New Work on Biosignatures

Christopher Cowie
University of Durham, UK
christopher.d.cowie@durham.ac.uk

The search for extraterrestrial life centres on the search for 'biosignatures'. Yet there is little agreement within the scientific community with respect to what exactly it is for something to be a biosignature. Existing accounts are presented and criticised. An alternative is provided that resolves problems with existing accounts by distinguishing clearly between types and tokens.

1. Introduction

Astrobiology is the field of scientific enquiry concerned with the nature of life in the universe. Given that there is at present only one data point – earth – the field is at times unavoidably speculative. This is part of what makes it exciting. It also explains why astrobiologists must sometimes venture into philosophical territory (for example, Duner, Persson and Holmberg 2013, Cleland 2019). This article concerns a core conceptual issue at the foundations of this emerging science, and by extension at the foundations of the search for extraterrestrial life itself.

Within astrobiology, the concept of a *biosignature* is frequently deployed.[1] It performs a useful function. We are unlikely to observe extraterrestrial life *directly*. At best we will observe something indirect from which we can draw an inference. These intermediaries are biosignatures. Consider an example. The robotic mission that is presently being conducted on the surface of Mars by the *Perseverance Rover* will study an area that scientists believe was once – perhaps a billion years ago – the bed of a river delta system (Mangold et al. 2021). It is conceivable that it will uncover complex structures that resulted from ancient life. It is searching for biosignatures (for example, Hickman-Lewis et al. 2022, Baque et al. 2022). Consider a second example. The newly commissioned James Webb Space Telescope will study the atmospheres of planets in distant solar systems in search of atmospheric compositions – such as telltale

[1] Scholarly articles in the field that feature 'biosignature' in the title number well into the tens of thousands. For a recent history see Duner (2019). Important recent studies that include accounts of what biosignatures are include Des Marais et al. (1999), (2002), (2008), Walker et al. (2018), Catling et al. (2018), Lingam and Loeb (2021).

Mind, Vol. 133 . 530 . April 2024
https://doi.org/10.1093/mind/fzad050

'chemical disequilibria' – that result from life (for example, Krissansen-Totten 2018, Thompson 2022). It too is searching for biosignatures.

Contrary to what the above passage might suggest, the nature of biosignatures is not as clear as it might be. By this I do not merely mean that it is not clear which things are biosignatures and which are not – though this is also true. Rather, I mean that it is unclear what it even is for something to be a biosignature. This is in part at least a conceptual, *philosophical* issue. I bring the issue into view and make use of some well-worn philosophical techniques to defend a simple, attractive account. Biosignatures are indicators of life: reliable indicators if understood at the level of types, factive indicators at the level of tokens.

2. The current literature

The following passage is taken from the start of Walker et al. (2018): a recent, influential and widely cited NASA Technical Report about biosignature detection, published in the field-leading journal *Astrobiology*. By working slowly through this passage we can helpfully unmask the lack of clarity in the current understanding of what a biosignature is. It reads:

> Des Marais et al. (2002) defined a *biosignature* as an 'object, substance, and/or pattern whose origin specifically requires a biological agent' (Des Marais and Walter 1999; Des Marais et al. 2008). In this paper we follow this convention and refer to a substance or pattern that is known to be an indicator of biological activity (in a given planetary context) as a biosignature, e.g., a 'biosignature molecule' or 'biosignature pattern'. More specifically, we quantify a *biosignature as a molecule, pattern or other signal that has a non-zero probability of occurring, conditioned on the presence of a living process* (see Section 3 where we define *P(data|life)* and provide a quantitative definition for a biosignature within a Bayesian framework). Importantly, a biosignature does not imply life, it only implies a signal consistent with life. To qualify as *evidence* for life, a biosignature should be much more likely to be produced by living processes than by abiotic ones (see… Catling et al. 2018). That is, a molecule, pattern or signal must be able to be produced by life to be a biosignature, but it does not qualify as evidence for life unless life is the best explanation for its production in a given environmental context. (Walker 2018, p. 781, emphasis added)

Start at the top. The authors begin by favourably setting up the Des Marais (2002) account – itself another influential and well-cited

Astrobiology piece, lead authored out of NASA's Ames Research Centre – as a basis on which they build. That account, Des Marais (2002), states that a biosignature *requires* a biological agent as its origin. This is very restrictive. Suppose that something both *in fact* and *uncontroversially* has a biological agent as its origin, but *could* – though does not – have an abiotic origin. On the Des Marais account it is not a biosignature as a biological agent was not *required* as its origin. Consider the following example. Suppose that methane-oxygen disequilibria are almost always – *but not in every case* – caused by life. And suppose we find a planet that enjoys such a disequilibrium and that it is in fact caused by life. The Des Marais account entails that we have not found a biosignature.

