employ different chemical structures and different cellular architectures, arranged into different gross morphologies in the brain, brain-analogue, or more distributed signaling system (think again, even on Earth, of octopuses and jellyfish). There will likely be different, complex and intermeshing functional relationships from small-scale chemical bonding up to large-scale functional differences in sensory, memory, and affective systems. To think that somehow, among this diversity, only entities with our particular architecture and functionality would be conscious, would be unmotivated terrocentrism.

7. Implications for Artificial Intelligence

Substrate flexibility is typically brought up in the context of AI consciousness, in particular, the possibility of entities made of the same basic material as our current best generative AI technology. While we have argued that consciousness is substrate flexible, we have not argued that it is capable of being realized in *that* substrate, which does use silicon as (the closest thing a computer chip has to) a scaffolding atom. Should we then assume that consciousness cannot exist in such a substrate? The two authors differ on this question.

One answer (preferred by JMP) is: until we have reason to believe otherwise, we should assume that our current computer chips cannot realize consciousness. Not being able to realize consciousness is the default property we attribute to substrates until we have a reason to think otherwise, and our argument does not give us a reason to think otherwise for the substrates of current AI’s. Note that this position does not imply a ‘biological naturalism’ as suggested by Seth, wherein “consciousness depends on intrinsic properties of its material biological basis”

(Seth forthcoming). It does not assume the necessity of any particular biological process, such as autopoiesis (Maturana and Varela 1972/1980; Seth forthcoming) or low-level biochemical metabolism (Godfrey-Smith 2016). It is instead a simpler default to assuming that specialness needs to be established.

The other answer (preferred by ES) is: we should be open to the possibility of AI consciousness. Once we acknowledge that consciousness does not require our particular substrate, it seems unmotivated to draw the line in any one specific place, as long as the substrate in question shows the capacity to support sufficient behavioral sophistication. To reject AI consciousness specifically on grounds of substrate would require some argument either (1) that silicon chips cannot support the degree of behavioral sophistication that would justify attributing consciousness to sophisticated but differently constructed aliens, or (2) that, even if it can support the requisite sophistication, silicon is for some other reason disprivileged relative to carbon. The failure of the neural replacement argument gives at best very weak evidence for such a general incapacity or disprivilege.

8. Conclusion: Relaxing Behavioral Sophistication

We have argued for the substrate flexibility of consciousness by supporting the three premises we set out in the introduction. In doing so, we used specifically human-level behavioral sophistication as our criterion, erring on the side of conservatism. Relaxing the required level of behavioral sophistication would presumably increase the likelihood of conscious species existing in different substrates.

Suppose your best guess estimate is that, on Earth, consciousness is present in all vertebrates, plus cephalopods and some insects (Birch 2024). And suppose that your best guess estimate is that on average each galaxy contains a million planets where species of approximately that level of behavioral sophistication eventually evolve (even if technological civilizations rarely arise). The observable universe would then host, over its lifetime, a quintillion (10^{18}) qualifying planets. With that many draws from the lottery, some of these life forms will be strange indeed. We do not think you will then also want to suppose that consciousness will be limited to only those life forms lucky enough to be made of the same substances that compose us here on Earth.

References

Adams, Fred (1979). "Properties, functionalism, and the identity theory." *Eidos* 1:153-179.

Antony, Louise, and Joseph Levine (1997). "Reduction with Autonomy." *Philosophical Perspectives* 11:83-105.

Aronld, Frances H. (2017). "Directed Evolution: Bringing new chemistry to life." *Angewante Chemie International Edition* 57:4143-4148.

Bains, William (2004). "Many Chemistries Could Be Used to Build Living Systems." *Astrobiology* 4:137-167.

Bains, William, Janusz J. Petkowski, and Sara Seager (2024a). "Venus' atmospheric chemistry and cloud characteristics are compatible with venusian life." *Astrobiology* 24:371-385.

Bains, William, Janusz J. Petkowski, and Sara Seager (2024b). "Alternative Solvents for Life: Framework for evaluation, current status, and future research." *Astrobiology* 24:1231-1256.

Bains, William, and Sara Seager (2012). "A combinatorial approach to biochemical space: description and application to the redox distribution of metabolism." *Astrobiology* 12: 271-281.

Ballesteros, Fernando Jesus, Alberto Fernandez Soto, and Vincente J. Martinez (2019). "Diving into exoplanets: Are water seas the most common?" *Astrobiology* 19:642-654.

Barrow, John D., and Frank J. Tipler (1986). *The Anthropic Cosmological Principle.* Oxford University Press.

Bechtel, William G. (2012). "Representing Time of Day in Circadian Clocks." In (Eds.) Albert Newen, Andreas Bartels, and Eva-Maria Young, *Knowledge and representation* (pp. 129-162). CSLI Publications.

