*Forthcoming in Persson and Sekrst (ed). *Special Issue of Philosophical Problems in Science, Astrobiology Special Issue*. Please cite only published version*

What is a Biosignature and Why Does it Matter? A Philosophical Analysis

Abstract: The concept of a *biosignature* is widely used in the contemporary search for life beyond Earth. Yet its articulation is contested. An influential recent proposal by a joint team of philosophers and scientists has begun to gather garner support, having been cited in the recent announcement of redox spots on the Cheyava Falls sample. This view is presented, analysed and found to be clearly incorrect. A simple alternative is proposed. It is argued however, that this skirmishing about the nature of biosignatures is of little scientific significance, and the worry is raised that the philosophy of astrobiology quite generally is of at best marginal significance to the science that it studies. An alternative picture of what philosophy could usefully contribute to the search for extraterrestrial life is then proposed.

Keywords: Biosignature; philosophy of astrobiology; search for extraterrestrial life; epistemology.

The search for extraterrestrial life is one of the most exciting frontiers in contemporary science. In 2012 Astronomer Royal Martin Rees declared that we would 'soon know' whether we are alone, citing a time period of 'decades'.[1] Three years later, Ellen Stofan, then NASA's Chief Scientist, predicted that we would find definitive evidence within only 20 to 30 years.[2] In 2024, the British astronaut Tim Peake predicted that we would find evidence of microbial life beyond Earth within just five years.[3] Even these optimists, however, would not think it likely that we will observe life *directly*. It is small, the distances to it are large and the obstacles forbidding. The best we can hope for is to find some *trace or sign* of it; something indirect from which we can infer

[1] https://www.theguardian.com/commentisfree/2012/sep/16/alone-in-universe-life-fluke-earth.

[2] https://www.space.com/29041-alien-life-evidence-by-2025-nasa.html

[3] https://www.dailymail.co.uk/news/article-13394581/tim-peake-alien-life-universe.html

life's presence. The name given to these indirect signs is 'biosignatures'. They are the proxies toward which the present search for extraterrestrial life is oriented. The following quotations from the literature are representative:

> "Astrobiology aims to determine the distribution and diversity of life in the universe. But as the word "biosignature" suggests, what will be detected is not life itself, but an observation that implicates living systems: out-of-equilibrium gasses [not] a writhing octopus."[4]

> "Why a concept of biosignature at all? When searching for life, we are faced with indirect evidence and a multitude of proxies. It is useful to have a way to claim that something - perhaps a substance or a signal - is a sign of biological activity, even if it is a by-product of life and not life itself."[5]

This sounds rather straightforward. On closer inspection however, it is not. Although just about every scholarly article in contemporary astrobiology makes use of the concept of a biosignature, there are almost as many articulations of that concept as there are articles. They are often mutually inconsistent. The phenomenon is quite striking. Indeed in *a single paragraph* in one of the most influential articles in the field - a highly influential and widely-cited NASA Technical Report about biosignature detection, published in the leading journal *Astrobiology* - the authors affirm *all three* of the following:

(A) A biosignature is an object, substance, and/or pattern whose origin specifically requires a biological agent.
(B) A biosignature is a substance or pattern that is known to be an indicator of biological activity in a given planetary context.

[4] Harrison Smith and Cole Mathis. Life detection in a universe of false positives: Can the Fatal Flaws of Exoplanet Biosignatures be Overcome Absent a Theory of Life? *Bioessays* 45(12): e2300050 (2023).

[5] Christophe Malaterre. Is There Such a Thing as a Biosignature? *Astrobiology* 23(11):1213-1227 (2023).

(C) A biosignature is a molecule, pattern or other signal that has a non-zero probability of occurring, conditioned on the presence of a living process.[6]

Yet these are manifestly different views. Something meets condition (A) but not (B) if its origin requires a biological agent but this is not known; (C) but not (B) if it has a non-zero probability of occurring conditional on life but is not known to be an indicator of life; and (C) but not (A) if it has a non-zero probability of occurring conditional on life but its origin does not require a biological agent. All of these conjunctions are real possibilities.

So how *should* astrobiologists understand the concept? Is there an interesting problem here or not? The plan is as follows. In §I a recent proposal owing to a team of philosophers and scientists is presented. In §II a problem for this view is presented; it is internally inconsistent. In §III this inconsistency is used to identify a positive criterion or desideratum for an account: it must be *objective* rather than *epistemic*. In §IV a second problem for the view is presented: it entails that biosignatures aren't signs of life. In §V this problem is used to identify a second positive criterion or desideratum for an account: it must be *factive* rather than *non-factive*. In §VI a simple account is presented and defended that meets these two criteria or desiderata. In §VII it is argued that the arguments of this paper are nonetheless of limited significance and the worry that this is a more general feature of the philosophy of astrobiology is developed. In §VIII an alternative, positive picture of what philosophy might contribute to astrobiology is tentatively proposed.