> Though this may seem like a simple mistake, we should be careful. The same phrasing reappears in Des Marais' subsequent articles on the topic, (2008) and (2018). It is also *quite explicitly* used elsewhere. In their recent paper 'False Biosignatures on Mars: Anticipating Ambiguity', Julie Cosmidis and Sean McMahon – a philosophically minded astrobiologist – write approvingly:We note that, by definition, a biosignature is more than simply a phenomenon produced by life: it is a phenomenon that specifically requires a biological agent – that is, it could not have been produced naturally by non-living systems (e.g. Des Marais et al. 2008). (McMahon and Cosmidis 2022, p. 2022)

McMahon and Cosmidis are making explicit use of the restrictive feature of the Des Marais account. This is puzzling. Let's set it to one side for now though and return to the lengthy passage from Walker with which we began. Focus on the next part of that passage in which Walker offers a *commentary* on the Des Marais account that has just been presented. Walker says: 'We follow this convention and refer to a substance or pattern that is known to be an indicator of biological activity (in a given planetary context) as a biosignature'. Here, Walker is explicitly stating that they are *following the convention* of the Des Marais account just discussed. This implies a continuity between these two accounts. But this is arguably not borne out by the view of what a biosignature is that Walker subsequently provides. Indeed, Walker's stated view actually strays a very long way from Des Marais. For Des Marais, recall, the key point is that:

> *DM*: The origin specifically requires a biological agent.

For Walker, by contrast, as we see in the quotation immediately above, the key point is that:

> *W1*: A biosignature is known to be an indicator of biological activity.

There are several big differences here. One is that W1 contains epistemic qualification in a sense that DM does not. W1 requires that the indication to which it refers is *known* to be the case, whereas DM does not require that the origination to which it refers is known, or believed, or indeed that it stands in any relation to any observer. Suppose that a fossilised microbe that in fact required a biological agent as origin is found on Mars, but is not known to be a fossilised microbe. DM entails that it is a biosignature, whereas W1 entails that it is not.

Consider also a second difference between the two accounts. W1 entails anything (known to be) an *indicator* of biological activity is a biosignature whereas DM makes use of a much more restrictive relation than indication: *origination*. Origination is more restrictive because, quite generally, one thing, *y*, can indicate another, *x*, without *y* being originated by *x*. Indication lacks the *directionality* of origination. For example, the fact that there are heavy clouds above indicates that it will rain. But the heavy clouds do not originate from rain. It's the other way round. This matters when it comes to biosignatures because there could, in principle, be things that indicate life without being caused by it; they could be either causes of it, or correlates of it. W1 would allow that these kinds of thing could be biosignatures whereas DM would not.

These differences are significant. But interestingly they are arguably *less significant* than those between the Walker account *just discussed*, W1, and the subsequent explication of that account *according to its own authors*; an explication according to which '[a] biosignature [i]s a molecule, pattern or other signal that has a non-zero probability of occurring, conditioned on the presence of a living process' (p.781). Focus on the key part of this account according to which:

> *W2*: A biosignature has a non-zero probability of occurring, conditioned on the presence of a living process.

Strikingly, W2 drops the epistemic qualifier that was present in W1. In this respect W2 is actually more like DM than the account of which it is meant to be an explication, W1. But in another respect W2 is unlike either DM or W1. The relation that W2 posits between things and life is neither *origination* (as in DM) nor *indication* (as in W1) but rather is probabilistic in the specific sense of 'a non-zero probability of occurring, conditioned on the presence of a living process'. This entails that something with a 0.01 probability of occurring conditional on 'the presence of a living process' is a biosignature. This sets the bar very low. How low exactly? Walker says: 'Importantly, a biosignature does not imply life, it only implies a signal consistent with life' (p. 782). This (the above

quotation) correctly explicates W2's preceding claim about probabilities: for the probability of *x* conditional on *y* to be non-zero is indeed simply for *x* to be consistent with *y*. This means that anything – or rather, any molecule, pattern or signal – consistent with the presence of life is in fact a biosignature of that life. The Andromeda galaxy, 2.5 million light years away, contains molecules. This is not inconsistent with the existence of life in that galaxy. So this is a biosignature. This is a very permissive view. Pretty much everything ends up being a biosignature. Perhaps aware of this, Walker notes that we should differentiate what a biosignature is from *evidence of life,* about which they write: '*Evidence* for life should be much more likely to be produced by living processes than by abiotic ones' (p. 782, emphasis in original).

This is peculiar. Why have an account of what a biosignature is that allows that *everything* is a biosignature and a completely different account of what evidence for life is? It would surely be better to say that the account that they provide of what it is to be *evidence of life* is in fact the correct account of what it is to be a *biosignature*. One of the other, most influential and widely cited views – a view that is in fact referred to by Walker in the lengthy passage with which we began – explicitly takes this approach. Catling et al. write: 'A biosignature is any substance, group of substances, or phenomenon that provides evidence of life' (Catling et al. 2018, p. 710).[2]

There is huge variation between Walker (both W1 and W2) and Catling. There are many things that Walker would think of as biosignatures that Catling would not; namely, all of those things that are consistent with life, but not evidence of it. And there is also significant variation between Catling's view and the view of Des Marais that we discussed earlier. To see this compare DM with (the equivalent part of) Catling, C:

DM: …origin specifically requires a biological agent.
C: …provides evidence of life.

