



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Confidence of Life Detection: The Problem of Unconceived Alternatives

**Citation for published version:**

Vickers, P, Cowie, C, Dick, SJ, Gillen, C, Jeancolas, C, Rothschild, LJ & McMahon, S 2023, 'Confidence of Life Detection: The Problem of Unconceived Alternatives', *Astrobiology*.  
<https://doi.org/10.1089/ast.2022.0084>

**Digital Object Identifier (DOI):**

[10.1089/ast.2022.0084](https://doi.org/10.1089/ast.2022.0084)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Astrobiology

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Confidence of Life Detection: The Problem of Unconceived Alternatives

---

Peter Vickers<sup>(a)</sup>, Christopher Cowie<sup>(a)</sup>, Steven J. Dick<sup>(b)</sup>, Catherine Gillen<sup>(a)</sup>, Cyrille Jeancolas<sup>(a)</sup>,  
Lynn J. Rothschild<sup>(c)</sup>, Sean McMahon<sup>(d)</sup>

(a): Durham University, UK; (b) NASA Chief Historian (retired); (c) NASA Ames Research Center, USA; (d) University of Edinburgh, UK

---

## Abstract

Potential biosignatures that offer the promise of extraterrestrial life (past or present) are to be expected in the coming years and decades, whether from within our own solar system, from an exoplanet atmosphere, or otherwise. With each such potential biosignature, the degree of our uncertainty will be the first question asked. Have we really identified extraterrestrial life? How sure are we? This paper considers the problem of unconceived alternative explanations. We stress that articulating our uncertainty requires an assessment of the extent to which we have explored the relevant possibility space. It is argued that, for most conceivable potential biosignatures, we currently have not explored the relevant possibility space very thoroughly at all. Not only does this severely limit the circumstances in which we could reasonably be confident in our detection of extraterrestrial life, it also poses a significant challenge to any attempt to quantify our degree of uncertainty. The discussion leads us to the following recommendation: when it comes specifically to an extraterrestrial life detection claim, the astrobiology community should follow the uncertainty assessment approach adopted by the IPCC.

## 1 Introduction

Due to important technological advances in analytical instruments and new space missions, we should be prepared for the detection of potential biosignatures of extraterrestrial life in the near future. Data from the surface of Mars, for example, will be abundant over the next twenty years as a consequence of rovers such as NASA's Perseverance, the Chinese National Space Administration's Zhurong, and the planned sample return missions. Data from exoplanets, including exoplanet atmosphere composition, will also be abundant over the next twenty years in particular because of the deployment of the James Webb Space Telescope. With each potential biosignature, the degree of our uncertainty will be the first question asked. How sure are we that the signal actually derives from extraterrestrial life? Despite its clear importance, the challenge of assessing uncertainty in biosignature detection has thus far only briefly, and only very recently, been seriously addressed in the literature (Almár & Race, 2011; Green et al., 2021; Meadows et al., 2022; NASEM, 2022b). Moreover, the "problem of unconceived alternatives" discussed by philosophers of science (Stanford, 2006) is highly relevant but almost completely absent from the astrobiology literature.

This paper highlights one particularly significant, underappreciated challenge for the uncertainty assessment frameworks proposed in the literature, especially Green et al. (2021) and Meadows et al. (2022). Specifically, we bring to bear the problem of unconceived abiotic explanations for phenomena of interest. Drawing on the NfoLD/NExSS community report (Meadows et al., 2022, p. 26), we stress that articulating our uncertainty requires an assessment of the extent to which we have explored the relevant possibility space. It is argued that, for most conceivable potential biosignatures, we currently have not explored the relevant possibility space very thoroughly at all. Not only does this severely limit the circumstances in which we could reasonably be confident in our detection of extraterrestrial life, it also poses a significant challenge to any attempt to quantify our degree of uncertainty. The discussion leads us to the following recommendation: when it comes specifically to an extraterrestrial life

detection claim, the astrobiology community would do well to adopt a thoroughly time-tested framework – that utilized within IPCC reports.

## 2 The challenges of known and unknown false positives

Increasingly, we have at our disposal technologies and methods for analyzing the composition of exoplanet atmospheres (Catling et al., 2018). Suppose that one day soon we detect oxygen in the atmosphere of an exoplanet.<sup>11</sup> Moreover, that exoplanet is a rocky planet in the habitable zone of a “friendly” star. Is it a biosignature? There are a range of definitions of biosignature out there (des Marais et al., 2003, p. 234; Catling et al., 2018, p. 710; Schwieterman et al., 2018, p. 666); the important question to ask is how sure we are that the oxygen we have detected has a biotic cause. At one time there were no plausible abiotic explanations of accumulation of oxygen in the atmosphere of a planet. That is no longer the case; we now know of several different abiotic pathways to such oxygen accumulation (see, e.g., Meadows *et al.* (2018)). These abiotic explanations – previously unconceived – would now give us pause for thought if oxygen was detected (greater pause for thought than *before* we had such abiotic explanations).

This is a cautionary tale. And in fact there are many such cautionary tales or “lessons from history” in the recent decades of the field of astrobiology (Green et al., 2021, p. 575; Meadows et al., 2022, p. 51). The worry is that we tend to jump too quickly from “There is no known abiotic explanation of  $\varphi$ ” to “ $\varphi$  is probably caused by life.” This is an inference that has been made on several occasions in the history of astrobiology. The following three examples serve to demonstrate that this is not merely a hypothetical concern.

First consider Sinton’s (1957) interpretation of a curious, apparently Martian absorption spectrum as evidence for the presence of organic molecules and even “vegetation” on Mars. He argued that the evidence made it “extremely likely that plant life exists on Mars” (p. 239).

Sinton (1959) continued the argument, dismissing carbonates as the cause and concluding that “[T]hese bands are most probably produced by organic molecules” (p. 1237). Several years later, Shirk et al. (1965) put forward an abiotic explanation, and soon after it turned out that the bands were due to deuterium in the atmosphere of Earth, not Mars (Rea et al., 1965). Sinton’s conclusions were surely too hasty, and the form of expression too confident; Dick (2020, p. 697) purported that he was “undoubtedly affected by preconceived ideas.”

Next consider the famous case of ALH84001, a Martian meteorite discovered in Antarctica in 1984. McKay et al (1996, p.929) accepted from the start that *individual* characteristics of the meteorite were easy enough to explain abiotically. But they contended that the full suite of characteristics – taken as a package – couldn’t be explained abiotically and was thus good evidence of past microbial life on Mars. This conclusion was gradually undermined by work exploring “unconceived alternatives” (Golden et al., 2001; Martel et al., 2012). Arguably, the final word on the matter came in 2022 when Steele et al. (2022) presented evidence that mineral carbonation and serpentinization reactions on early Mars were the cause of the claimed evidence of microorganisms.

Finally, consider the work of Nutman et al. (2016). The authors proposed the existence of stromatolites in the Isua supercrustal belt of southwestern Greenland that dated back 3,700 million years, extending the fossil record back by over 200 million years. To reach this conclusion, the authors addressed four areas whereby the structures fail to be explained by known abiotic processes, and hence: “on these grounds, we rule out an abiogenic origin for Isua stromatolites” (p. 3). But then, two years later, Allwood et al. (2018) offered a plausible abiotic explanation for the proposed stromatolites. The explanation largely attributes the apparent stromatolites to structural deformation and chemical alterations of layered rock.

Nutman et al. (2016) *might* still be right – see their response (Nutman et al. (2019))– but the existence of an abiotic explanation will certainly give the community much pause for thought.