I. A Recent View

As a useful departure point, I will consider the recently-proposed view of a joint team of scientists and philosophers.[7] This proposal, from Gillen et al, is intended to be both philosophically rigorous and scientifically useful. It is:

[6] Sara Walker et al. Exoplanet Biosignatures: Future Directions. *Astrobiology* 18(6): 779-824, p. 781 (2018). This case is picked up on and discussed in detail in Christopher Cowie. New Work on Biosignatures. *Mind* 133(530): 452-471 (2024).

[7] Catherine Gillen et al. The Call for a New Definition of 'Biosignature' *Astrobiology* 23(11) (2023). The only other explicitly philosophical engagement is Cowie (2024).

(D) A *biosignature* is any phenomenon for which (i) biological processes are a known possible explanation and (ii) potential abiotic causes have been reasonably explored and ruled out.

This is supplemented by an account of 'potential biosignature'. This widely-used expression denotes, roughly, things that could or might be biosignatures. So, for example, the redox spots on the Cheyava Falls sample discovered on Mars last year were described by NASA as 'potential biosignatures'; the idea being that they could be, or might be, biosignatures.[8] As Gillen et al present it:

(E) A *potential biosignature* is any phenomenon for which (i) biological processes are a known possible explanation but (ii) potential abiotic causes have *not yet* been reasonably explored and ruled out.

What is the appeal of (D) and (E)? Clause (i) of (D) tells us that for something to be a biosignature, it must be known that life *could* have caused it. That's a necessary condition. But it's not sufficient. Clause (ii) tells us that we *also* need to rule out some candidate abiotic causes. And it's not good enough that we have ruled out abiotic causes that we are presently aware of. For there may be abiotic causes that we are not immediately aware of. We also need to have *reasonably explored* the space of abiotic causes. This also makes sense of (E). (E) tells us that something is merely a *potential* biosignature if it fails to meet clause (ii) of (E); that is, if we haven't yet managed to explore and rule out those possibilities that we would need to.

One of the features of this account that its authors note is its *usefulness*. It is, they claim, a view that practicing scientists can use when making assessments. They might, for example, ask: *have we* explored and ruled out the alternatives? If not, then they merely have a potential biosignature. This is really important. The concept of a biosignature is a recent invention, intended to serve a purpose within astrobiological inquiry. The aim should be to come up with something useful; something that *does a job* for us. That's exactly what (D) and (E) claim to do.

II. The First Problem for Gillen et al

[8] https://www.nasa.gov/news-release/nasa-says-mars-rover-discovered-potential-biosignature-last-year/

Gillen et al's view has gained traction. In a recent paper reporting arguably the most important lead yet in the search for life beyond Earth, *their* account was cited.[9] Nonetheless, it should not be adopted. Seeing why will lead us to a better view.

Begin by noting that if (E) is true, then if something is a potential biosignature, it is something whose potential abiotic causes have *not* yet been reasonably explored and ruled out. Note also that if (D) is true, then if something is a biosignature, then it is something whose potential abiotic causes *have* been reasonably explored and ruled out. This means that

(F) If both (D) and (E) are true, then necessarily, if something *is* a potential biosignature it *is not* a biosignature.

This is a simple consequence of their view. But it is a big problem for Gillen et al. It is a problem because, surely, the consequent of (F) is false. Much as a potential sighting of mountain lion in one's back garden is a sighting that might be of a mountain lion (and so a cause for concern!), so a potential biosignature is something that could be a biosignature (and so a cause of interest!). So the following is true:

(G) It could be that some potential biosignatures are in fact biosignatures.

Yet from (F) and (G) jointly it follows that:

(H) Either (D) or (E) is false.

This is a deeply problematic result for Gillen et al. It seems that their accounts of 'potential biosignature' and 'biosignature' are effectively inconsistent with one another. The only way out is to deny (G), but this is an obviously bad option. If a potential biosignature isn't something that could be a biosignature, what is it?! Something has gone badly wrong for Gillen et al. How did they get into this bind?