Now, presumably something can provide evidence of life without its origin requiring a biological agent. In fact, there are lots of ways in which this could happen. Option one: something could originate in life and so be evidence of life but not *require* life for its origination, so it would be a biosignature by C though not by DM. Option two: something could be evidence of life but not *originate in* life, so it would be a biosignature by C but not DM. This could be the case if, for example, something correlates strongly with life but isn't caused by it. Option three: something could

[2] Compare also Pohorille and Sokolowska (2020).

originate in life but not be evidence of life, and so be a biosignature by DM but not by C. This would be the case on certain 'subjective' readings of 'evidence' that I will return to later; roughly, readings according to which things caused by life are only *evidence* of life if recognisable as such against a background of information on the part of some observer.

What does the foregoing brief examination show? Most obviously, that the most influential existing accounts of what a biosignature is vary hugely. Secondly, that some of these very different accounts are presented as similar to one another when the differences are at least as great as the similarities. And thirdly, that some of the existing accounts have some very peculiar consequences indeed. My aim is to delineate the issues and defend a view. Three questions emerge from the preceding analysis and will be useful as a guide:

1. *The ontological question*: What kinds of things – in the sense of *ontology* – are biosignatures? Are they, for example, things, facts, events, processes, kinds or of some other category?
2. *The substantive question*: What type of *relation* do these things stand in to life? Are they, for example, evidence of life, statistical correlates of life, or effects of life, or do they stand in some other relation to it?
3. *The factivity question*: Must a biosignature *always*, *actually* be a sign of life? Or could something be a biosignature though it is only sometimes, or often, or *apparently* a sign of life?

These questions allow us to group the obvious points of difference between the accounts discussed above. My focus is on providing an analysis of what it is to be a biosignature that answers the second and third questions. This means setting two further issues to one side. Firstly, I shall set the first question to one side. I use the neutral expression 'thing' to refer to all biosignatures. As I stipulatively understand it, 'thing' is neutral across all ontological categories and as such does not bias our inquiry in any particular direction. So, for example, objects, properties, propositions, facts, events and processes are all *things*, albeit different categories of thing. Categories that essentially involve an *experiential* element – such as observations and, on some uses, phenomena – are also things. If someone wishes to take up the ontological question in detail, so be it. It is not, I think, where the interesting action or controversy lies. Secondly, I shall also set aside the question – on which there is a significant existing literature – of how we should understand *life* itself.[3]

[3] See especially Cleland (2019) for a summary of recent work on this issue, and a broadly sceptical approach to attempts to define 'life'.

My focus is simply on the nature of the relation that things must stand in to life if they are to be biosignatures. These two issues – the nature of the biosignature relation and the nature of life respectively – are in principle independent of one another; differing accounts of the nature of life do not speak for or against differing accounts of the nature of the biosignature relation. Once the correct account of 'life' is found, it can simply be substituted in to fully explicate the second relatum of the relation.

3. A question about method

Before we begin to answer these two questions and so to unmask the nature of the concept of a biosignature, it is necessary to say something about *method*. I am engaged in a project of conceptual analysis. Within this, we can distinguish between *descriptive* and *normative* analyses of a target concept. A descriptive analysis delineates the concept as it figures in ordinary use. A normative analysis states which concept we *should* be working with. Which am I attempting to provide? Both.

My view is that there is a core concept that – at least most – astronomers and astrobiologists have in mind when they use the term 'biosignature'. I aim to delineate it. That's a descriptive project. There is also a normative dimension. I think that the core concept I delineate is the one that astronomers and astrobiologists *ought* to be using given the nature of their project. That's because *this* concept makes sense of what practitioners are actually doing. It plays a foundational, explanatory role more effectively than the other candidate delineations of the concept.

The co-incidence of the descriptive and normative should not be a surprise. The very same fact explains both why astronomers and astrobiologists will recognise my delineation of the concept as the *true* delineation of that concept at the descriptive level, and why they will recognise it as the delineation that they *ought* to be working with. It is that there is a need for something to fill this specific conceptual role within the science of astrobiology; and the concept of a biosignature was introduced for this very purpose. The role for philosophy is to reveal this.