Given such examples (and there are many more<sup>21</sup>) we need to be highly cautious in a situation where we *currently* can't think of a plausible abiotic explanation for the phenomenon in question. A scholar heavily influenced by the noted cautionary tales might even *expect* that, in due course, a plausible abiotic explanation of a current *biosignature* will likely be developed. It isn't obvious when such an expectation would be irrational.

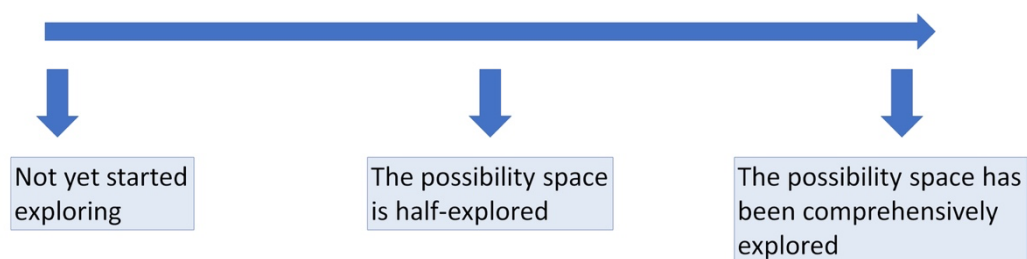
To illustrate, consider now a much stronger candidate for a genuine biosignature: oxygen-methane disequilibrium. Such a disequilibrium in the atmosphere of a planet is readily explained by the presence of life (Krissansen-Totton et al., 2018). We have a strong oxygen-methane disequilibrium in the atmosphere of *our* planet, for example, and it is caused by life. We do have plausible abiotic stories to tell concerning very *weak* oxygen-methane disequilibrium in the atmosphere of a planet, but not (yet) concerning *strong* oxygen-methane disequilibrium, such as that found on Earth (Simoncini et al., 2013; Thompson et al., 2022).

Suppose now that we detect such a signal. It would certainly be called a “biosignature” by some scientists and journalists. But how confident could we be that it was caused by life? The fact that we can't think of an abiotic cause would no doubt tempt some commentators to conclude that it is *probably* caused by life. In the work of Krissansen-Totton et al. (2018), we find the following:

The methane flux required to sustain observed quantities of methane in the modern Earth's oxidizing atmosphere is greater than what abiotic processes could plausibly provide, and thus, biological methane leakage must be invoked to explain the persistent disequilibrium.

But we might worry that the word *must* is too strong, especially when we look at other cases – such as those sketched above – where abiotic explanations were previously unconceived and then later conceived. The fact that we have a range of such examples from the history of astrobiology constitutes a cautionary tale: perhaps we should really *expect* that a plausible abiotic explanation of a strong oxygen-methane disequilibrium will one day be developed.

This is one consideration. Another comes from considering the extent to which we have explored the relevant possibility space. Consider two extremes: one where we haven't even started exploring possible abiotic explanations of an oxygen-methane disequilibrium, and another where we have explored the topic for decades without finding a plausible abiotic explanation. In the first scenario, there should be no excitement at all if we detect an oxygen-methane disequilibrium since, as far as we know, highly plausible abiotic explanations may exist. In the second scenario, we should be *very* excited, since we are approaching the point where we are *sure* that such a disequilibrium must have a biotic cause. Thus, to judge how excited we should be about detecting an oxygen-methane disequilibrium, we need to judge where we are on the *spectrum of exploration*, with the two scenarios given above at its extremes (Fig. 1).



**Fig. 1** The spectrum of exploration; the extent to which we have explored a certain possibility space, ranging from “not yet started” to ‘fully explored.’ In many scenarios we won’t actually know when the space is half-explored or fully explored.

As Meadows et al. (2022, p. 26) noted, “[I]f the scope of possible abiotic explanations is known to be poorly explored, it suggests we cannot adequately reject abiotic mechanisms.”

Conversely, if it is known to be *thoroughly* explored, we probably *can* reject abiotic mechanisms.

When it comes to the specific case of atmospheric methane, Krissansen-Totton et al. (2018) stated that, “On the basis of current understanding, the conditions required to generate large fluxes of abiotic methane are specific and implausible [p. 8].” But they say nothing about the *extent* of our current understanding. Thus, even though they discuss a handful of possible abiotic methane sources, the reader is left in the dark on the question of just how much relevant possibility space might still be “out there” waiting to be explored. But if we don’t know this, we can’t assess the relevant uncertainty.

To put it another way, there are two completely different problems here – currently not clearly distinguished in the literature<sup>21</sup> – both coming under the general heading of “abiotic mimics,” or “false positives.” On the one hand, it can already be extremely challenging to consider all currently known potential abiotic mimics for a given signature and adequately rule them out. The extent to which this has been done in a given case is going to be challenging to measure but needs to be measured if the degree of uncertainty is to be adequately articulated. On the other hand, *even if* one has thoroughly ruled out all plausible abiotic mimics *given current knowledge*, we cannot articulate our degree of uncertainty – or “confidence of life detection” – until we have also somehow factored in the extent to which we have explored the space of possible abiotic mimics. Ruling out all *known* abiotic mimics is little comfort if our knowledge of possible abiotic mimics is in its infancy.

We argue that astrobiology is still a young science, and research into “abiotic mimics” is in its infancy. That is, in most relevant contexts, we are much closer to the “under-explored” end of the spectrum than the “thoroughly explored” end. Consider McMahon and Cosmidis (2022) writing on “false biosignatures on Mars”:



[K]ey evidence [for confirming possible biosignatures] will come from the investigation of abiotic physicochemical systems and their capacity to mimic the forms and properties of life. Yet this area of enquiry has received rather scant and unsystematic attention from astrobiologists, who have tended to focus their published work on expanding our knowledge of life's signatures rather than its abiotic mimics. [...] The reliability of any detected biosignatures on Mars therefore depends crucially on our understanding of the abiotic processes that might mimic them. [...] [B]iogenicity criteria are unable to discriminate sensitively and reliably between biosignatures and pseudobiosignatures unless they are grounded in extensive knowledge and understanding of both classes of phenomena. [...] However, most known varieties of pseudobiosignature have not been characterized or understood in sufficient detail for this to be possible. Moreover, given the haphazard and unsystematic way in which varieties of false biosignature have so far been identified, we can only assume that many others remain undiscovered.<sup>[4]</sup>

The authors here considered abiotic mimics on Mars, similar to the case of Martian meteorite ALH84001. But biosignatures associated with exoplanet atmospheres are not disanalogous in the relevant respects. Here, too, we may reasonably assume that many “false biosignatures” remain undiscovered.

Thus, we must embrace the thought that we have *not* thoroughly explored the relevant possibility space. We simply can't say whether an abiotic mimic of a strong methane-oxygen disequilibrium is a serious possibility. The fact that we currently can't imagine such an abiotic mimic seems like scant reason to believe that no such mimic exists. When we hear scholars say of oxygen and methane, “that combination is *very* hard to explain [abiotically]”<sup>[5]</sup>, perhaps we should take that as saying more about our current state of knowledge than about what is or isn't possible.

In fact, Krissansen-Totton et al. (2018) provided one example of significant abiotic methane production: “For terrestrial planets with a more reducing mantle than Earth, significant CH<sub>4</sub> outgassing is conceivable” (p. 7). They go on to claim that such an abiotic methane source could be identified via the presence of CO, on the grounds that “CO has few abiotic sinks.” Thus, the issues multiply: how thoroughly explored is the possibility space of abiotic CO sinks? Not very thoroughly at all, we would suggest, but the more fundamental point is that the extent of possibility space exploration needs to be included in any (un)certainty assessment.