III. An Alternative Proposal

The root of the problem is that their account of what a biosignature is, (D), is what we might call an *epistemic* account. By this I mean that if (D) is true, then what it is for something to be a

[9] Joel Hurowitz et al. Redox Driven Mineral and Organic Associations in Jezero Crater, Mars. *Nature* 645: 332-340 (2025).

biosignature is for that thing to stand in certain epistemic relations to us; namely, that it is a thing we *know* to be a possible cause of life, and for which we have *reasonably explored and ruled out* abiotic causes. This all but ensures the inconsistency. For unless potential biosignatures and biosignatures respectively are just *the very same thing*, then the epistemic relation referenced in the account of 'potential biosignature' must be *different* from that in the account of 'biosignature'. And that's what we find in (D) and (E); a potential biosignature is characterised as something for which we have *not* reasonably explored and ruled out abiotic causes. So it is bound to follow that if something is a potential biosignature it is not a biosignature.

Consider the following easy way around this: take the *epistemic* properties *out* of one's account of what it is to be a biosignature. That way we don't preclude the possibility that a potential biosignature is something that might be (where this is an epistemic modal) a biosignature. I'll call such accounts (of 'biosignature') *objective*. There are many objective accounts that one could offer. According to a simple version:

(I) A biosignature is something that is caused by life.
(J) A potential biosignature is something that could be or might be a biosignature.

Unlike (D) and (E), this simple view allows for consistency between biosignatures and potential biosignatures. It also makes sense of why we are interested in biosignatures. We are interested in them because they are things that are caused by life and we want to find life; so if we find a biosignature, then what we want is there to be found too.

One *objection* concerns usefulness. I mentioned above that one of the strengths of Gillen et al's account, as they see it, is its usefulness: that it tells scientists *what to do* in the search for life beyond Earth. Gillen et al would, I think, contrast this with simple accounts like (I) and (J), which do not.

This is a genuine distinction, but it is not a plausible *objection*. It is not generally true that the concept of an X biosignature is useful only if the correct explication of it tells one how to go about searching for an X, and it is not true for biosignatures either. The concept of a biosignature is useful, and indeed is used widely, because it denotes the proxies for life for which we are searching. Having something that plays this role is pretty-well essential for astrobiology and the search for extraterrestrial life. It is essential because they are what we are looking for, and we need something to denote that. The simple objective articulation of the concept speaks to this straight-forwardly. Of course there are other things that it *doesn't* do; like telling us how to go

about conducting scientific inquiries. But that's hardly an objection (especially when the cost of *doing* those things is inconsistency).

IV. The Second Problem for Gillen et al

The epistemic nature of Gillen et al's view is one problem. There is another. Begin with the following premise, which is simply definitional of Gillen et al's account of 'biosignature:

(K) If (D) is true, then a biosignature is any phenomenon for which (i) biological processes are a known possible explanation and (ii) potential abiotic causes have been reasonably explored and ruled out.

Now note that it is possible to satisfy conditions (i) and (ii) though the 'phenomenon' has not been caused by life nor is life present at the site in question. It might not have been. Its source *could* be entirely abiotic. This is simply because there is no guarantee that having reasonably explored the potential abiotic causes with respect to something will have ruled all of them out. Reasonable exploration is not the same as exhaustive exploration. This is explicitly conceded by the authors. Indeed they take it to be a virtue of their account. For example, they write:

> "[We] have acknowledged that discovering a biosignature is not the same as discovering life itself; biosignatures are limited, observable proxies for life and are inherently uncertain. Definitions of biosignature that overlook this gap (e.g., Des Marais et al., 2003) are impracticable."

The Des Marais account to which they are referring is one on which something's being a biosignature *does* entail the presence of life. By contrasting with this, Gillen et al's are stressing that their view is *not* of this kind. So the following premise – all sides concede - is true

(L) It is possible that there is some phenomenon, P, at some site, S, for which biological processes are a known possible explanation, and potential abiotic causes have been reasonably explored and ruled out, yet life does not exist, and never has existed, at S.

Now put (K) and (L) together. By modus ponens they entail:

(M) If (D) is true, then it is possible that P, at S, is a biosignature, though life does not exist, and never has existed, at S.

Now add the following platitude:

(N) For any X: If X does not exist and never has existed, then Y is not a sign of X.

This is, I take it, clearly true. For while something might *seem to be* a sign of something else that does not exist, it cannot, if that thing does not exist and never has done, *actually be* a sign of it. Together (M) and (N) entail:

(O) If (D) is true, then it is possible that there is a biosignature that is not a sign of life.

Add the following additional platitude:

(P) Necessarily, if something is a biosignature, then it is a sign of life.

The result, which follows from (O) and (P), is:

(Q) (D) is false.

And (D), recall, is just Gillen et al's account of what it is to be a biosignature.