4. Proposing the view (1): the factivity question

The key to answering the factivity question is to make an important distinction that has not yet been made in the large and growing literature

on biosignatures; between *types* and *tokens*. Consider the following three cases:

> *Case 1*: We discover a planet, Planet A, with a methane-oxygen disequilibrium in its atmosphere. A methane-oxygen disequilibrium is something that, as it happens, is almost always caused by life. And in the case of A it is caused by life.
> *Case 2*: We discover a planet, Planet B, with a methane-oxygen disequilibrium. A methane-oxygen disequilibrium is something that is almost always caused by life. But in the case of B it is not so caused.
> *Case 3*: We discover a planet, Planet C, with a methane-oxygen disequilibrium. A methane-oxygen disequilibrium is very rarely caused by life. But in the case of C it is caused by life.

The purpose of these cases is to bring to our attention a distinction between types and tokens. When we ask whether an atmospheric methane-oxygen disequilibrium is a biosignature, we are asking about types. Is this *type* of thing – a methane-oxygen disequilibrium – a type biosignature? When we ask whether a *specific* methane-oxygen disequilibrium is a biosignature, we are asking about tokens. Is this specific, token thing – *this* methane-oxygen disequilibrium (for example, on Planet A) – a token biosignature?

Quite generally, when we ask whether or not something is a biosignature we could be asking either about types or about tokens. The question is ambiguous. With respect to Case 1 the ambiguity goes under the radar because the answer to both the question about the type and the question about the token is 'yes'. With respect to Cases 2 and 3, the ambiguity comes to the fore. It is natural to say in Case 2 that there is a sense in which the answer to the question is 'yes' and a sense in which it is 'no'. It is 'yes' because, at the level of types, methane-oxygen disequilibria are almost always caused by life. So surely they are a biosignature. And yet it is also tempting to answer 'no' because in the case of Planet B the disequilibrium has not in fact been caused by life. Given that it has not been caused by life it would be most peculiar to say that it is a sign of life. For any x and y, how could x be a sign or mark of y if y isn't there? It could not. Similarly for Case 3. There is a sense in which it is natural to answer 'yes' and a sense in which it is natural to answer 'no'. Only here, compared to Case 2, the roles are reversed. In this case it is natural to answer 'no' at the level of types; after all, this disequilibrium does not typically accompany life, even where life is present. But it is natural to say that the answer is 'yes' at the level of tokens. Planet C's atmosphere is *in fact* caused by life. So surely it is actually a sign of that life, though this is unusual.

I propose then that we distinguish between biosignatures at the level of types and at the level of tokens. When we ask whether something is a biosignature, we could be asking about either or both.

There is a very simple and attractive way of thinking about *token* biosignatures. It is that some particular, token thing is a token biosignature if and only if it is in fact causally dependent on some living thing, or as I'll simply say 'life'. I'll put it like this:

> *Token Biosignature*: A token thing, *t*, (for example, a specific planet's atmospheric disequilibrium) is a token biosignature if and only if it is in fact causally dependent on life.

Note an important feature of this account. It is *factive*. When we're talking about tokens, this is the obvious way to go. Suppose, as in Case 2, we find that the disequilibrium in the atmosphere of Planet B is *not* caused by life. Then surely that *specific* disequilibrium was *not* a biosignature, even if disequilibria of this kind (or type) in general are. Some evidence that this is how the expression 'biosignature' is understood in the astrobiology community can be found by looking at cases in which token, specific token things that it was once hoped were caused by life turned out not to be. Once the hope was lost, the term 'biosignature' stopped being used, or the inappropriateness of its use was strongly implied, often by some modifier like 'candidate' or 'possible' or even 'false'. So it looks very much as though token biosignatures are understood factively. This is arguably clearest in discussion of the infamous 'Alan Hills' meteorite fragment (McKay 1996). Now that they are known to be abiotic, the unusual structures within that fragment are rarely if ever described as biosignatures. Consider for example the following – I think fairly representative – passage from a recent article on the subject:

> Proposed biosignatures in the ancient Allan Hills 84001 Martian meteorite are most plausibly explained as abiotic features… The controversial hypothesis that ALH 84001 contains evidence of extraterrestrial biology has mostly subsided, but it has fuelled a Mars exploration program focused on the search for life and has helped refine the criteria for the recognition of biosignatures. (McSween 2019, p. 167)

The unusual molecules in Alan Hills are here described as 'proposed biosignatures… most plausibly explained by abiotic features'. This sounds very much as though the proposed biosignatures have turned out not to be biosignatures at all. They have turned out to be *abiotic* and hence *merely* proposed. The impression is reinforced by the claim

that the falsity of the proposal has helped to 'refine the criteria for the recognition of biosignatures', the obvious reading being that the organic molecules in Alan Hills were *not* in fact biosignatures and that discovering this has helped us develop better criteria for recognising what *would* be. Or similarly, consider the recent *International Journal of Astrobiology* publication 'On Biosignatures for Mars' (Westall 2021). Writing about the Alan Hills controversy – and the paper that made it famous, McKay (1996) – the authors state:

> McKay (1996)… kick-started the field of astrobiology. Today we have a far better understanding of microbes and how their traces can be preserved. We also have a deeper awareness of associated problems, such as the distinction between a microbial biosignature and a similar abiotic signature. (Westall et al. 2021, p. 378)

Again, this sounds very much as though a distinction is being drawn between (microbial) biosignatures and abiotic signatures, the latter and not the former being that which was in fact present in Alan Hills.