### 3 Frameworks for life detection and the certainty continuum

Green et al. (2021) are to be commended for offering an initially plausible “Confidence of Life Detection” scale, the “CoLD” scale (Fig. 2). The proposed scale runs from 1 to 7, where 7 indicates *very* high confidence of detection of extraterrestrial life: the presence of life has been confirmed. It is explicitly put forward as a framework for clear communication with the general public; it is a “progressive one-dimensional scale” that can serve to communicate, with a number between 1 and 7, how confident we are that we have detected extraterrestrial life. We may now ask how it handles the two false positive challenges noted in the previous section, since both must be factored into any certainty calculation.



**Fig. 2:** The CoLD scale, intended as a measurement of the degree of confidence for a particular life detection claim (from (Green et al. 2021))

Most obviously, at Level 4 of the CoLD scale we find,

*“All known non-biological sources of signal shown to be implausible in that environment.”*

As noted in the previous section, the extent to which we have ruled out *known* non-biological sources can itself be extremely challenging. But what of currently unconceived abiotic mimics? If research on abiotic mimics is in its infancy, then the fact that all *known* non-biological sources have been ruled out is very weak evidence that there isn't a non-biological source. This thought is strengthened by the fact that there have been unconceived abiotic mimics in the past (later conceived).

We argue that meeting the “Level 4” requirement on the CoLD scale (having already met Level 1 – 3 requirements) can correspond to both (i) *very high* confidence (when we have more-or-less exhausted the space of possible abiotic mimics, and shown them all to be implausible), but also (ii) *very low* confidence (when scientific knowledge of possible abiotic mimics is in its infancy). In case (i), we imagine that we have *exhausted* the relevant possibility space (many teams working over many decades), and we haven't found *any* plausible abiotic explanations of the phenomenon in question. In that scenario, the fact that all *known* abiotic sources of the signal are implausible entails that a biotic cause is the only plausible explanation; this is the “no alternatives argument” (Dawid et al., 2015) or what Cowie (2022) calls the “argument from elimination” in the context of the Oumuamua debate. On a scale of confidence with 7 steps, this should be close to a “7”.<sup>64</sup> In case (ii) we imagine that we haven't even started exploring the relevant possibility space. In this scenario, the fact that all *known* abiotic sources of the

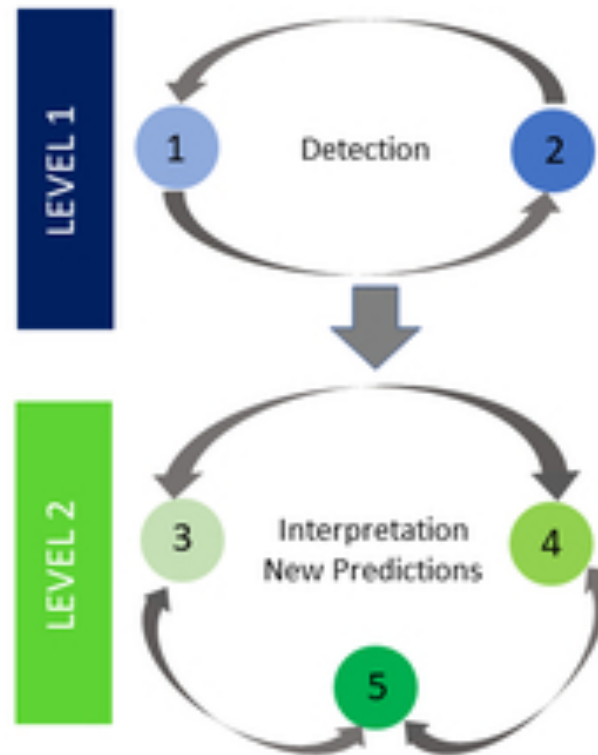
signal have been ruled out (Level 4 met) means nothing. As far as we know, there may be many plausible abiotic explanations.

If meeting the “Level 4” requirement on the CoLD scale can sometimes mean “high confidence” and sometimes mean “low confidence,” then this is a potential point of confusion for the general public. The basic idea of the 1 – 7 CoLD scale framework was to map distinct and specifiable scientific developments to the “certainty continuum,”<sup>[2]</sup> where a bigger number means that one is more certain. The problem of unconceived alternatives poses a serious challenge to this goal.

Any criticism of the CoLD scale shouldn’t be overstated, since it was merely “the beginning of an important dialogue” (Green et al., 2021, p. 575) and “Discourse within the broader community should modify or supplant the scale” (p. 578). In 2021-22, the option to “supplant the scale” was taken up in a serious way by the “Standards of Evidence for Life Detection Community Workshop” (July 19-22, 2021), which ultimately led to a many-authored white paper: Meadows et al. (2022). Here the CoLD scale is left behind, and a new framework proposed. The NfoLD/NExSS community of scholars behind this white paper apparently felt uncomfortable with the Green et al. (2021) attempt to map specifiable scientific developments to a one-dimensional numeric scale representing overall “confidence.”<sup>[3]</sup> Though they don’t explicitly criticize the CoLD framework, they do present their own framework based on five “framework questions” (see Fig. 3), and state:

While the framework questions are presented in order, there was a strong sense at the workshop that the application of the steps do not need to be in a particular order, and in some cases may be difficult to implement linearly. (Meadows et al., 2022, p. 7)

The rejection of “linearity” (other than a distinction between “Level 1” and “Level 2” – see Fig. 3), and preference for “iteration,” is stressed throughout the white paper.



**Fig. 3:** The NfōLD/NExSS framework for life detection assessments. The five questions are: (1) Have you detected an authentic signal? (2) Have you adequately identified the signal? (3) Are there abiotic sources for your detection? (4) Is it likely that life would produce this expression in this environment? (5) Are there independent lines of evidence to support a biological (or non-biological) explanation? From Meadows et al. (2022), with permission.

We may now ask: Where does this leave the “continuum of confidence,” or “certainty continuum” that Green et al. (2021) took to be at the heart of their framework? We agree with Meadows et al. (2022) that scientific methodology is far less linear, and far more iterative, than

the CoLD framework allowed. And yet it remains the case that *confidence of life detection* should be something one can in principle map onto a linear scale ranging from “no confidence at all” to “full confidence.” With the move from the CoLD scale to the NfoLD/NExSS framework, we correct the presentation of scientific methodology, at the expense of de-prioritizing the goal to effectively communicate confidence of life detection. The NfoLD/NExSS white paper leaves this question behind, except to say that “additional community discussion” will be needed to develop a “ranking or numerical scheme for certainty in biosignature detection and interpretation.” (p. 60). Apparently, the best we can do is say that we can be confident to the extent that questions 1 – 5 in the NfoLD/NExSS framework have been satisfactorily answered.

#### 4 Bayes to the rescue?

Bayesian techniques are often put forward to quantify uncertainties in science quite generally. And Bayesian techniques are often stressed in the astrobiology literature in particular.<sup>[9]</sup> In principle, the Bayes’ formula can be used: we can input values into the formula in order to determine the probability of life, given the evidence, where in this particular case the evidence is a strong methane-oxygen disequilibrium (without accompanying CO). But do we know how to fill in the terms in the equation? If we don’t, then we have a case of *obscurum per obscurius*,<sup>[10]</sup> but applied to assessing uncertainty: employment of the equation introduces as much or more uncertainty than the uncertainty we hoped to address.