V. Gillen et al's Defence

How did Gillen et al get into this predicament? At one level, the answer is straightforward; as is the lesson that we should take from it. Let's say that an account of what it is to be a biosignature is *factive* iff it entails that if some thing, P, at some site, S, is a biosignature, then life exists or has existed at some point, at S. And let's say that it is *non-factive* otherwise. The root of the problem is – pretty clearly - that Gillen et al's view, (D), is non-factive. So the above argument shows that whatever the correct account of 'biosignature' is, it must be factive too.

Ultimately, I think this is correct. I will use it to help develop my own posit e account below. But before doing so it is worth pausing to think about *why* Gillen et al presented a non-factive account and whether there is a compelling defence that they might give. Note that it is not a simple oversight on their part. As we have seen in the above quotation, they present the non-factivity of their view as a *virtue* of it, a virtue that they contrast with the view of Des Marais. What (as they see it) is its virtue? In that quotation they make two distinct claims. They are:

(R) Discovering a biosignature is not the same as discovering life itself; biosignatures are limited, observable proxies for life and are inherently uncertain.
(S) Definitions of biosignature that overlook this gap are impracticable.

Some features of claim (R) are puzzling. *Of course* discovering a biosignature is not the same as discovering life itself. And indeed, as they note, the reason for this is that biosignatures are observable proxies. But neither of these two points implies that biosignatures cannot be things that are caused by life's presence or otherwise entail it. Koplik spots are not the same as measles. They are observable proxies for it. But it would be quite bizarre to conclude from this that Koplik spots aren't things that are caused by measles. Indeed that's precisely what they are.

At the end of (R), however, is the key claim: biosignatures are inherently *uncertain*. I take it that what is meant is, roughly, that a biosignature is the kind of thing that one cannot be certain is caused by life. And this also explains (S). For the thought is presumably that an account of biosignatures that would require that one *is* certain (that what one has discovered is caused by life) is impractical. It is impractical because we won't ever be able to correctly deploy it. We won't be able to correctly deploy it because we will never be certain that what we have discovered is caused by life. Here, then, is the key claim:

(T) One cannot correctly deploy a *factive* concept of 'biosignature' unless one is certain that one's findings are caused by life.

This is false for at least the following two reasons.

Firstly, one can correctly deploy the concept – in the sense of applying it in a true proposition (e.g. 'the X on Planet Z is a biosignature') – without being certain that one's findings are caused by life. All that is required is that the thing one is applying it to is in fact caused by life. If so, the proposition is true. This seems obvious. Suppose I deploy the concept 'star' while pointing at a bright, stationary object in the night sky. Suppose, furthermore, that the thing I'm pointing at not only looks just like a star, but is a star. I have, presumably, correctly deployed the concept 'star'. This is the case even if I can't be *certain* that it is a star, because there is an outlying possibility that it is a UFO.

The obviousness of this point makes Gillen et al's position puzzling. It is as though they are working with a criterion for correct concept application whereby one ought to apply some concept, C, only if one is certain that the subject falls under it. Call this *the certainty criterion*. This is a very strong and very unusual criterion, not generally applicable elsewhere. Perhaps however Gillen et al think the certainty criterion applies to the concept of a *biosignature* only. This would be very strange. We would need to see some argument for it. And none is present in the text. Suppose, however, that some such argument *were* provided. *Even this* would not be good for Gillen et al. For it would cut equally against *their own view* just as much as it would against factive

views. To see this, consider Gillen et al's view, (D), of what a biosignature is. On this view, if something is a biosignature, then its potential abiotic causes have been reasonably explored and ruled out. This too would be vulnerable to the certainty criterion. For presumably it will rarely if ever be certain that we have reasonably explored and ruled out potential abiotic causes. One could, after all, always doubt whether one's explorations had been reasonable, or whether abiotic caused explored had indeed been ruled out. So if factive views are impractical, then so is (D).

Secondly, one can correctly deploy the concept – of a biosignature, factively understood - in an auxiliary sense to frame and structure one's inquires even if one is not certain of any specific objects that fall under it. Indeed this is precisely how it is ordinarily used. When (for example) a NASA white paper states that the James Webb Space telescope will be used to search for biosignatures, this is what has happened. The concept of a biosignature has been unproblematically deployed – in a way that structures actions and future inquiries – even though we don't know of anything at present that falls under it. This way of using concepts isn't peculiar to biosignatures. An historian may spend their career trying to work out whether King Arthur really existed without knowing whether he ever did. A physicist may look for a grand unified theory of gravitation and the standard model without knowing whether there is any such thing. In both cases the search may be in vain. Certainty may be impossible to achieve precisely because there is no such figure or theory to be found. But that doesn't mean the concept hasn't been correctly applied in the search. It has been. And it has been extremely useful to this end. Suppose that the historian of physicists had been *debarred* by Gillen et al from referring to King Arthur or the Grand Unified Theory. This would hardly have made their task easier. Indeed if these concepts hadn't existed, they would have invented them precisely because it helped to articulate what they were doing.