The lesson is that – as one might expect – for token cases the astrobiology community is taking the concept of a biosignature to be factive. But this does not apply to types. If it did – if we required that types were similarly factive – then for some type of thing to be a type of biosignature, every token of that type would have to have been causally dependent on life. It would have to be what is sometimes called a 'universal biosignature'. If life is diverse there may be no conditions under which this restriction would be met. And so while we *could* say that there would consequently be *no* types of thing that are biosignatures this would seem unnecessarily restrictive. Suppose that the universe is teeming with life and that there are certain effects of that life which, while not universally caused by it, are very often caused by it and allow us to predict its presence as reliably as we can predict many other phenomena. Why say that, simply because these effects are not infallible indicators, they are not actually type biosignature? It would surely be better to say that it is enough for some type of thing to be a biosignature that it is *reliably* caused by life. With this in mind, the simple type-based correlate of the above account would be:

> *Type Biosignature*: For any type of thing, *T*, (for example, a type of atmospheric disequilibrium) *T* is a type biosignature if and only if tokens of *T* are reliably causally dependent on life.

We could then say that *T*'s strength as a biosignature is determined, all else being equal, by the strength of the correlation, or degree of

reliability. This is a matter of degree. There will be no correlative distinction between the strength and weakness of a token biosignature. This is an all-or-nothing matter—though we may of course be more or less sure that some token thing is a token biosignature. So epistemically, though not metaphysically, it will be a matter of degree.

5. Proposing the view (2): the substantive question

I have been trying to make progress on the factivity question. To this end I have introduced the type/token distinction. I have simply *assumed* an answer to the substantive question. Specifically, I have assumed that the relation between the things that are biosignatures and the life of which they are signatures is one of causal dependency of the former on the latter. By doing this I am ignoring a number of alternative options. It is worth setting out the possibility-space clearly here and determining whether my proposal is *too* narrow.

One option that I am ignoring is a causal relationship that runs in the opposite direction: upstream from biosignature to life. Consider the following case:

> *Case 4*: The environment of Planet D consists of some particular combination of chemical and physical properties that can properly be said to have *caused* life to emerge on that planet.

My account would entail that the (token) environment on D is not a (token) biosignature because it is not caused by life, even though it is the cause of that life. And there are yet more substantive relations in which things can stand to life; it is not limited to cause or effect. Consider the following case:

> *Case 5*: The environment of Planet E consists of some particular set of chemical and physical properties that has caused both life and, quite independently, a peculiar weather pattern. Furthermore, this is a common combination: where this particular set of chemical and physical properties exists, both the weather pattern and life reliably exist.

Now consider the weather pattern on planet E. This is neither an effect nor a cause of the life on E. But Planet E does harbour life, and this type of weather pattern predicts as much. So there is certainly a sensible question with respect to whether it (that is, the weather pattern) is a biosignature.

These cases suggest a broader account of the substantive relation between biosignature and life than that defended so far. It is easiest to see at the level of types. According to this *broad* view: a type of thing is a type biosignature iff that type *reliably indicates* life. Or:

> *Type Biosignature (Broad)*: For any type of thing, *T*, *T* is a type biosignature if and only if it reliably indicates life.

This is an account that is neutral with respect to *why* the type reliably indicates life: whether by (reliably) being an effect, or a cause, or standing in some other relation. All that matters is that it *does* reliably indicate life. This leaves open precisely how the threshold of 'reliability' is to be specified. I suggested above that this is a matter of degree. Types may more or less reliably indicate life and as such be more or less reliable biosignatures. One may regard this as overly permissive. For example, one may suggest that a T-type property counts as reliably indicating life to a sufficient degree only if life is present in *at least half* of the cases in which the T-type property is present. I shall not attempt to provide such a precise specification here. The important point for present purposes is simply that reliable indication in the present context does not imply factivity as type biosignatures, unlike tokens, are not factive.

Phrasing a correlatively broad account of *token* biosignatures is actually surprisingly tricky. While types are the kinds of things that reliably indicate, tokens aren't or at least are not obviously so. Dark clouds overhead reliably indicate rain. But it is questionably coherent to say of this *specific* dark cloud overhead that *it* reliably indicates rain. Either it will rain or it will not. But it's not the kind of thing that can *reliably* rain. This means that we can't just say that a token thing is a biosignature in the broad sense if and only if that token thing reliably indicates life. And there's a second problem. Reliable indication isn't factive. So even if we could say that tokens reliably indicate (which I've just said we can't), we *shouldn't*, because this would compromise the factivity that we want for token cases. Now you might think that we could get around both of these problems at once by saying this: a token thing is a biosignature if and only if the type of which that token is an instance reliably indicates life, and in the token case at hand, life is in fact present. But this wouldn't be right either. To see this, consider again:

> *Case 3*: We discover a planet, Planet C, with a methane-oxygen disequilibrium. A methane-oxygen disequilibrium is very rarely caused by life. But in the case of C it is caused by life.