Benner, Steven A., Alonso Ricardo, and Matthew A. Carrigan (2004). "Is there a common chemical model for life in the universe?" *Current Opinion in Chemical Biology* 8:672-689.

Berg, Jeremy M., John L. Tymoczko, Gregory J. Gatto, Jr., and Lubert Stryer. (2018). *Stryer's Biochemistry, 8th Edition.* W.H. Freeman & Co.

Birch, Jonathan (2024). *The Edge of Sentience.* Oxford University Press.

Block, Ned (2023). *The Border Between Seeing and Thinking.* Oxford University Press.

Block, Ned (2025). "Can only meat machines be conscious?" *Trends in Cognitive Sciences* 30:298-308.

Brown, Sean M., Václav Voráček, and Stephen Freeland (2023). "What would an alien amino acid alphabet look like and why?." *Astrobiology* 23:536-549.Bondi 1968.

Budelman, B. U. (1995). The cephalopod nervous system. In O. Breidbach, O and W. Kutsch, eds. *The Nervous Systems of Invertebrates. Experientia Supplementum, 72.* Birkhäuser Basel.

Caldwell, R. R., and A. Stebbins (2008). "A test of the Copernican Principle." *Physical Review Letters* 100:191302.

Camarena, David, Valerio Marra, Ziad Sakr, and Chris Clarkson (2022). "The Copernican principle in light of the latest cosmological data." *Monthly Notices of the Royal Astronomical Society* 509:1291-1302.

Cao, Rosa (2022). "Multiple realizability and the spirit of functionalism." *Synthese* 200:506.Chalmers, David J. (1996). *The Conscious Mind.* Oxford University Press.

Chase, Ronald A. (2002). *Behavior and Its Neural Control in Gastropod Molluscs.* Oxford University Press.

Ciaunica, Anna, Evgeniya V. Shmeleva, and Michael Levin (2023) "The brain is not mental! coupling neuronal and immune cellular processing in human organisms." *Frontiers in integrative neuroscience* 17:1057622.Clarkson, Bassett, and Lu 2008

Cranford, Jerry L. (2015). *Astrobiological Neurosystems.* Springer.

Cuda, Tom (1985). "Against neural chauvinism." *Philosophical Studies*, 48:111-127.Egan, Greg (1997). *Diaspora.* Millenium.

Figdor, Carrie (2026). A defense of substrate dependence in the mind sciences. Unpublished manuscript.

Fodor, Jerry A.(1974). "Special sciences (or: The disunity of science as a working hypothesis)." *Synthese* 28:97-115.

Frank, A., and W. T. Sullivan, III (2016). "A new empirical constraint on the prevalence of technological species in the universe." *Astrobiology* 16:359-362.

Galakhova, A.A., S. Hunt, R. Wilbers, D.B. Heyer, et al. (2022). "Evolution of cortical neurons supporting human cognition." *Trends in Cognitive Sciences* 26:909-922.

Gillet, Carl (2003). "The metaphysics of realization, multiple realizability, and the special sciences." *The Journal of Philosophy* 100:591-603.

Godfrey-Smith, Peter (2016). "Mind, matter, and metabolism." *Journal of Philosophy* 113:481-506.Godfrey-Smith, Peter (2024). "Simulation scenarios and philosophy." *Philosophy & Phenomenological Research* 109:1036-1041.

Grefensette, Natalie, Luoth Chou, Stephanie Colon-Santos, et al. (2024). "Chapter 9: Life as We Don't Know It." *Astrobiology* 24, Supplement 1:S186-S201.

Ilardo, Melissa, Markus Meringer, Stephen Freeland, et al. (2015). "Extraordinarily Adaptive Properties of the Genetically Encoded Amino Acids." *Scientific Reports* 5:9414.

Ionnidis, Stavros, Gal Vishne, Meir Hemmo, and Orly Shenker, Eds. (2022). *Levels of Reality in Science and Philosophy.* Springer.

Kan, S.B. Jennifer, Russell D. Lewis, Kai Chen, and Frances H. Arnold (2016). “Directed Evolution of cytochrome c for carbon-silicon bond formation: Bringing silicon to life.” *Science* 354:1048-1051.

Keller, Markus A., Domen Kampjut, Stuart A. Harrison, and Markus Raiser (2017). "Sulfate radicals enable a non-enzymatic Krebs cycle precursor." *Nature ecology & evolution* 1: 0083.

Kershenbaum, Arik (2021). *The Zoologist’s Guide to the Galaxy.* Penguin Press.

Kim, Jaegwon (1992). “Multiple Realization and the Metaphysics of Reduction.” *Philosophy and Phenomenological Research* 52:1-26.

Koshland, Daniel E. (1980). *Bacterial chemotaxis as a model behavioral system.* Raven Press.

Lewis 1980 Mad Pain and Martian Pain

Li, Jingyi, Xi Zhao, and Jing Liu (2024). “Biosimilar liquid metal living matter.” *Matter* 7:2033-2065.