VI. My Account

Gillen et al's view, (D), has two extremely implausible and unsupported implications. Firstly, it entails that potential biosignatures aren't things that might be biosignatures; indeed they couldn't be. Secondly, it entails that biosignatures aren't signs of life. What should we take from this? The lesson from the first implausible implication is that the account of biosignatures should be objective, not epistemic. The lesson from the second implausible implication is that the account

should be factive, not non-factive. I proposed above an account that meets the first of these two criteria. According to that account, recall:

(I) A biosignature is something that is caused by life.

(J) A potential biosignature is something that could be or might be a biosignature.

Note that this account also satisfies the second criterion. For if something is caused by life then life exists. So the view is factive. Taken together, (I) and (J) are, I submit, the simplest easiest view of biosignatures that avoids the twin pitfalls of Gillen et al's view. But mightn't it be vulnerable to objections. I'll consider two: from *uselessness* and *universality* respectively.

The first concerns *uselessness*. Gillen et al think it a strength of their view that it tells scientists what to do in the search for life. This is contrasted (by them) with factive accounts, which don't. Are they right about to take this as a mark against such views? In short, no. While it is true that factive accounts do not tell one how to go about finding life, it would be a mistake to think that they are not thereby useful. Indeed the role that these accounts play in framing and structuring inquires is not merely useful (as in the case of King Arthur or the search for a grand unified theory of physics), it is indispensable. A simple, factive conception like (I) and (J) performs two important tasks to this end. The first is simply to mark *indicators* of life as opposed to *life itself*. This is useful because, as noted earlier, we cannot search *directly* for life. If we are to find it, it must be by finding an intermediary. This might seem trivial. But given our present epistemic predicament we *need* something to mark this. To see this, suppose for example, that we had access to a machine that could scan for life, or a clairvoyant who could reliably intuit its existence. There would be no need to focus our attention on *indicators* of life. We could, and I think would, just focus on the life itself. If this were the case, it is not clear that we would need the concept of a biosignature at all. But we are not in this position. And so we need a concept to play this role. The second function, which the concept simultaneously plays, is to *group* those indicators of life given that we do not presently know what they are, and that for all we know they may be extremely varied and perhaps even surprising. To see why this is important, suppose that we could not detect life directly, but that we knew it existed in all and only those places that some *other* thing that we *could* detect more easily existed; for example, water. There would not, now, be any real need for the concept of a biosignature. We could just think and talk directly about water. That, I think, is exactly what we would do. So the concept of a biosignature, as I understand it, jointly fills two roles that make it an indispensable tool for structuring and conceiving of inquiry. Firstly, it allows us to talk about indicators of life. Secondly, it groups them given uncertainty about their nature. Without *both* of these there would be no need for the

concept. But as it stands, both roles do need to be filled, and filled together. Having something that does this is pretty-well essential for astrobiology and the search for extraterrestrial life. The simple factive articulation of the concept of a biosignature speaks to this by being the minimal articulation that jointly plays these two roles. That's why it is foundational for astrobiology.

Consider a second worry: *from universality*. My view is that something is a biosignature only if it is caused by life. But do I mean caused by life *in all instances*? Or *only in some*? It sounds like I am committed to the former: the insistence on factivity suggests as much. But this means that something is a biosignature only if it is (what astrobiologists call) *a universal biosignature*. And, even if the universe is teeming with life, there may be no such things. Perhaps, for example, there are many methane-oxygen disequilibria in a planetary atmospheres, and perhaps furthermore they are very reliably caused by life; 999 times out of 1000. My view would nonetheless entail that a methane-oxygen disequilibrium is *not* a biosignature. This seems peculiar: to set up the concept of a proxy for life in a way that fails to capture just about every instance that we would want it to.

Responding to this worry allows me to articulate a nice feature of my view. All of the existing literature on biosignatures has missed an important distinction: between types and tokens. A car is a type of thing. My car is a token of that type. Countries are types of thing. Germany is a token of that type. Birds are a type of thing. The bird in my garden right now is a token of that type. And so on. We can – indeed should - similarly distinguish between tokens and types in the analysis of biosignatures. A methane-oxygen disequilibrium is a type of thing. The specific methane oxygen disequilibrium on planet Zog is a token of that type. It is a curiosity that no existing account of biosignatures in the litertaure has drawn this distinction or specified which it is attempting to capture. Consider for example Gillen et al. Is their view meant to apply to token things (e.g. the methane oxygen disequilibrium on planet Zog) or to types (e.g. methane oxygen disequilibria)?