I think we should say that C's atmosphere is a token biosignature. It was in fact caused by life. But this would be false if we were to say, as proposed above, that a token thing is a biosignature if and only if the type of which that token is an instance reliably indicates life, and in the token case at hand, life is in fact present.

So how *should* we phrase the broad, token conception of a biosignature? One option would be to simply enumerate all of the different substantive relations in which token biosignatures can stand to life. These include cause, effect, and via a 'third factor' (that is, as in Case 5). Ideally however there would be a single expression that could capture and usefully group all of these substantive relations. The simplest option is to minimally modify the expression used in discussing *type* biosignatures above. In that case we used the familiar expression 'reliably indicates'. This expression cannot be correctly applied to token biosignatures without modification; *reliability* is a property that types, not tokens, can possess. We can however say of a token thing that it *actually* or *in fact* indicates life. By this I intend that the (token) thing in question is *actually* caused by life, is an effect of life, or stands in a third-factor relation to life. If any of these obtain I shall say that the thing actually, or in fact indicates life. So, I shall say:

> *Token Biosignature (Broad)*: A token thing, t, is a biosignature if and only if it in fact indicates life.

Admittedly, in this account the expression 'in fact indicates' functions as a catch-all placeholder for the relations in question (that is, cause, effect, and third-factor). It does not provide an underlying explanation of what these have in common that makes each an instance of the same relation; the relation that biosignatures stand in to life. This is non-ideal but I shall continue to make use of this simple locution in what follows as I detail and defend the account further.

The result of the foregoing is that we now have two different kinds of account of biosignatures: narrow and broad. The former understands biosignatures as effects, the latter as indicators that need not be effects. Which should we prefer? Each has virtues. The virtue of the narrower account is that – I think – it captures what astrobiologists are in fact looking for today: things caused by life. The word 'biosignature', or 'biomarker' rather indicates this. Signatures and marks are *effects*. I shall work with the broad account however. So as I'll understand it:

> *Token*: A token thing, *t*, is a token biosignature if and only if it in fact indicates life.

> *Type*: For any type of thing, *T*, *T* is a type biosignature if and only if it reliably indicates life.

Why work with this broad account? When it comes to the search for life it doesn't much matter whether we find it because we find something that it causes, something it is caused by, or something that indicates it in some other way. What matters is simply that we find it; whether it's a cause, an effect or something else entirely doesn't really matter. The broad account speaks to this. If, however, the reader either prefers the narrow account, or prefers to be ecumenical, that's fine. The arguments that follow would be equally applicable whether they are used in defence of narrow or broad accounts.

6. Subjective accounts

On the view defended above, whether something is a biosignature is an *objective* matter. It is determined entirely by features of the world and not in any way by features of our understanding or cognition. This is clear on the narrow account of biosignatures according to which biosignatures are effects. Whether *y* is an effect of *x* is dependent only on the relation between *x* and *y* and in no way on features of the human understanding. It could be that *y* is an effect of *x* but that nobody knows it or could ever know it. The same is true on the broad account. Suppose, as is possible on that account, that life is the cause of some biosignature and not the effect of it. Again, this is an objective matter. Generally, as I am understanding it, *indication* is an objective matter. For *a* to indicate *b* is for *a* to stand in some objective relation to *b*; a relation that obtains if at all independently of features of the human understanding.[4]

This objective way of understanding biosignatures contrasts with at least some of the accounts from the literature discussed earlier. Consider for example:

> *W1*: A biosignature is known to be an indicator of biological activity (in a given planetary context).

This is *subjective* in the following sense: whether something is known to indicate biological activity is determined – in part – by facts about people; specifically, what they know. This is a very different kind of view. To see the difference consider the following case:

[4] This is roughly in keeping with standard ways of using the word. See for example Kelly (2008, 2014).

> *Case 6*: The environment of Planet F displays an environmental profile that is caused by life both reliably and in this specific case. So the environmental profile of F indicates life. But we do not presently *know* that this environmental profile indicates life (either reliably at the level of types, or in fact, at the level of tokens).

On Walker's view (W1) this is *not* a biosignature because it is not known that the atmosphere indicates life, even though it does. Which view is correct; the subjective or the objective? My aim in this section is to argue for the objective view and against the subjective view. I'll take W1 as an example of a subjective view, though I intend the basic point to generalise to other subjective views too.

6.1. First Argument

The first argument is that subjective views struggle with the many modifications of 'biosignature' that figure in the literature, such as 'potential biosignature' or 'known biosignature'. Making this point requires a brief preliminary.