Lipinski, Christopher, and Andrew Hopkins (2004). “Navigating Chemical Space for Biology and Medicine.” *Nature* 432:855-861.

Lyon, Pamela, Fred Keijzer, Detlev Arendt, and Michael Levin (2021). “Reframing cognition: getting down to biological basics.” *Philosophical Transactions of the Royal Society of London B* 376:20190750.

Maturana, Humberto R., and Francisco J. Varela (1972/1980). *Autopoiesis and cognition.* D. Reidel.Morris, Kevin(2018). *Physicalism Deconstructed.* Cambridge University Press.

Munoz, Leonid L., Daria Y. Romanova, and Andrea B. Kohn. (2021). “Neural versus alternative integrative systems: molecular insights into origins of neurotransmitters.” *Philosophical Transactions of the Royal Society of London B*, 376:20190762.

Needham, Paul (2003). "Chemical substances and intensive properties." *Annals of the New York Academy of Sciences* 988.1: 99-113.

Needham, Paul (2005). “Mixtures and Modality.” *Foundations of Chemistry* 7:103-118.

Nixon, Marion, and John Z. Young (2003). *The Brains and Lives of Cephalopods.* Oxford University Press.

De Oliveira, A. L., Calcino, A., and Wanninger, A. (2019). Extensive conservation of the proneuropeptide and peptide prohormone complement in mollusks. *Scientific Reports* 9:4846.

Petkowski, Janusz J., William Bains, and Sara Seager (2020). “On the potential of silicon as a building block for life.” *Life* 10:84.

Piccinini, Gualtiero (2020). *Neurocognitive Mechanisms.* Oxford University Press.

Pober, Jeremy M. (under review). “Hybrid Functionalism: Between Superficial and Deep”

Putnam, Hilary (1967). “The Nature of Mental States.” Reprinted in H. Putnam (1975) *Mind, Language, and Reality*, Cambridge: Cambridge University Press 429-440

Putnam, Hilary (1975). “Philosophy and our mental life.” In H. Putnam (1975) *Mind, Language, and Reality: Philosophical Papers Volume 2.* Cambridge, UK: Cambridge University Press.

Raulin, F., P. Bruston, P. Paillous, and R. Sternberg (1995). “The low temperature organic chemistry of Titan’s geofluid.” *Advances in Space Research* 15:321-333.

Scharf, Caleb (2014). *The Copernicus Complex.* Farrar, Straus and Giroux.

Schwitzgebel, Eric (2022). “An objection to Chalmers’s fading qualia argument”. Blog post at *The Splintered Mind* (Dec 13). URL: https://schwitzsplinters.blogspot.com/2022/12/an-objection-to-chalmerss-fading-qualia.html

Schwitzgebel, Eric (forthcoming). *AI and Consciousness.* Cambridge University Press.

Schwitzgebel, Eric and Jeremy M. Pober (2026). "The Copernican Argument for Alien Consciousness; The Mimicry Argument Against Robot Consciousness." *Philosophers' Imprint*, forthcoming; DOI https://doi.org/10.3998/phimp.7823

SETI (2021). "The Drake Equation". URL: https://www.seti.org/research/seti-101/drake-equation [accessed May 22, 2026].

Seth, Anil K. (2025). "Conscious artificial intelligence and biological naturalism." *Behavioral and Brain Sciences.* Published online 2025:1-42. doi:10.1017/S0140525X25000032

Shapiro, Lawrence A. (2000). "Multiple Realizations." *The Journal of Philosophy* 97:635-654.

Shapiro, Lawrence A. and Thomas Polger (2016). *The multiple realization book.* Oxford University Press.

Shoemaker, Sydney (1981). "Some Varieties of Functionalism." *Philosophical Topics* 12:93-119.

Snyder-Beattie, Andrew E., Anders Sandberg, K. Eric Drexler, and Michael B. Bonsall (2021). "The timing of evolutionary transitions suggests intelligent life is rare." *Astrobiology* 21:265-278.Trestman, Michael, Jonathan Birch, and Colin Allen (1995/2026). "Animal consciousness", *Stanford Encyclopedia of Philosophy* (spring 2026 edition).

Udell, David B., and Eric Schwitzgebel (2021). "Susan Schneider's proposed Tests for AI Consciousness: Promising but Flawed". *Journal of Consciousness Studies* 28(5-6):121-144.

Verkhatsky, Alexei, Amaia M. Arranz, Katarzyna Ciuba, and Aleksandra Pekowska (2022). "Evolution of neuroglia." *Annals of the New York Academy of Sciences* 1518:120-130.

Wright, Jason T., Jacob Haqq-Misra, Adam Frank, Ravi Kopparapu, Manasvi Lingam, and Sofia Z. Sheikh (2022). "The case for technosignatures: Why they may be abundant, long-lived, highly detectable, and unambiguous". *Astrophysical Journal Letters* 927:L30.