Here is the relevant Bayesian formula, where we have this new evidence ‘E’ and we want to know whether we have really detected life ‘L’:

$$p(L, E) = \frac{p(E, L)p(L)}{p(E)} = \frac{p(E, L)p(L)}{p(E, L)p(L) + p(E, \neg L)p(\neg L)}$$

Perhaps the most difficult thing we need to know is the term  $p(E \mid \neg L)$ , which asks for the probability of a strong methane-oxygen disequilibrium (the evidence ‘E’) caused by some *abiotic* means (not life, ‘ $\neg L$ ’). On the one hand, we (the relevant scientific community) can’t currently think of any such abiotic story, and this might suggest a very *low* probability. On the other hand, we have a bunch of examples, from the history of astrobiology, of previously unconceived abiotic explanations of phenomena being developed where previously the only known explanation involved the presence of life. Moreover, we know that we are very far from exhausting the possibility space of abiotic mimics (recall the point, above, about abiotic CO sinks). This all needs factoring in.

As discussed, it seems reasonable to say that we’ve only just started exploring abiotic mimics. Thus, the responsible thing to say to someone who demands a value for  $p(E \mid \neg L)$  seems to be, “We just don’t know; we haven’t done the research.” If we are subjective Bayesians then we have the option to input a subjective value for  $p(E \mid \neg L)$ , and if we really have no idea, we might choose 0.5. But this seems controversial, since we really have no idea if the probability is anywhere near to 0.5. As Stanford (2011, p. 898) wrote in another context,

The austere Bayesian apparatus does promise to allow us to formally integrate the confirmational significance of various diverse forms of evidence, but this remains a promissory note when *we have no way to responsibly determine likelihoods*. (emphasis added)

If an individual insists that 0.5 is her actual credence for  $p(E \mid \neg L)$ , then the posterior probability merely has meaning for her as an individual, and doesn’t say anything about the actual, objective probability that the signal has a biotic cause.

Another option is to use imprecise probability theory, essentially introducing an interval for  $p(E \mid \neg L)$ . We might even conduct a survey of community opinion and (discounting outliers)

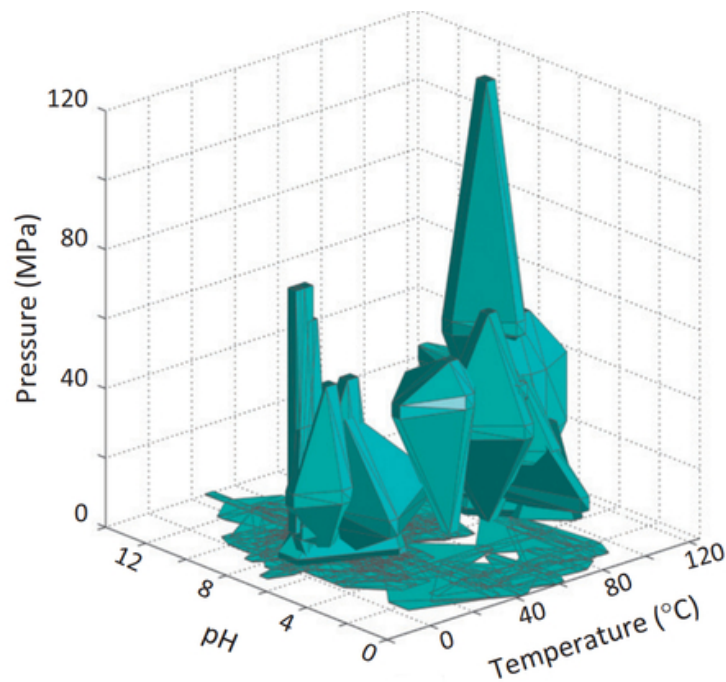
use the resultant range of values to define the interval. Of course, many of those surveyed might reasonably respond “I have no idea until much more research has been done,” which could only correspond to a maximum interval of [0,1].

The Bayes approach isn't completely devoid of information. It does show us clearly that, in a scenario where the possibility space has *definitely been exhausted* and no plausible abiotic explanation has been found,  $p(E \mid \neg L)$  would be zero, and thus our posterior would be 1, *whatever* our prior  $p(L)$  – this is the “no alternatives argument” again (Dawid et al., 2015). But most other scenarios present immediate problems. In a more realistic scenario where we have (roughly speaking) *half-explored* the possibility space, it isn't clear what would be a reasonable value for  $p(E \mid \neg L)$ . In the half of the space that we have *not* explored there could either be zero plausible abiotic stories to tell or several.

## 5 A way forward – the IPCC uncertainty framework

One option for moving forward is to try to assess the extent to which the relevant possibility space has been explored. For some such spaces, the parameters are known and well-defined, such as those considered by Harrison et al. (2013), who deliberated the limits for life under multiple extremes. Their figures 1-4 show *not* the limits for life under multiple extremes (temperature, pH, pressure, salt concentration) but, rather, our knowledge given the current extent of exploration of the space (see Fig. 4 for an example). One can more-or-less see, in the figures, the extent to which the space has been explored.<sup>uuu</sup>





**Fig. 4:** The *currently known* boundaries for terrestrial life determined as a function of temperature, pH, and pressure. From Figure 3(b) of Harrison et al. (2013), with permission.

But in the case of possible abiotic mimics for, say, a  $\text{CH}_4 - \text{O}_2$  atmospheric disequilibrium, it isn't nearly as clear how to define the possibility space. Probably it shouldn't be conceived literally as a "space," comparable with the figures of Harrison et al. (2013); probably we should think in terms of a *list* of possible abiotic causes.<sup>[12]</sup> In which case, how would we know that the list was complete? Or even half-way complete? In a scenario where we have only half-way explored the possibilities, there would still be no responsible value to input for  $p(E | -L)$ , as discussed in the previous section. Inputting a responsible value for this term requires that our exploration of the space of possibilities has been thorough (ideally exhaustive).

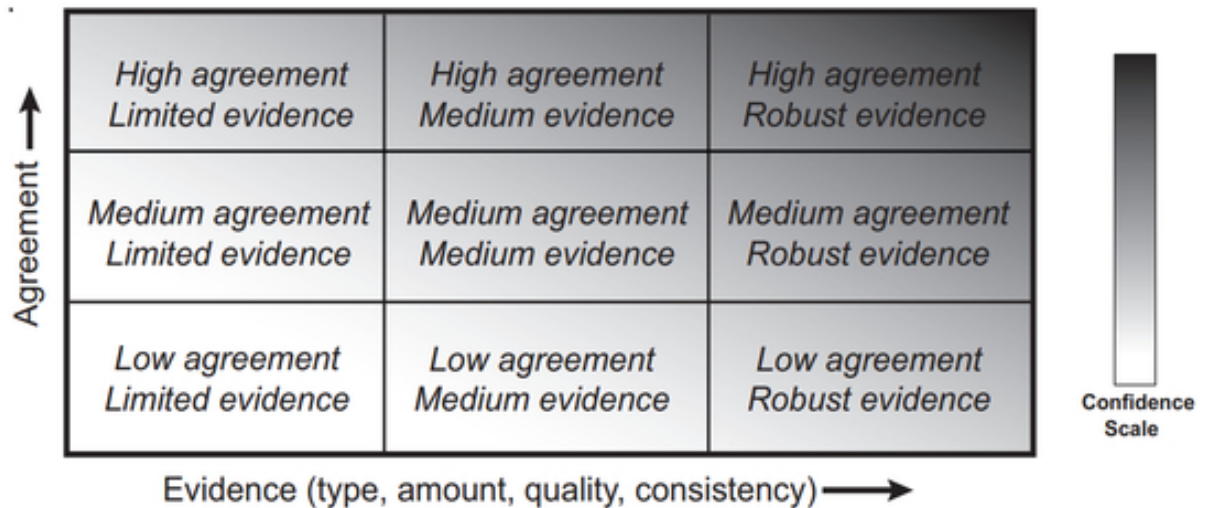
Thus, we propose a rather different approach, drawing on the uncertainty language framework (see Fig. 5) used in IPCC reports, including the most recent set of reports known as “AR6” (reports of Working Groups I, II, and III; see [www.ipcc.ch](http://www.ipcc.ch)).<sup>[13]</sup> To give just one of many possible examples, consider this statement from Hoegh-Guldberg et al. (2018, p. 178):

The ocean has absorbed about 30% of the anthropogenic carbon dioxide, resulting in ocean acidification and changes to carbonate chemistry that are unprecedented for at least the last 65 million years (high confidence).