There are facts about whether a thing indicates the presence of life, and facts about whether we know this. Objective accounts of biosignatures keep these separate and identify biosignatures with the former. Subjective accounts do not keep them separate. On subjective accounts, the facts about what we know fall within the scope of what it is to be a biosignature. That makes the subjective concept of a biosignature a complex thing. To see this, look at W1. Here, the objective concept 'indicator of life' really falls *within* the subjective concept of what it is to be a biosignature. The structure is complex—much more so than my objective account in which we do not find the complex nesting of different concepts within the account.

With this in mind we can turn to the argument against the subjective view. The simple structure of the objective concept of biosignature makes it much easier to modify in simple, intuitive ways. The more complex structure of the subjective concept makes it much harder. On the subjective view, simple and commonly used modifications have to be interpreted in very peculiar ways.

Consider, for example, the complex concept of a *potential biosignature*, frequently deployed in the astrobiology literature. On my objective account, this is something that is potentially an indicator of life. That is to say, it is something that could be an indicator of life. Suppose that a subjective account like W1 is true, however. Then a potential biosignature is something that could be known to be an indicator of life. This

is quite peculiar. When practitioners say that something is a potential biosignature, they're (surely) not saying that *it could be that we know* it is an indicator of life. If this were what they were saying, they would be telling us a sociological fact about themselves: a fact about what it could be that they know. I don't think that's what they're trying to do. They're trying to tell us that it could be an indicator of life. Only the objective view gets this result.

Or consider another complex concept: that of a *known biosignature*. Suppose that a scientist makes an incredible discovery in a sample taken from below the surface of Mars: a pattern left by a mini fossil! After careful study he presents his findings, claiming that he has discovered a biosignature. A colleague asks: 'Do you know that this is a biosignature?'. This sounds like a perfectly normal thing to ask. And on the objective view it is. On the objective view, the colleague is asking whether the scientist knows that what he has found is an indicator of life. On the subjective view, on the other hand, the colleague is asking a very different question. He is asking a question about higher order knowledge. While he *could* be asking this, it would be a most peculiar question, and I would be surprised if anyone would interpret the question in this way. It's surely not what he meant. But the subjective view entails that it is.

The argument can be summarised as follows. Complex modifications of the concept of a biosignature such as that of a *potential biosignature* or a *known biosignature* are commonly used. Their meanings are much more plausibly understood following the simpler, objective conception than the subjective conception.

6.2. Second Argument

If subjective views are right, something's status as a biosignature varies as our subjective states vary. This is not plausible. Consider again:

> *Case 6*: Planet F displays an environmental disequilibrium that is caused by life both reliably and in this specific case. So the environmental profile of F indicates life. But we do not presently know that this environmental profile indicates life.

Suppose now that we look at Planet F's atmosphere through a telescope. On the objective view we are in fact looking at a biosignature but we do not presently know this. On the subjective view, we are not looking at a biosignature at all. Suppose now that we come to know that Planet F's atmosphere is in fact an indicator of life. On the objective view we have *discovered* that it is a biosignature. Our knowledge has allowed us to make this discovery. On the subjective view by contrast our knowledge

has turned something that was not a biosignature into a biosignature. So we haven't – or haven't merely – discovered that it is a biosignature. Rather, we have made it into or caused it to be a biosignature. This seems wrong. We did not create or cause a biosignature to come into being. It was extraterrestrial life that did *that*! Rather, we discovered that something that was always a biosignature is a biosignature. So the subjective view is wrong and the objective view is right. This is *the argument from discovery and creation*. It shows that something's status as a biosignature does not vary with our subjective states in the way that subjective views predict.

The second argument is similar. Consider the following case.

> *Case 7*: Planet G appears to have bright, artificial lights on its surface. We correctly identify these as being very possibly caused by life. We are right. We attempt to send probes to this distant planet. They will take ten thousand years to reach it. However, human civilisation regresses in this time and at some point we cease to know that these lights exist or are an indicator of life on Planet G.

What should we say in this case? If the objective view is true, then the lights on Planet G are a biosignature, we discover this, and then we lose this knowledge. On the subjective view we do not *merely* lose this knowledge. Rather, when, human civilisation regresses, Planet G's bright artificial lights cease to be a biosignature. This seems wrong. Planet G's lights do not cease to be a biosignature. We have simply lost this knowledge. Call this *the argument from lost knowledge*.

In both of these arguments – *from discovery and creation* and *from lost knowledge* – we have drawn out a consequence of letting something's status as a biosignature vary with our subjective states. Biosignatures become the kinds of things that exist and cease to exist as our knowledge varies. This is awkward. It is certainly much less intuitive than the simple results that the objective view allows: results according to which it is not the *existence* of biosignatures that varies with our knowledge, but rather our knowledge of their existence.