My view is this. (I) and (J) apply to both tokens and types. Understanding what this means answers the above objection (from universality). Consider the following two sentences:

(U) My grandfather's smoking caused his cancer.

(V) Smoking causes cancer.

Both of these sentences are true. (U) concerns tokens: my grandfather's smoking and his cancer. (V) concerns types: smoking and cancer. (U) is true only if my grandfather smoked and got cancer. But the truth of (V) does not require that everyone who ever smokes gets cancer. If it did, it would be false. Generally, token causation of the form 'X caused Y' requires that X is

instantiated and that Y is too. But type causation of the form 'X causes Y' does not require that every time that X is instantiated, Y is. Now think about biosignatures. If we are talking about token things, for example the methane-oxygen disequilibrium on Planet Zog, then my account, (I), entails that it is a biosignature only if life exists, or did exist, on Planet Zog. But if we're talking about types of things, for example methane oxygen disequilibria, then my account does not entail that it is a biosignature only if life exists at all tokens of methane oxygen disequilibria. So it does not entail that something is a biosignature only if it is a universal biosignature. So it is not vulnerable to the objection from universality. Indeed this objection brings out a neat feature of the view. By making a biosignature a matter of causation, it allows us to say, as surely we want to, that (e.g.) if the redox spots on Cheyava Falls weren't caused by life then they aren't a biosignature, they merely seem to be, and that if the DMS in the atmosphere of K2-18b isn't caused by life on that planet, then it isn't a biosignature, whilst avoiding the slide into the view that some type of thing is a biosignature only if universally so.

VII. Is There Any Point in a Philosophy of Astrobiology?

The concept of a biosignature is widely used in contemporary astrobiology. Yet when authors attempt to 'define' it, things get confusing. Gillen et al's view is a fascinating case in point. In attempting to provide clarity they end up with a view of biosignatures that is equivalent, roughly, to a view of what it is to be a bird-spotter that entails that someone is a bird-spotter only if they haven't spotted a bird and couldn't do so. I recommend (I) and (J) as a simple view that remedies this. It is both objective and factive and captures what astrobiologists use the concept for in their ordinary practices, without complication or inconsistency. It can be profitably supplemented by the token/type distinction.

One might wonder how Gillen et al ended up in the place that they did. When I initially read the account, I thought that the answer was obvious: they had confused what it is to be a biosignature with one would need to do to know that one had found one. Their account was of the latter, but presented as being of the former. But on closer examination this could not be the case. For their account is non-factive. One could meet its criteria and yet life not be present. It cannot, then, be an account of the conditions for knowledge that one has found life. For knowledge is factive.

A second option is more deeply philosophical. It is that Gillen et al are being driven by a broadly *positivistic* philosophy. They are trying to write empirical, verification conditions for the

application of a concept into the meaning of the concept itself. So the meaning of the concept in effect becomes a specification of the method for its own application. This *could* be the case. But it seems like over-interpretation. Certainly, there is no reference to any such inspiration in the account. I now think that a less specific diagnosis is more accurate. Gillen et al are simply being driven by the attempt to come up with something *useful*: something that can give direction to practicing scientists by writing what one ought to *do* into the nature of the concept itself. This comes through clearly in various passages (as noted earlier). Suppose it is right. It opens up a new way that they might respond to my criticisms. I have argued that their views could not be correct; indeed could not be *true*. Perhaps, however, they are *not interested* in truth or correctness. As such, my criticisms are accurate but nonetheless miss their target.

One test of this is whether the concept is actually used by the scientific community. And here, one might think, they are on solid ground. For their account is in fact referenced in one of the most influential and important papers about the search for extraterrestrial life in recent years; Hurowitz et al's announcement of the discovery of redox spots on Cheyava Falls on Mars in *Nature* (2025). This looks like *exactly* the kind of usefulness that they might want; scientists taking up their account and running with it in what could prove to be one of the most important discoveries ever made. When we look more closely, however, the picture looks rather different. For in fact the authors of this paper cite *two different sources* with respect to how they understand what biosignatures are. One is Gillen et al, the other is Des Marais (2003).[10] And as we have seen, these accounts are not only different, they are quite incompatible, Des Marais' being an objective, factive account that Gillen et al explicitly distance themselves from; indeed that they characterise their view in contrast to. And as it happens, indeed Hurowitz et al go on to cite directly from the Des Marais account, not Gillen et al.