Note the term “high confidence.” Most statements in this report include some reference to confidence, with six variations:

1. low confidence
2. low-medium confidence
3. medium confidence
4. medium-high confidence
5. high confidence
6. very high confidence

One finds almost exactly the same framework in the work of Moss and Schneider (2000, pp. 44, 45), Figs. 3 and 4, where “continuum of confidence” was analyzed in preparation for the IPCC Assessment Report 3 (AR3), published the following year in 2001. If this uncertainty framework can be considered adequate for the IPCC reports over a period of more than 20 years – when the stakes are so high, and clear communication with laypersons and policymakers couldn’t be more important – then mightn’t it also work adequately for astrobiologists eager to adopt an exactly similar continuum, but now for confidence of (extraterrestrial) life detection?



**Fig. 5:** The IPCC uncertainty language framework: The X-axis corresponds to the degree of robustness of the evidence for a claim (assessed by a working group), and the Y-axis represents the extent of agreement of the scientific community regarding the same claim. From Mastrandrea et al. (2010), with permission.

Here, we do not intend a full analysis of pros and cons of the IPCC uncertainty framework. Over the past 20+ years – including four IPCC reports AR3, 4, 5, and 6 – it has been heavily scrutinized (Budescu et al., 2012, 2014; Kause et al., 2022; Mach et al., 2017; M. D. Mastrandrea et al., 2011; M. D. Mastrandrea & Mach, 2011; Molina & Abadal, 2021; Rehg & Staley, 2017). Whatever your view, in the face of extraordinary scrutiny and the highest possible stakes, this uncertainty assessment framework has stood the test of time.

One possible objection might stem from the thought that, whilst the IPCC uncertainty framework is acceptable for IPCC statements, when it comes to confidence of life detection, we would prefer to somehow add *detail*, for example concerning the extent to which known abiotic mimics have been ruled out. This is precisely what Green et al. (2021) attempted, of course, with the CoLD framework. However, it is precisely such details that lead to criticism and community disagreement. As NASEM (2022b, p. 1) note in their review of Meadows et al. (2022), “The diversity of opinion from the broader community indicates that universal adoption of any one framework or scale would be challenging.” They later refer to “strong opinions” and “lack of community consensus” (p. 5). Thus, if we want to bring the astrobiology community together – as the earth sciences community has come together vis-à-vis uncertainty assessment in IPCC reports – we may want to avoid “overly structured frameworks” (NASEM, 2022b, p. 18). We agree that “Any proposed framework should be kept as general as possible” (NASEM, 2022b, p. 19). In other words, the lack of structure and detail in the IPCC framework should be considered a strength, not a weakness.

Are we offering here yet *another* framework to compete with both the CoLD scale and the NfoLD/NExSS framework, thus potentially further fracturing the community? We don’t think so. The NfoLD/NExSS framework still has its place, since it was never seriously put forward as a way to “read off” confidence of life detection. Instead, it serves a different function: that of describing/prescribing good practice scientific methodology when a new potential biosignature is put forward. Thus, we tend to agree with NASEM (2022b, p. 19), when they write of the NfoLD/NExSS framework: “It could also be argued that approaching all five of these proposed questions in any order is just the execution of the standard scientific process rather than a new assessment framework.”

It may be asked: how would the astrobiology community's adoption of the IPCC uncertainty framework actually work, in practice? This can't be the place for a full answer to this question, but we can make some initial suggestions as follows. Imagine that a committee (not unlike an IPCC "Working Group") has been put in place to assess the uncertainty vis-à-vis a particular "biosignature" that has been put forward. Making use of the IPCC framework, that committee is required to think not only about the "first-order" scientific evidence for the claim under scrutiny (X-axis of Fig. 3), but also the "second-order" contribution coming from the extent of community agreement or "consensus" (Y-axis of Fig. 3). This ensures that the committee considers where the *community* stands on the matter. This is worthwhile, since there may be cases where, although the evidence seems robust to a particular team of scientists, for some reason significant disagreement within the community remains. As Mastrandrea et al. (2011, p. 678) put it, "The degree of agreement is a measure of the consensus across the scientific community on a given topic and not just across an author team." This constraint limits the circumstances in which a team can say they are "confident." It prevents them from being too confident too soon since it will take some time for a community opinion to take shape. If we can't say anything concrete about 'agreement,' we can't (yet) say anything about 'confidence,' whatever we might think of the (first-order) evidence.<sup>[14]</sup>

What about possible abiotic mimics? How do they get factored into the uncertainty assessment? Consider again the particular case of a  $\text{CH}_4 - \text{O}_2$  disequilibrium in the atmosphere of an exoplanet. To assess the degree of confidence, our hypothetical committee must consider (i) the first-order scientific evidence and (ii) the extent of agreement in the community of experts. When it comes to (i), a judgement needs to be made concerning the extent to which available evidence links a  $\text{CH}_4 - \text{O}_2$  disequilibrium with life. The consideration of possible unconceived abiotic mimics is considered implicitly within this judgement. Whilst some scientists might jump too quickly from "There is no known abiotic explanation" to "It is highly likely to be

caused by life,” others will be more cautious, and this expresses itself in the measure of *agreement* – a lack of consensus. The final judgement concerning our overall degree of confidence comes from combining (i) and (ii) and considering where we land on the “Confidence” chart.

To see more fully how it would work in practice, consider the case of Nutman et al. (2016) and Allwood et al. (2018) on the proposed 3.7-billion-year-old stromatolite. Imagine our committee, shortly after the publication of Nutman et al. (2016), attempting to fill in “evidence” and “agreement” to reach a degree of confidence. Even if they considered the evidence strong, they wouldn’t be able to give any score for “agreement,” since at that time there hadn’t yet been the kind of peer scrutiny required to establish the extent of agreement. Filling in this score requires waiting for scrutiny to take place. And, once that scrutiny *did* take place, obvious community *disagreement* was expressed by way of Allwood et al. (2018). This case demonstrates how the IPCC uncertainty framework will sometimes deliver an undefined confidence score, since in the early stages of a new research program we won’t yet have any value for “scientific agreement.” We see this as a virtue of using the IPCC uncertainty framework – when asked “How sure are we that this is extraterrestrial life?” we should indeed sometimes reply “It’s too early to say.” Thus, adoption of the IPCC framework requires scientific community activity over a significant period of time. This is just as it should be, we contend, since sensible scores for overall confidence need to be determined by a process that washes out individual biases, perspectives, and blind spots that any individual scientist, or any individual team of scientists, may have.<sup>[15]</sup>

These are early thoughts on the adoption of the IPCC uncertainty framework within the field of astrobiology, and with application to “confidence of life detection” in particular. A key question remains how, exactly, the IPCC framework is supposed to factor in the possibility of

unconceived abiotic mimics. Application to a case such as the proposed 3.7-billion-year-old stromatolites is easy, since an abiotic explanation was so easy to find (within two years), as soon as the wider scientific community turned its attention to the case. In terms of the confidence scale, we move from “undefined” to “low confidence” over the course of 2016-2019 (not *very* low confidence, given the response: Nutman *et al.* (2019)).