7. Conclusion: the nature and function of the concept *biosignature*

I have proposed and defended an account of what it is to be a biosignature. Biosignatures are factive indicators of life in the token case, or reliable indicators of life in the type case. Indication is an

objective relation that is paradigmatically, but not exclusively, one of cause and effect. I claimed earlier – in outlining my method – that the conception of a biosignature that I would delineate plays a useful role in astrobiology. We are now in a position to ask what, exactly, that role is.

The concept I have delineated performs two important tasks; tasks that we need a single concept to play simultaneously. 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. 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 given that for all we know they may be extremely varied and perhaps even surprising (Vickers 2020). 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, fills two roles. 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 all but essential for astrobiology and the search for extraterrestrial life. My 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. Accounts that do *less* will be inadequate to the task that the concept is needed to perform. Accounts that try to do more, like subjective accounts, will in any case have to rely on the concept that I have articulated as a constituent. So there is a need for *this* concept. And it is *this* concept that the word 'biosignature' functions to pick out in its core, ordinary use.

References

Baque, Mickael et al. 2022, 'Biosignature stability in space enables their use for life detection on Mars', *Science Advances* 8(36)

Catling, David et al. 2018, 'A Framework for assessing exoplanet biosignatures', *Astrobiology* 18(6): 709-738

Cleland, Carol 2019, *The Quest for a Universal Theory of Life. Searching for Life as We Don't Know It* (Cambridge: Cambridge Astrobiology)

Duner, David, Persson, Erik, and Holmberg, Gustav 2013, *The History and Philosophy of Astrobiology* (Cambridge Scholars Publishing: Newcastle Upon Tyne)

Duner, David 2019, 'The History and Philosophy of Biosignatures', in B. Cavalazzi and F. Westall (eds.) *Biosignatues for Astrobiology* (Cham, Switzerland: Springer)

Des Marais, David and Walter, Malcolm 1999, 'Astrobiology: Exploring the origins, evolution, and distribution of life in the universe', *Annual Review of Ecology and Systematics* 30: 397–420

Des Marais, David et al. 2002, 'Remote sensing of planetary properties and biosignatures on extrasolar terrestrial planets', *Astrobiology* 2: 153–181

Des Marais, David et al. 2008, 'The NASA Astrobiology Roadmap', *Astrobiology* 8: 715–30

Des Marais, David 2018, 'Astrobiology Goals' in V. Kold (ed.) *Handbook of Astrobiology* (Boca Raton: CRC Press).

Hickman-Lewis, Keyron et al. 2022, '*In Situ* Identification of Paleoarchean Biosignatures Using Colocated Perseverance Rover Analyses: Perspectives for *In Situ* Mars Science and Sample Return', *Astrobiology* 22(9): 1143-1163

Kelly, Tom 2008, 'Evidence: Fundamental Concepts and the Phenomenal Conception', *Philosophy Compass* 3(5): 933-955

Kelly, Tom 2014, 'Evidence', *Stanford Encyclopedia of Philosophy*. (Winter 2016 Edition)

Krissansen-Totten, Joshua et al. 2018, 'Disequilibrium biosignatures over Earth history and implications for detecting exoplanet life', *Science Advances* 4(1): eaao5747

Lingam, Manasvi and Loeb, Avi 2021, *Life in the Cosmos: From Biosignatures to Technosignatures* (Cambridge, Mass: Harvard University Press)

Mangold, Nicolas et al. 2021, 'Perseverance rover reveals an ancient delta-lake system and flood deposits at Jezero crater, Mars', *Science* 374(6568): 711-717

McKay, David et al. 1996, 'Search for past life on Mars: Possible relic biogenic activity in Martian meteorite ALH84001', *Science* 273: 924-930

McMahon, Sean and Cosmidis, Julie 2022 'False Biosignatures on Mars: Anticipating Ambiguity', *Journal of the Geological Society*: 179(2): 2021-2050

McSween, Harry 2019, 'The Search for Biosignatures in Martian Meteorite Allan Hills 84001', in B. Cavalazzi and F. Westall (eds.) *Biosignatures for Astrobiology* (Cha, Switzerland: Springer)

Pohorille, Andrew and Sokolowska, Joanna 2020, 'Evaluating Biosignatures for Life Detection', *Astrobiology* 20(10): 1236-1250

Thompson, Maggie et al. 2022, 'The case and context for atmospheric methane as an exoplanet biosignature', *Proceedings of the National Academy of Sciences* 119(14): e2117933119

Vickers, Peter 2020, 'Expecting the unexpected in the search for extraterrestrial life' *International Journal of Astrobiology* 19(6): 482-491

Walker, Sara et al. 2018, 'Exoplanet biosignatures: Future directions' *Astrobiology* 18(6): 779-824

Westall, Frances et al., 2021 'On biosignatures for Mars' *International Journal of Astrobiology* 20(6): 377-393