This is interesting. Indeed, it is worrying. The Cheyava Falls paper is surely one of the most influential and important papers in the search for life beyond Earth *ever*. It is entirely about a potential biosignature, and it uses this term repeatedly, as do the NASA press releases about it. Yet it cites, together, two incompatible views of what a biosignature is, one of which (Gillen et al's) is both internally inconsistent and entails that a biosignature isn't a sign of life at all. Suppose that one were to point this out to the authors; to ask them about this peculiar feature of their paper. What would they say? Would it give them pause; reason to go back and revisit their

[10] They also cite a third source, The (2013) Report of the Mars 2020 Science Definition Team. But this source simply cites Des Marais.

arguments and conclusions, or to downwardly revise the significance of their results? Would it cause them to be more or less confident in any way that the finding on the surface of Mars is of biogenic origin? If we are honest, it would not. Whether they refer to Des Marais, Gillen et al, some other account (mine included) would not actually influence what they are doing – the substance of their thought and argument – *in any way at all*. One could take out the reference to biosignatures altogether and the paper would run just as before; nor would the excitement in the scientific community be any the less. The correct philosophical articulation of the concept of a biosignature is, if we are honest, completely irrelevant to everything that makes arguably the most important paper about potential biosignatures ever interesting and important, as well as to the assessment of its conclusions.

All of which makes one wonder whether papers like this one, *the one you're reading right now* – papers which provide an analysis of what it is to be a biosignature - play any useful role in advancing the goals of astrobiology or the search for extraterrestrial life at all; whether they help anyone, in any way, in the search for life beyond Earth. My strong suspicion is that they don't. And so, to put the point perhaps overly dramatically, there is a sense in which what you've been reading to this point is a waste of time. There was really nothing to be gained form getting your head around all of those bracketed letters, (A), (B), (C) and so on in the first place. Of course, this isn't necessarily a problem. A philosophy of X needn't advance the goals of X. It may just provide a commentary on it in a way that is interesting in its own right. But there is at least something unsatisfactory about admitting this in the present context. Surely the analysis of the concept of a biosignature is an exercise engaged in precisely because it is meant to be of use to practitioners. So the fact that it is quite irrelevant to their outcomes whether they cite one internally inconsistent account (Gillen et al), two mutually inconsistent accounts (Gillen and Des Marais), or another altogether (mine, yours or someone else's) is somewhat down-heartening.

More worryingly still, the same issue arises elsewhere in the philosophy of astrobiology. The *other* area – aside from thinking about the nature of biosignatures - that philosophers have had a lot to say about is the nature of life itself. There is a somewhat significant literature on this subject now, much of which concerns how we should understand what life is, for the purposes of astrobiology and the search for extraterrestrial life more generally. This makes sense. It is often claimed that we lack an adequate account of the nature of life itself and that this makes the

search for life especially problematic.[11] The point appears in not only in academic presentations, but also popular. For example, NASA's website section Can We Find Life? begins with:

> "So far, the only life we know of it right here on our planet Earth. But we're looking. The big question – is there life beyond Earth? – comes with an ironic asterisk: we don't really have a universally accepted definition of life itself."

The problem, I take it is, is that without an adequate account of the nature of life, we don't really know what we're looking for. And so of course it will be difficult to find it. This is a version of Plato's paradox of inquiry. If you don't know what you're looking for, you won't be able to recognise it when you see it, and so you won't be able to close out an inquiry that has its discovery as the aim.

Can philosophy help get us out of this bind by providing us with a theory of what life is? Interestingly, by far the most influential and important work on this answers pretty firmly in the negative. In her *The Quest for a Universal Theory of Life: Searching for Life As We Don't Know It*, Carol Cleland firmly rejects the attempt to provide definitions of 'life' or, for that matter, philosophical theories of life. They are, she argues, misguided. This is in large part because we are working from such a limited sample ('N=1'). Rather, she suggests, we should take our cue in thinking about what life is from ongoing empirical investigations. If this is right, and Cleland effectively says as much, philosophical contributions to astrobiology that take the form of papers offering a 'theory of life' (or indeed that offer definitions of 'life') are, much like papers offering theories of biosignatures (or definitions of 'biosignature'), of limited use.[12]

---

[11] For a public-facing expression see e.g. https://science.nasa.gov/exoplanet. For the attempt to provide a definition of 'life' for the purposes of astrobiology see e.g. Steven Benner. Defining Life. *Astrobiology* 10:10 (2010). For the worry that successful discovery is difficult in the absence of a theory of life (and a proposed solution) see e.g. Harrison Smith and Cole Mathis. Life detection in a universe of false positives: Can the Fatal Flaws of Exoplanet Biosignatures be Overcome Absent a Theory of Life? *Bioessays* 45:12 (2023). There is more detailed discussion in Cleland (2019).