There are important questions regarding how the framework might apply to other historical examples, such as that of ALH84001, which is notable both for the complexity of the evidence involved and the slow accumulation of comprehensive abiotic explanations. In the years following the publication McKay *et al.* (1996) announcing the possible evidence of life on Mars, the community reception was mixed. As Dick (2020, p. 705) put it, “the stakes were high and the skeptics numerous.” This suggests that the “agreement” score would never have gotten beyond “low”. Similarly, the various lines of evidence were individually and collectively ambiguous, and it’s hard to see how the “evidence” score could get beyond “limited” (see Fig. 5). Thus, in this case we propose that the overall confidence score would have started out low, and then over the course of 1996-2022 would have gotten gradually lower in a series of steps coinciding with important publications such as Golden *et al.* (2001), Martel *et al.* (2012), and Steele *et al.* (2022). The final resting place for “confidence” is extremely low.

What would the framework say about a much harder case, such as an oxygen-methane disequilibrium biosignature? It must be accepted that many in the scientific community would be tempted by the move from “No currently-conceived plausible abiotic explanation” to “It’s probably life.” But a scientific community is heterogenous – some scientists are more cautious than others<sup>[16]</sup> – and some proportion of the community would no doubt be more open to possible, currently unconceived, abiotic explanations. We propose that there is no better way to measure the appropriate degree of uncertainty vis-à-vis possible abiotic mimics than to distil our assessment from the cut-and-thrust of scholarly debate – at conferences and in the

professional literature – in the years following the detection of the disequilibrium. A committee tasked with determining a confidence score would then hope to factor in the possibility of abiotic mimics via consideration of the degree of community disagreement, following substantial community debate (which may or may not take several years, depending, for one thing, on the overall measure of community attention afforded to the issue).

Despite the argument from community heterogeneity, a worry may still linger that the scientific community would be skewed towards over-confidence, since the inference from “No currently conceived abiotic explanation” to “It’s probably life” is tempting, and scientists aren’t used to factoring into their evidential assessments cautionary tales from the history of science, such as those sketched above, in Section 2. A partial solution to this would be to increase the interaction of scientists with science scholars (philosophers, historians, sociologists, anthropologists), who would offer such perspectives. Whilst cross-disciplinary interactions in the broad field of astrobiology do take place to a certain extent (e.g., at AbSciCon), much more could be done to enrich interdisciplinary learning and dialogue (see Denning and Dick (2019), especially Section 5.0 ‘Recommendations’).

## 6 Conclusion

This paper makes three positive proposals. First and foremost, that the astrobiology community, searching for a way to assess biosignature uncertainty, would profit from adopting the time-tested IPCC uncertainty framework, and that this would help to address challenges faced by the assessment frameworks found in the works of Green et al. (2021) and Meadows et al. (2022). The IPCC framework can be adopted *alongside* another framework such as that found in the work of Meadows et al. (2022), with the latter interpreted as describing, or perhaps prescribing, good practice scientific methodology (key questions to be asked and answered)



when a potential biosignature is put forward. Second, that if we want to assess biosignature uncertainty more rigorously than is currently possible, we need much more data on the topic of abiotic mimics (in the spirit of Rouillard et al. (2018)). This is the only way we will be able to say something meaningful concerning what National Academies of Sciences, Engineering, and Medicine (2022a) call “the reliability of a biosignature” (Q11.2). Third, we argue that astrobiologists who make biosignature judgements often need to factor in considerations that are best described as historical, philosophical, or sociological. This adds to calls for greater support (e.g., funding) for meaningful interdisciplinary engagement in the field (Denning and Dick (2019), Section 5.0).

## **Acknowledgements**

An early version of his paper was presented at the Isaac Newton Institute for Mathematical Sciences, in Cambridge, UK, in February 2022; huge thanks to delegates for helpful feedback, and to Peter Coveney for the invitation. Thanks also to EURiCA project reading group participants for helpful discussion, especially Catriona Sellick, Martin Ward, and Chris Greenwell. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

## **Authors' Contributions**

All authors contributed equally to the research; P.V. wrote the article.

## **Funding Information**

This work is part of the EURiCA project: ‘Exploring Uncertainty and Risk in Contemporary Astrobiology’, funded by a Leverhulme Trust Research Project Grant (RPG-2021-274).

## Author Disclosure Statement

The authors declare no conflict of interest.

## References

- Allwood, A. C., Rosing, M. T., Flannery, D. T., Hurowitz, J. A., & Heirweh, C. M. (2018). Reassessing evidence of life in 3,700-million-year-old rocks of Greenland. *Nature*, *563*(7730), Article 7730. <https://doi.org/10.1038/s41586-018-0610-4>
- Almár, I., & Race, M. S. (2011). Discovery of extra-terrestrial life: Assessment by scales of its importance and associated risks. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *369*(1936), 679–692. <https://doi.org/10.1098/rsta.2010.0227>
- Budescu, D. V., Por, H.-H., & Broomell, S. B. (2012). Effective communication of uncertainty in the IPCC reports. *Climatic Change*, *113*(2), 181–200. <https://doi.org/10.1007/s10584-011-0330-3>
- Budescu, D. V., Por, H.-H., Broomell, S. B., & Smithson, M. (2014). The interpretation of IPCC probabilistic statements around the world. *Nature Climate Change*, *4*(6), Article 6. <https://doi.org/10.1038/nclimate2194>
- Cady, S. L., Farmer, J. D., Grotzinger, J. P., Schopf, J. W., & Steele, A. (2003). Morphological Biosignatures and the Search for Life on Mars. *Astrobiology*, *3*(2), 351–368. <https://doi.org/10.1089/153110703769016442>

- Capece, M. C., Clark, E., Saleh, J. K., Halford, D., Heinl, N., Hoskins, S., & Rothschild, L. J. (2013). Polyextremophiles and the Constraints for Terrestrial Habitability. In J. Seckbach, A. Oren, & H. Stan-Lotter (Eds.), *Polyextremophiles: Life Under Multiple Forms of Stress* (pp. 3–59). Springer Netherlands. [https://doi.org/10.1007/978-94-007-6488-0\\_1](https://doi.org/10.1007/978-94-007-6488-0_1)
- Catling, D. C., Krissansen-Totton, J., Kiang, N. Y., Crisp, D., Robinson, T. D., DasSarma, S., Rushby, A. J., Del Genio, A., Bains, W., & Domagal-Goldman, S. (2018). Exoplanet Biosignatures: A Framework for Their Assessment. *Astrobiology*, 18(6), 709–738. <https://doi.org/10.1089/ast.2017.1737>
- Cowie, C. (2022). Arguing About Extraterrestrial Intelligence: Philosophical Lessons from the ‘Oumuamua Controversy’. *Philosophical Quarterly*, Forthcoming.
- Dawid, R., Hartmann, S., & Sprenger, J. (2015). The No Alternatives Argument. *The British Journal for the Philosophy of Science*, 66(1), 213–234.
- Denning, K., & Dick, S. (2019). Preparing for the Discovery of Life Beyond Earth. *Astro2020: Decadal Survey on Astronomy and Astrophysics, APC White Papers, No. 183, Bulletin of the American Astronomical Society, Vol. 51*(Issue 7, id. 183).
- des Marais, D. J., Allamandola, L. J., Benner, S. A., Boss, A. P., Deamer, D., Falkowski, P. G., Farmer, J. D., Hedges, S. B., Jakosky, B. M., Knoll, A. H., Liskowsky, D. R., Meadows, V. S., Meyer, M. A., Pilcher, C. B., Nealson, K. H., Spormann, A. M., Trent, J. D., Turner, W. W., Woolf, N. J., & Yorke, H. W. (2003). *The NASA Astrobiology Roadmap*. 3(2), 219–235.
- Dick, S. J. (2020). Lessons Learned from the Twentieth-Century Extraterrestrial Life Debate. In S. J. Dick (Ed.), *Space, Time, and Aliens: Collected Works on Cosmos and Culture* (pp. 695–721). Springer International Publishing. [https://doi.org/10.1007/978-3-030-41614-0\\_39](https://doi.org/10.1007/978-3-030-41614-0_39)
- Golden, D. C., Ming, D. W., Schwandt, C. S., Lauer, H. V., Jr., Socki, R. A., Morris, R. V., Lofgren, G. E., & McKay, G. A. (2001). A simple inorganic process for formation of carbonates, magnetite, and sulfides in Martian meteorite ALH84001. *American Mineralogist*, 86(3), 370–375. <https://doi.org/10.2138/am-2001-2-321>