[12] I think that Cleland is right about this. I have defended this at greater length elsewhere [REDACTED].

VIII. What a Philosophy of Astrobiology Might Usefully Be

The emerging picture is intentionally pessimistic. Two of the main preoccupations of the philosophy of astrobiology such as it presently is have been to offer definitions of 'biosignature' and of 'life'. Yet it's not clear how these are to contribute to the search for life beyond Earth. The thought that it probably doesn't is one that, I should add, I am not the first to make. The point is nicely expressed by eminent astrobiology Steven Benner. He writes clearly:

> "Many (and perhaps most) of those scientists who are aware of philosophers on their periphery find their approach not particularly useful."[13]

Many would, I suspect, agree with him. Nonetheless, I am in fact optimistic. There *are* ways in which philosophy might make a meaningful contribution to astrobiology. One of the most interesting and important is by providing tools and resources familiar from epistemology for conducting and assessing inquiries of the kind that astrobiologists might find themselves engaged in. Indeed some of the authors whose work we have already encountered offer the best examples of this. One is Cleland herself, who has urged that the search for life beyond Earth be conducted via the search for *anomalies*; a strategy based in philosophical reflection on studies from the history of science. Another, somewhat similar, is Peter Vickers, a co-author on Gillen et al, whose application of the role of 'unconceived alternatives' in the history of science to astrobiology is fruitful in undermining inference to the best explanation as a method for confirming claims to have found extraterrestrial life.[14] These are both examples of bringing the tools of epistemology and philosophy of science to the attention of astrobiologists in a way that might prove fruitful for their studies.

There are many more respects in which this approach can be developed. My own view is that astrobiologists could profit from exposure to the *modal* approach to hypothesis assessment as developed in contemporary analytical epistemology.[15] Modal frameworks of hypothesis

---

[13] Steven Benner, Defining Life. *Astrobiology* 10:10 (2010): 1029.

[14] See especially Peter Vickers Expecting the Unexpected in the Search for Extraterrestrial Life. *International Journal of Astrobiology* 19(6): 482-491 (2020); Peter Vickers et al. Confidence of Life Detection: The Problem of Unconceived Alternatives. *Astrobiology* 23(11): 1202-1212 (2023).

[15] E.g. Timothy Williamson. Knowledge and Its Limits. OUP (2000). Duncan Pritchard. Anti-Luck Epistemology. *Synthese* 158(3): 277-297 (2007).

assessment - in which the basic metric of assessment is possibility rather than probability - offer an alternative from the familiar Bayesian approach. Much existing astrobiological practice is, I think, naturally couched in these terms. I have attempted to bring this to astrobiological practice in recent work.[16] No doubt there are many more examples waiting to be fruitfully developed.

References

- Steven Benner, Defining Life. *Astrobiology* 10:10 (2010): 1021-1030.
- Christopher Cowie. New Work on Biosignatures. *Mind* 133(530): 452-471 (2024).
- Catherine Gillen et al. The Call for a New Definition of 'Biosignature' *Astrobiology* 23(11) (2023).
- Joel Hurowitz et al. Redox Driven Mineral and Organic Associations in Jezero Crater, Mars. *Nature* 645: 332-340 (2025).
- Christophe Malaterre. Is There Such a Thing as a Biosignature? *Astrobiology* 23(11):1213-1227 (2023).
- Duncan Pritchard. Anti-Luck Epistemology. *Synthese* 158(3): 277-297 (2007).
- Harrison Smith and Cole Mathis. Life detection in a universe of false positives: Can the Fatal Flaws of Exoplanet Biosignatures be Overcome Absent a Theory of Life? *Bioessays* 45(12): e2300050 (2023).
- Peter Vickers Expecting the Unexpected in the Search for Extraterrestrial Life. *International Journal of Astrobiology* 19(6): 482-491 (2020);
- Peter Vickers et al. Confidence of Life Detection: The Problem of Unconceived Alternatives. *Astrobiology* 23(11): 1202-1212 (2023).
- Sara Walker et al. Exoplanet Biosignatures: Future Directions. *Astrobiology* 18(6): 779-824, p. 781 (2018).
- Timothy Williamson. Knowledge and Its Limits. OUP (2000).

[16] [REDACTED]