- Green, J., Hoehler, T., Neveu, M., Domagal-Goldman, S., Scalice, D., & Voytek, M. (2021). Call for a framework for reporting evidence for life beyond Earth. *Nature*, *598*(7882), 575–579. <https://doi.org/10.1038/s41586-021-03804-9>
- Harrison, J. P., Gheeraert, N., Tsigelnitskiy, D., & Cockell, C. S. (2013). The limits for life under multiple extremes. *Trends in Microbiology*, *21*(4), 204–212. <https://doi.org/10.1016/j.tim.2013.01.006>
- Hoegh-Guldberg, O., Jacob, D., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K., Engelbrecht, F., Guiot, J., Hijjoka, Y., Mehrotra, S., Payne, A., Seneviratne, S. I., Thomas, A., Warren, R., Zhou, G., Halim, S. A., Achlatis, M., ... Zougmore, R. B. (2018). *Impacts of 1.5°C Global Warming on Natural and Human Systems*. <https://helda.helsinki.fi/handle/10138/311749>
- Kause, A., Bruine de Bruin, W., Persson, J., Thorén, H., Olsson, L., Wallin, A., Dessai, S., & Varemán, N. (2022). Confidence levels and likelihood terms in IPCC reports: A survey of experts from different scientific disciplines. *Climatic Change*, *173*(1), 2. <https://doi.org/10.1007/s10584-022-03382-3>
- Krissansen-Totton, J., Olson, S., & Catling, D. C. (2018). Disequilibrium biosignatures over Earth history and implications for detecting exoplanet life. *Science Advances*, *4*(1), eaao5747. <https://doi.org/10.1126/sciadv.aao5747>
- Mach, K. J., Mastrandrea, M. D., Freeman, P. T., & Field, C. B. (2017). Unleashing expert judgment in assessment. *Global Environmental Change*, *44*, 1–14. <https://doi.org/10.1016/j.gloenvcha.2017.02.005>
- Martel, J., Young, D., Peng, H.-H., Wu, C.-Y., & Young, J. D. (2012). Biomimetic Properties of Minerals and the Search for Life in the Martian Meteorite ALH84001. *Annual Review of Earth and Planetary Sciences*, *40*(1), 167–193. <https://doi.org/10.1146/annurev-earth-042711-105401>

- Mastrandrea, M. D., & Mach, K. J. (2011). Treatment of uncertainties in IPCC Assessment Reports: Past approaches and considerations for the Fifth Assessment Report. *Climatic Change*, *108*(4), 659. <https://doi.org/10.1007/s10584-011-0177-7>
- Mastrandrea, M. D., Mach, K. J., Plattner, G.-K., Edenhofer, O., Stocker, T. F., Field, C. B., Ebi, K. L., & Matschoss, P. R. (2011). The IPCC AR5 guidance note on consistent treatment of uncertainties: A common approach across the working groups. *Climatic Change*, *108*(4), 675–691. <https://doi.org/10.1007/s10584-011-0178-6>
- Mastrandrea, M., Field, C. B., Stocker, T. F., Edenhofer, O., Ebi, K. L., Frame, D. J., Held, H., Kriegler, E., Mach, K. J., Matschoss, P. R., Plattner, G.-K., Zwiers, F. W., & Matschoss, P. R. (2010). *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties* (p. 7). IPCC.
- McKay, D. S., Gibson, E. K., Thomas-Keprta, K. L., Vali, H., Romanek, C. S., Clemett, S. J., Chillier, X. D. F., Macchling, C. R., & Zare, R. N. (1996). Search for Past Life on Mars: Possible Relic Biogenic Activity in Martian Meteorite ALH84001. *Science*, *273*(5277), 924–930. <https://doi.org/10.1126/science.273.5277.924>
- McMahon, S., & Cosmidis, J. (2022). False biosignatures on Mars: Anticipating ambiguity. *Journal of the Geological Society*, *179*(2), jgs2021-050. <https://doi.org/10.1144/jgs2021-050>
- Meadows, Graham, H., Abrahamsson, V., Adam, Z., Amador-French, E., Arney, G., Barge, L., Barlow, E., Berea, A., Bose, M., Bower, D., Chan, M., Cleaves, J., Corpolongo, A., Currie, M., Domagal-Goldman, S., Dong, C., Eigenbrode, J., Enright, A., ... Young, L. (2022). *Community Report from the Biosignatures Standards of Evidence Workshop*. <http://arxiv.org/abs/2210.14293>
- Meadows, V. S., Reinhard, C. T., Arney, G. N., Parenteau, M. N., Schwieterman, E. W., Domagal-Goldman, S. D., Lincowski, A. P., Stapelfeldt, K. R., Rauer, H., DasSarma, S., Hegde, S., Narita, N., Deitrick, R., Lustig-Yaeger, J., Lyons, T. W., Siegler, N., & Grenfell, J. L. (2018). Exoplanet Biosignatures: Understanding Oxygen as a Biosignature in the Context of Its Environment. *Astrobiology*, *18*(6), 630–662. <https://doi.org/10.1089/ast.2017.1727>

- Molina, T., & Abadal, E. (2021). The Evolution of Communicating the Uncertainty of Climate Change to Policymakers: A Study of IPCC Synthesis Reports. *Sustainability*, *13*(5), Article 5. <https://doi.org/10.3390/su13052466>
- Moss, R. H., & Schneider, S. H. (2000). Recommendations to lead authors for more consistent assessment and reporting. In R. Pachauri, T. Taniguchi, & K. Tanaka (Eds.), *Guidance Papers on the Cross Cutting Issues of the Third Assessment Report of the IPCC* (p. 33:51). World Meteorological Organization.
- NASEM. (2022a). *Origins, Worlds, and Life: A Decadal Strategy for Planetary Science and Astrobiology 2023-2032*. The National Academies Press. <https://doi.org/10.17226/26522>.
- NASEM. (2022b). *Independent Review of the Community Report from the Biosignature Standards of Evidence Workshop: Report Series—Committee on Astrobiology and Planetary Sciences*. The National Academies Press. <https://doi.org/10.17226/26621>
- Nutman, A. P., Bennett, V. C., Friend, C. R. L., Van Kranendonk, M. J., & Chivas, A. R. (2016). Rapid emergence of life shown by discovery of 3,700-million-year-old microbial structures. *Nature*, *537*(7621), Article 7621. <https://doi.org/10.1038/nature19355>
- Nutman, A. P., Bennett, V. C., Friend, C. R. L., Van Kranendonk, M. J., Rothacker, L., & Chivas, A. R. (2019). Cross-examining Earth's oldest stromatolites: Seeing through the effects of heterogeneous deformation, metamorphism and metasomatism affecting Isua (Greenland) ~3700 Ma sedimentary rocks. *Precambrian Research*, *331*, 105347. <https://doi.org/10.1016/j.precamres.2019.105347>
- Rea, D. G., O'Leary, B. T., & Sinton, W. M. (1965). Mars: The Origin of the 3.58- and 3.69-Micron Minima in the Infrared Spectra. *Science*, *147*(3663), 1286–1288. <https://doi.org/10.1126/science.147.3663.1286>
- Rehg, W., & Staley, K. (2017). “Agreement” in the IPCC Confidence measure. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *57*, 126–134. <https://doi.org/10.1016/j.shpsb.2016.10.008>

- Rouillard, J., García-Ruiz, J.-M., Gong, J., & van Zuilen, M. A. (2018). A morphogram for silica-witherite biomorphs and its application to microfossil identification in the early earth rock record. *Geobiology*, *16*(3), 279–296. <https://doi.org/10.1111/gbi.12278>
- Schwieterman, E. W., Kiang, N. Y., Parenteau, M. N., Harman, C. E., DasSarma, S., Fisher, T. M., Arney, G. N., Hartnett, H. E., Reinhard, C. T., Olson, S. L., Meadows, V. S., Cockell, C. S., Walker, S. I., Grenfell, J. L., Hegde, S., Rugheimer, S., Hu, R., & Lyons, T. W. (2018). Exoplanet Biosignatures: A Review of Remotely Detectable Signs of Life. *Astrobiology*, *18*(6), 663–708. <https://doi.org/10.1089/ast.2017.1729>
- Shirk, J. S., Haseltine, W. A., & Pimentel, G. C. (1965). Sinton Bands: Evidence for Deuterated Water on Mars. *Science*, *147*(3653), 48–49. <https://doi.org/10.1126/science.147.3653.48>
- Simoncini, E., Virgo, N., & Kleidon, A. (2013). Quantifying drivers of chemical disequilibrium: Theory and application to methane in the Earth's atmosphere. *Earth System Dynamics*, *4*(2), 317–331. <https://doi.org/10.5194/esd-4-317-2013>
- Sinton, W. M. (1957). Spectroscopic Evidence for Vegetation on Mars. *The Astrophysical Journal*, *126*, 231. <https://doi.org/10.1086/146393>
- Sinton, W. M. (1959). Further Evidence of Vegetation on Mars: The presence of large organic molecules is indicated by recent infrared-spectroscopic tests. *Science*, *130*(3384), 1234–1237. <https://doi.org/10.1126/science.130.3384.1234>
- Stanford, P. K. (2006). *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. Oxford University Press. <https://doi.org/10.1093/0195174089.001.0001>
- Stanford, P. K. (2011). Damn the Consequences: Projective Evidence and the Heterogeneity of Scientific Confirmation. *Philosophy of Science*, *78*(5), 887–899. <https://doi.org/10.1086/662283>
- Steele, A., Benning, L. G., Wirth, R., Schreiber, A., Araki, T., McCubbin, F. M., Fries, M. D., Nittler, L. R., Wang, J., Hallis, L. J., Conrad, P. G., Conley, C., Vitale, S., O'Brien, A. C., Riggi, V., & Rogers, K. (2022). Organic synthesis associated with serpentinization and carbonation on early Mars. *Science*. <https://doi.org/10.1126/science.abg7905>

Thompson, M. A., Krissansen-Totton, J., Wogan, N., Telus, M., & Fortney, J. J. (2022). The case and context for atmospheric methane as an exoplanet biosignature. *Proceedings of the National Academy of Sciences*, 119(14), e2117933119. <https://doi.org/10.1073/pnas.2117933119>

Vickers, P. (2022). *Identifying Future-Proof Science* (Oxford University Press).

Walker, S. I., Bains, W., Cronin, L., DasSarma, S., Danielache, S., Domagal-Goldman, S., Kacar, B., Kiang, N. Y., Lenardic, A., Reinhard, C. T., Moore, W., Schwieterman, E. W., Shkolnik, E. L., & Smith, H. B. (2018). Exoplanet Biosignatures: Future Directions. *Astrobiology*, 18(6), 779–824. <https://doi.org/10.1089/ast.2017.1738>

1

---

<sup>[1]</sup> In this paper, for simplicity, we assume that the phenomenon in question has been detected with certainty. The reader may suppose, for example, that the detection of oxygen has been double and triple checked by several teams using a range of techniques and instruments. Thus the case of ‘phosphine on Venus’ is less relevant here.

<sup>[2]</sup> Another rich and interesting example is the Schopf-Brasier dispute, and resultant literature. In this case Schopf and Packer (1987) announced bacteria fossils in a 3.5 billion year old rock: “among the oldest fossils yet described” (p. 72). Braiser et al. (2002) disputed this. To date, this has led to three serious ‘unconceived alternative’ abiotic explanations for the phenomena in question (see García-Ruiz et al. 2003; Wacey et al. 2016).

<sup>[3]</sup> Green et al. (2021), Meadows et al. (2022), and NASEM (2022b) all mention the problem of false positives and/or abiotic mimics, but without distinguishing the two separate challenges.

<sup>[4]</sup> Earlier discussion of these issues can be found in Cady et al. (2003).

<sup>[5]</sup> This is taken from NASA Sagan Fellow Joshua Krissansen-Totton, from a SETI talk: <https://youtu.be/iNSFKLb8oAY> (see 31mins in). But it is a sentiment widely found in the literature.

<sup>[6]</sup> One could always claim that there might remain unconceived alternatives. But sometimes this is clearly unreasonable, and the matter is beyond reasonable doubt, something even Stanford accepts: “Often enough . . . we are rightly confident in our ability to have exhausted the space of likely or plausible explanations” (Stanford, 2006, p. 31). These issues are discussed in detail in Vickers (2022), Chapter 2, Section 2.

<sup>[7]</sup> “It has been stressed previously (Green et al., 2021) that there is a continuum from a low to higher certainty level in whether or not a given phenomenon is due to life” (Meadows et al., 2022, p. 59). See also NASEM (2022b, p. 1), which refers to a “continuum of confidence”.



<sup>[8]</sup> Green et al. (2021) had already stated that “much remains to be established [including] whether a one-dimensional progressive scale is the right approach” (p. 578).

<sup>[9]</sup> E.g. Walker et al., (2018), especially Section 3: ‘Detecting Unknown Biology on Unknown Worlds: A Bayesian Framework’; Catling et al. (2018), especially Section 2: ‘A Bayesian Framework for Biosignature Assessment and Life Detection’.

<sup>[10]</sup> *Obscurum per obscurius*: an explanation that is obscurer than the thing to be explained.

<sup>[11]</sup> Also highly relevant to this point is Capece et al. (2013).

<sup>[12]</sup> See Rouillard et al. (2018) for an interesting example of an attempt to go through a kind of phase space looking for abiotic mimics.

<sup>[13]</sup> Meadows et al. (2022) include a very brief section entitled ‘Lessons learned from other fields’ (pp. 58-9), where they mention (but do not discuss) the IPCC uncertainty language framework.

<sup>[14]</sup> It is essential to include the qualification ‘first-order’ here, since strength of community agreement *is* an important type of evidence in science, albeit second-order evidence. See e.g. Parker (2020) and Vickers (2022).

<sup>[15]</sup> Here we align with Almár and Race (2011), who write that claims are “reviewed over time by general consensus of the scientific community” (p. 688f.). However, Almár and Race (2011) lump together two things we definitely want to keep separate: the *scientific importance* of a finding, and the *credibility* of the claim being made (p. 685).

<sup>[16]</sup> See e.g. Section 3 of Chapter 5 of Vickers (2022).